


Keith K. Niall
Editor

Vision and Displays for Military and Security Applications

**The Advanced Deployable Day/Night
Simulation Project**

 **Springer**

Vision and Displays for Military and Security Applications

Keith K. Niall
Editor

Vision and Displays for Military and Security Applications

The Advanced Deployable Day/Night
Simulation Project



Springer

Editor

Keith K. Niall
Defence Research & Development
Embassy of Canada
Washington DC
USA
Keith.Niall@forces.gc.ca

ISBN 978-1-4419-1722-5 e-ISBN 978-1-4419-1723-2

DOI 10.1007/978-1-4419-1723-2

Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2009943541

© Springer Science+Business Media, LLC 2010 All rights reserved.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface



Realistic and immersive simulations of land, sea, and sky are requisite to the military use of visual simulation for mission planning. Until recently, the simulation of natural environments has been limited first of all by the pixel resolution of visual displays. Visual simulation of those natural environments has also been limited by the scarcity of detailed and accurate physical descriptions of them. Our aim has been to change all that. To this end, many of us have labored in adjacent fields of psychology, engineering, human factors, and computer science. Our efforts in these areas were occasioned by a single question: how distantly can fast-jet pilots discern the aspect angle of an opposing aircraft, in visual simulation? This question needs some elaboration: it concerns fast jets, because those simulations involve the representation of high speeds over wide swaths of landscape. It concerns pilots, since they begin their careers with above-average acuity of vision, as a population. And it concerns aspect angle, which is as much as to say that the three-dimensional orientation of an opposing aircraft relative to one's own, as revealed by motion and solid form.

The single question is by no means simple. It demands a criterion for eye-limiting resolution in simulation. That notion is a central one to our study, though much abused in general discussion. The question at hand, as it was posed in the 1990s, has been accompanied by others. Questions of the visibility of vehicles and features of the landscape (air-to-ground issues) have taken precedence over questions of the speed and direction of other aircraft (air-to-air issues). Other questions have arisen about the visibility and comprehensibility of landscapes seen by new spectra (issues of night vision and infrared imagery) as well as questions of the geographic fidelity of simulation in synoptic view (issues of photogrammetry and mapping). All this will enable us to see better in simulation. (The answer to the original question, by the way, is more than a couple of nautical miles, but the answer will be elaborated elsewhere. Cf. the chapters by Covas et al., and Gaska et al., on this theme.)

The present volume began as the proceedings of a two-day symposium hosted by Defence Research & Development Canada (the Advanced Deployable Day/Night Symposium, November 13th and 14th, 2007, in Toronto). The background for the work lies in a long-standing collaboration between Defence Research & Development Canada in Toronto, and the Arizona facility of the US Air Force Research Laboratory. We document a number of new technologies, and also methods for their evaluation. One may ask the benefit of developing advanced technologies, when often governments can procure simulation and display technologies “off the shelf”. We have gathered many interested parties from industry and academe, as can be told from the many affiliations of the authors in this volume. Through our efforts, we trust that we have succeeded in nudging North American industry in a new direction. Several of our concerns – high-resolution projection and night-vision simulation serve as examples – were once considered “niche markets”. That phrase was meant shorthand for technologies as yet unprofitable for exploitation. It’s not a phrase that would aptly be applied to these areas any longer. We did predict that the course of industrial development would fill some gaps in technology, and generally we have been rewarded, with progress in distributed computing for simulation, and with the development of small and reliable solid state lasers for light sources. This volume is a condensed compilation of our own efforts, and their unified representation. Alternate technologies do exist in healthy competition with these, of course, and we remain more attentive than dogmatic about such new possibilities.

Prototype production is unlike engineering design is unlike experimental psychology. What an art it is, then, to balance these things. It would be an exaggeration – it could be called a pragmatic exaggeration – to claim they can be complementary in the business of simulation. Balancing them has been one of the pleasures of this enterprise, which does not end with the present volume. Instead there are many directions open for the future. Here are some favorites: the development of a four-colour projection system, the experimental evaluation of sensory substitution for image fusion, and the development of full atmospheric for night vision simulation. The first idea is to extend traditional projector systems, which depend on three lights such as red, green, and blue to develop a full gamut of display colours in simulation. The addition of a fourth channel in the near-infrared would enable the use of night vision devices in a simulator – given careful attention

to the simulation of environmental reflectance in the near infrared. Currently there are several display systems that are being developed along these lines. The second idea is to use sensory channels to carry new scenes of the same multiplicity or complexity, where a display is meant to represent the appearance of non-visible parts of the spectrum, as well as the visible scene. The use of haptic devices, and the fused representation of scenes in the true thermal range, are examples that have been put forward, and which could well be elaborated in future. The third idea is that variations in appearance due to clouds and moisture are even more important to night vision and infrared sensing than such variations are within the ordinarily visible range. The development of full models of these variations, based on physical principles, is then proportionally important to the realism of night vision simulation. Such an effort may sound fairly abstract, until you want to find your way about when the weather has turned. The latter theme is developed in chapters by Clark et al. and by Zhu, Church and Labrie in this volume. As a philosopher once prefaced his book: may others come along and improve on this.

Washington, DC

Keith K. Niall

Contents

1	Creating Day and Night: Past, Present, and Future	1
	Mary K. Kaiser	
2	Development of a DVI-Compatible VGA Projector Engine Based on Flexible Reflective Analog Modulators	13
	Francis Picard, François Duchesne, Michel Jacob, Carl Larouche, Carl Vachon, and Keith K. Niall	
3	Brightness and Contrast of Images with Laser-Based Video Projectors	27
	Michel Doucet, Mélanie Leclerc, Francis Picard, and Keith K. Niall	
4	Physics Based Simulation of Light Sources	49
	Jeff Clark, Brad Colbert, Karl Mathia, and Brett Chladny	
5	Integration of a Deployable CIGI-Based Image Generator in an Existing Simulation	61
	Robert Wolfe and Tony Ghoman	
6	Advances in Scalable Generic Image Generator Technology for the Advanced Deployable Day/Night Simulation Project	75
	Ray M. Obidowski and Rajesh Jha	
7	Detection Threshold of Visual Displacement in a Networked Flight Simulator	87
	Christine M. Covas, James P. Gaska, George A. Geri, Megan E. Shamp, and Byron J. Pierce	
8	Evaluation of the Spatial and Temporal Resolution of Digital Projectors for use in Full-Field Flight Simulation	97
	James P. Gaska, George A. Geri, Marc D. Winterbottom, and Byron J. Pierce	

9 A Spatial Cognition Paradigm to Assess the Impact of Night Vision Goggles on Way-Finding Performance 111
Michelle Gauthier and Avi Parush

10 Psychophysics of Night Vision Device Halos..... 123
Robert S. Allison, Tracey Brandwood, Margarita Vinnikov,
James E. Zacher, Sion Jennings, Todd Macuda, Paul Thomas,
and Stephen A. Palmisano

11 Effects of Screen Resolution and Training Variation on a Simulated Flight Control Task..... 141
Don C. Donderi, Keith K. Niall, Karyn Fish, and Benjamin Goldstein

12 Video-to-Reference Image Indexing..... 163
Vitaly Zholudev and Richard Wildes

13 AVS LIDAR for Detecting Obstacles Inside Aerosol 177
Xiang Zhu, Philip Church, and M. Labrie

Author Biographies 189

Glossary 193

References..... 203

Index..... 209



Chapter 1

Creating Day and Night: Past, Present, and Future

Mary K. Kaiser

Abstract The capability to produce compelling renderings of daylight and nocturnal environments has been a challenge both for flight simulation and for film and related entertainment venues. In this chapter, I compare and contrast the technologies and techniques that these communities have employed during the last half-century. In particular, I will examine the varying criteria (aesthetics-based, performance-based, information-based) employed to evaluate the “goodness” of the rendering. I will also discuss the future challenges awaiting the simulation community, as humans migrate to space environments whose lighting characteristics are decidedly non-terrestrial, and where the transition between “day” and “night” occurs, not at dusk, but at the edge of a boulder or crevasse.

I was delighted to be invited to deliver the keynote address at the Advanced Deployable Day/Night Simulation (ADDNS) Symposium because I knew the other speakers would present fascinating talks on state-of-the-art physics-based night vision simulation, ultra-high resolution display technologies, advances in registering video imagery with geographic databases, refined tools and techniques to measure performance in operational scenarios, and the ADDNS system itself. I realized that I could add relatively little to the technical discussion of the latest advances in these domains. Nonetheless, I hoped that I could play the role of Janus, the two-headed Greek god who looked to the past as well as the future, and ensure that the symposium members considered how the history of day/night simulation might inform our future endeavors. In this chapter, I will summarize my historical overview and the insights I believe we can glean from the road already traveled.

M.K. Kaiser (✉)

Human Factors Research & Technology Division, NASA Ames Research Center,
Mail Stop 262-2, Moffett Federal Airfield, California, 94035-1000, USA
e-mail: mkaiser@mail.arc.nasa.gov; mary.k.kaiser@nasa.gov

I began my review of the prior art with a quotation from the first chapter of Genesis: “And God said, Let there be light: and there was light. And God saw... that it was good. And God divided the light from the darkness. And God called the light Day, and the darkness He called Night.” I hoped the audience was amused by the reference (and not offended – readers should feel free to substitute their preferred creation account). But moreover, I hoped the audience would recognize that even this allegorical account touches on several important aspects of creating/simulating visual worlds: the quality of the resulting product; the feasibility and “share-ability” of the simulation methodology; and the criteria used to evaluate the quality of the simulation. With these three touchstones in hand (product, methodology, and evaluation criteria), we embarked on an overview of “Day and Night Simulation in Five Acts.”

1.1 Act I: Day and Night

Ultimately, the simulation community would like its technologies to support both daylight and night-time conditions with equal fidelity. Historically, systems have not possessed such flexible capabilities. One might suggest, tongue in check, that the Link General Aviation Trainer (the infamous “Blue Box” used to train over a half-million pilots and aviators during and after World War II) was equally adept at simulating day and night; but that is simply because, as an instrument-conditions trainer, it simulated contact displays for neither (Fig. 1.1a).

It wasn’t until the 1990s that simulators could meet the visual systems criteria for FAA Level D approval (i.e., “Daylight, dusk, and night visual scenes with sufficient scene content to recognize airport, the terrain, and major landmarks around the airport and to successfully complete a visual landing.” (Federal Aviation Administration 1991)). By mid-decade, 54 such systems were operational (Wentzel 1995; for an enlightening comparison, see CAT’s 2007–2008 Flight Simulator Census of over 1,100 full-flight simulators: CAT 2007). It is worth noting that computer graphics technologies were able to meet the FAA’s dusk/night visual system requirements (“...4,000 light points for night and dusk scenes”) years before they were capable of achieving the daylight benchmark (full color presentation, scene content comparable in detail to that produced by 4,000 edges or 1,000 surfaces); calligraphic (vector) devices were capable of generating the requisite point lights, but the edges/surfaces required for daylight scenes required later-generation raster technologies.

Current simulator visual systems, such as Flight Safety International’s VITAL VIIIi (Fig. 1.1b), greatly exceed these graphics generation requirements. Interestingly, the FAA did not heighten these criteria when they last revised its Advisory Circular on Airplane Simulator Qualification (AC Number 120-40C); this revision has been in draft form since 1995 (Federal Aviation Administration 1995). Thus, despite the lack of a regulatory driver, simulators have advanced their day/night scene-generation capabilities, incorporating advances in computer graphics and display technologies.

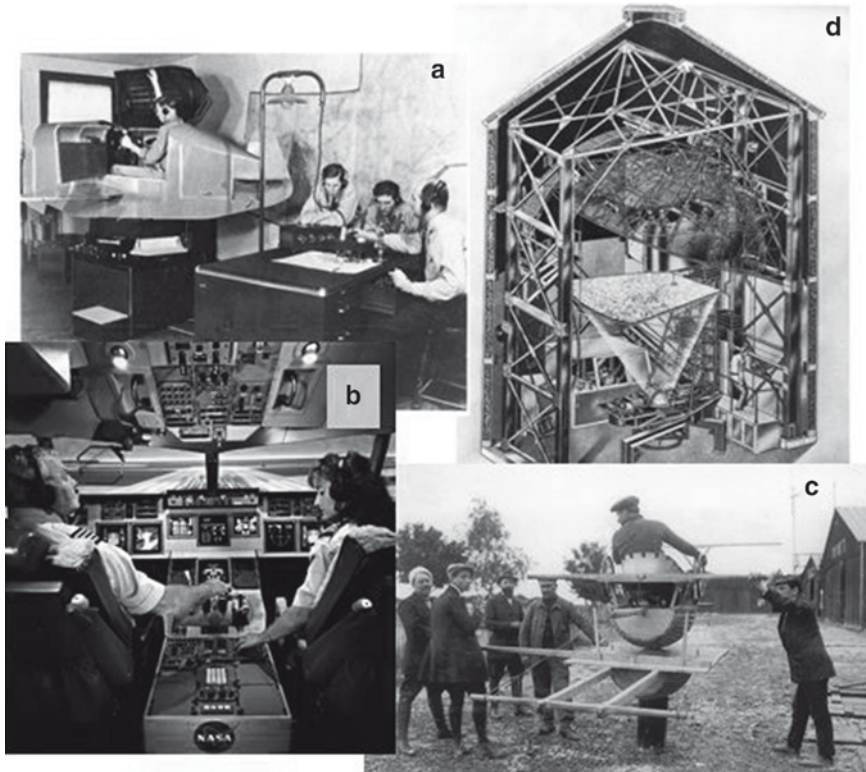


Fig. 1.1 A collage of flight simulators, past and present. Counter-clockwise from the upper left: (a) the Link General Aviation Trainer; (b) the B747-400 simulator at NASA Ames' Crew-Vehicle Systems Research Facility; (c) the Antoinette Trainer; and (d) Link's Celestial Navigation Trainer

1.2 Act II: Day and Night

Before possessing the capability and flexibility that computer graphics systems afford, the simulation community developed elegant (and not so elegant) approaches to visual scene generation. Typically, these visual scenes were limited in scope (e.g., providing horizon displays to aid with attitude-control training) and specific to daylight or night-time conditions.

The value of simulation-based training was recognized virtually at the dawn of aviation. In 1910, an article in *Flight International* stated, "The invention, therefore, of a device which will enable the novice to obtain a clear conception of the workings of the control of an aeroplane, and of the conditions existent in the air, without any risk personally or otherwise, is to be welcomed without a doubt." (quoted in Wentzel 1995). And, in fact, innovative aviators were putting these words to deed. The Sanders Teacher and the Antoinette Trainer (a.k.a. the "apprenticeship barrel" – Fig. 1.1c)

were both introduced that same year, each providing the student pilot with rudimentary “hands-on” exposure to the mapping between control input and change in aircraft attitude.

Not all simulators focused on pilotage skills. To aid WW II bomber crews, Link developed the Celestial Navigation Trainer (Fig. 1.1d). This elegant (if not especially compact) trainer recreated the night sky with 379 point-light “stars” embedded in a dome that could rotate and translate to reflect changes in time and aircraft position. Further, a screen mounted under the fuselage could display terrain projections for training on recognizing navigational landmarks and bombing targets (Kilgore 1989).

The Celestial Navigation Trainer’s creative use of divergent display technologies (i.e., both a “planetarium” dome and film projectors) reminds us that the simulation community often draws upon advances in other disciplines to advance our own. Nowadays, the most obvious contributions for visual systems are drawn from the computer graphics and displays community. Still, it is informative to examine other domains that have influenced past simulation technologies. In our next act, we will focus on techniques, tricks, and tweaks gleaned from cinematography and theme-park rides.

1.3 Act III: Day for Night

Filming motion pictures or television shows at night is logistically difficult and expensive. Consequently, cinematographers developed what has been termed “Day for Night” (DFN) filming techniques (also called “La nuit américaine” or American Night, because the early developers were American filmmakers working on limited-budget productions).

Simply stated, DFN is the process of filming in the daytime a scene that occurs at night. The effect of night-time is accomplished both by the pattern of lighting and by the use of special lenses (typically blue tint) and/or underexposure levels (typically 2–3 f-stops) to darken the scene and alter its hue. There are additional heuristics (e.g., shoot under tree-branch shadows to create the illusion of moonlight, avoid excessive amounts of sky in the frame), but the technique involves more art than science, and the resulting quality remain unknown until the film is developed.

In fact, experts in the field often disagree about the effectiveness of the application. Some practitioners praise David Lean’s *Lawrence of Arabia* (1962) as a paragon of DFN shooting (e.g., www.digitallyobsessed.com raves that the film’s “day for night scenes look realistic and possess outstanding shadow detail”); others contend it is one of the worst examples, at least among films with production budgets greater than my yearly civil servant salary (as Mark Smith comments on www.cinematography.net, “*Lawrence of Arabia* has some DFN which perhaps asks us to suspend a little too much disbelief”). In fact, the unusual nature of the scenery

being “simulated” (the craggy desert landscape under full-moon illumination – an environment often described as “otherworldly”) makes it difficult for the average viewer to render a valid fidelity judgment.

Actual day-for-night filming is gradually being displaced with computer-based post-production processing techniques that achieve similar effects. The processing software affords several advantages, including allowing the filmmaker to iteratively adjust the vDFN effects and to apply different adjustments to different areas of the frame (e.g., greater darkening to the sky). The “cost” of migrating to a post-production methodology is the lost of real-time implementation; for cinematographers, this cost is zero – DFN is virtually never applied to live broadcasts. For the simulation community, however, real-time implementation is a requirement – virtually all of our “broadcasts” are “live.”¹

The simulation community also shares technological challenges (and solutions) with the amusement park industry. One prominent example of this is the increased use of motion platforms, initially developed for flight simulation applications, in theme park rides. Park managers recognize that motion platforms provide compelling vestibular cues with a minimum cost of valuable staging space. When properly integrated with a visual display, the resulting immersive experience is stunning (e.g., the “Soarin’ over California” at Disney’s California Adventure theme park). Further, park managers (like simulator users) appreciate the flexibility motion platforms afford. Just as NASA Ames’ Vertical Motion Simulator can shift its simulation mode from Space Shuttle to rotorcraft with a swap-out of cabs and platform reprogramming, Paramount’s Great American was able to “re-theme” its motion-based Action FX Theater from a “Days of Thunder” car-racing experience to “Stan Lee’s 7th Portal 3D Simulation Experience” (the nature of which I’m at a loss to explain, even having experienced it myself).

Prior to motion platforms, theme rides employed a “transport the rider along a fixed track in a staged area” technique. These “dark rides” emerged in the early twentieth century with “Tunnel of Love” and “Adventure Train” rides, and likely reached their pinnacle of sophistication in the Disney parks. “Peter Pan’s Flight,” installed as an original attraction in 1955, remains one of the Anaheim Park’s most

¹A notable exception to real-time visual scene generation was Link’s Variable Anamorphic Motion Picture (VAMP). In this commercial aircraft simulator, full-color, 70 mm films were taken of nominal approach, landing, and taxi operations. The VAMP could simulate limited deviations from the recorded trajectories by utilizing two methods of image processing: (a) subsampling the area of the film frame to correspond to the aircraft’s pitch and heading, and (b) optically warping the image with anamorphic lenses to simulate the aircraft’s roll. While the advantages of such a visual system are apparent (e.g., high scene resolution and fidelity), the limitations are equally striking. Because only a limited number of filmed sequences were available, pilots became familiar with them and started using extraneous cues to aid their pilotage (e.g., using highway traffic to “time” their approach trajectories). Further, pilots learned to null the optical distortion to correct their flight path errors. Clearly, such strategies do not generalize to actual flight environments.

popular rides. Even today, Disney continues make major investments in “dark ride” attractions (e.g., the recently refurbished “Space Mountain” and the newly opened “Finding Nemo Submarine Voyage,” whose track just happens to be under water).

For over a decade starting in the late 1960s, the flight simulation community at NASA Ames embraced a visual-scene generation technology that borrowed both cinematic and dark ride techniques and tricks. Our Visual Flight Attachment (VFA) used intricate model boards and a mechanical tracking system to “fly” a video camera over terrains and airfields. The image was then transmitted to the pilot in the motion-based simulator cab. Figure 1.2 shows the VFA and the resulting image for a landing approach to the model board’s aircraft carrier.

Late in its service, the VFA was adapted for a series of simulated night studies by adding runway lighting to the model board. But just as cinematographers often discover during DFN shooting, unanticipated artifacts resulted. Most notably, the runway lights tended to “bloom” on the video image, creating false cues to distance (i.e., the runway appeared closer than it should during final approach). One of the VFA engineers (Hollow 1981) designed a dynamic lighting system that modulated the voltage fed to the lights as a function of distance from the camera; this innovation provided an effective mitigation against the blooming artifact. Clearly, Ames engineers can be just as clever as Disney’s Imagineers.

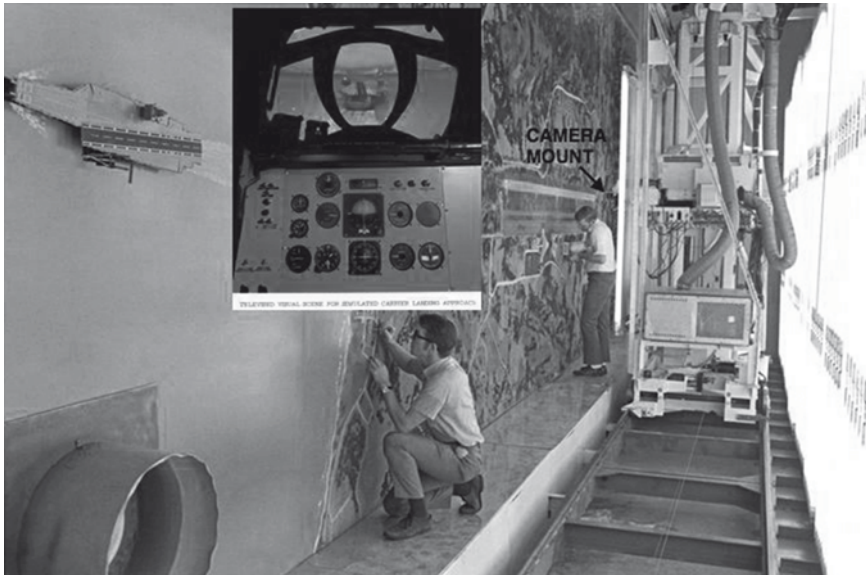


Fig. 1.2 NASA Ames’ Visual Flight Attachment system (ca. 1971). A video camera (initially black-and-white, later color) was mounted at the rail-based tracking system shown on the *right*. The camera was “flown” over the terrain model board via instructions from the simulator cab. The *inset* shows an example of the imagery displayed to the pilot, in this case for a landing approach to the aircraft carrier

1.4 Act IV: Night for Day

As challenging as simulating natural night conditions can be, even more daunting is simulating the visual imagery generated by night vision devices (NVDs), most notably image intensifying (I^2) and infrared (IR) sensors. The research presented by others at this symposium provides testament to the advances that have transpired in physics-based sensor-imagery simulation in the past decade, and provides much greater insight into the current state-of-the-art that I could ever hope. So what I shall focus on instead is a simulation technique from decades past, when physics-based models were simplistic (or non-existent) and tended to overwhelm our graphic systems' real-time processing capabilities. In this primitive research environment of the distant past (c. 1985), we turned in desperation to inverse video as a technique for IR simulation.

Inverse video (a.k.a. “reverse” video or “reversed-polarity” video) simulation method simply takes a monochrome video image and reverses its polarity (i.e., black pixels are rendered white, etc.). Because this is a computationally trivial transformation, even the most underpowered video board can handle a real-time feed. The resulting image, of course, portrays none of the scene's actual temperature or radiance properties; it is simply a “negative” of the visible light image. The question one must ask, then, isn't why inverse video doesn't work better as a simulation of IR imagery; the real question is: why does it work at all?

Consider the images of Comet Tempel 1 shown in Fig. 1.3 (as a NASA employee, I am required to include at least one space-related image per paper, gratuitous or not). The left frame shows a visible light image, the center frame shows the same image with reversed polarity, and the right frame shows the IR image. If we were to perform pixel-based analyses, the greatest (albeit negative) correlation would be between the left and middle images – a perfect -1.0 . The correlation of the right image with the middle one would be in the neighborhood of

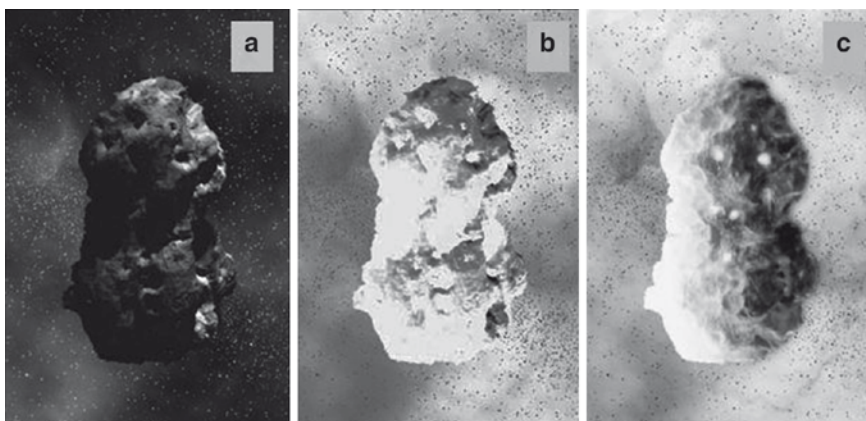


Fig. 1.3 Three views of Comet Tempel 1 provide a comparison of visible light (*left*), inverse video (*middle*), and infrared imagery (*right*)

0.6; there is an equal degree of correlation between the right image and the left, although (once again) this correlation is negative.

However, this mathematical analysis (at least, the magnitude of correlation) does not match our psychological impression. Perceptually, the middle image appears much more similar to the one on the right than on the left. And the right image appears similar to the middle image, but very different from the left image. Clearly, our perceptual criteria for similarity are different from our mathematical one. Likely, we are relatively insensitive to negative correlations (as anyone who's tried to match a photograph to its proper match among a pile of negatives can attest). Further, it seems that the IR and inverse-polarity images both disrupt our typical feature-recognition schema and force us to engage in alternative, strategy-driven image interpretation. At some level, both IR and inverse-polarity imagery share a "differentness"; this quality affords the inverse video simulation technique some limited training utility.

1.5 Act V: Day in Night

Actually, many additional NASA-supplied space images were used in this section of my presentation, but I'm not including them here because they suffer in the color-to-monochrome translation. Instead, I encourage readers to peruse the lovely Quicktime panoramas Hans Nyberg has assembled from imagery collected during the Apollo missions (<http://www.panoramas.dk/moon/mission-apollo.html>). One of the most striking aspects of these lunar landscapes is how "day" and "night" co-exist in a single scene. Due to the lack of atmospheric diffusion and attenuation, lighting is exceptionally harsh; shadows are utterly black. Lawrence of Arabia's moonlit desert crags may be called otherworldly, but the Moon's barren craters truly are: these ambient illumination conditions are truly alien to our terrestrially evolved visual system.

In order to perform tasks effectively in space environments, astronauts need to train in high-fidelity simulations of those environments. To date, only a few efforts have focused on simulating the visual conditions. Jim Maida's group at NASA's Johnson Space Center has developed lighting models to aid training for Shuttle extravehicular or EVA tasks (where, in addition to dealing with harsh illumination, astronauts must deal with 90-min day/night cycles). Maida and his colleagues have also applied their models to simulate what external cameras will capture during the Shuttle's launch and main tank separation (Maida et al. 2006), and have investigated whether augmented-reality displays can aid control of the Shuttle's remote manipulator arm when the ambient lighting degrades depth and distance cues (Maida et al. 2007).

NASA's current Vision for Space Exploration calls for an ongoing human presence on the Moon beginning in the 2020s. If we are going to live and work effectively in the lunar environment, it will be critical to ensure that we can simulate its visual characteristics for training and planning exercises. I invite my colleagues in the simulation community to engage their expertise toward this next challenging domain of visual scene simulation.

1.6 Simulation and Evaluation

Our five-act review has focused on two of the three touchstones of simulation: the methodologies used to create the simulation and the resulting product. We now turn to the third touchstone: evaluating the simulation quality. In my Genesis example, the resulting “day” and “night” were deemed “Good” by a single (albeit highly respected) evaluator. In the simulation community, even today, significant weight can be placed on the “goodness” judgment of a single esteemed evaluator, be it a test pilot, astronaut, or multi-starred general (none of whom, technically, is God or even godlike ... despite rumors to the contrary).

The Genesis evaluation process can also be criticized for allowing the designer – not a terribly objective party – to judge the product’s effectiveness. The designer’s lack of objectivity is not due simply (or only) to their personal investment in the project. In addition, their intimate knowledge of the simulation’s intended purpose can prevent them from evaluating the product with the “naïve” eye of the user. The Genesis example also points to the most critical aspect of evaluation: what is the function of the world we’ve created? In order to determine whether a simulation is of sufficient fidelity, we must first understand the function it must support.

For filmmakers, the world they create is there to support the story they wish to tell. Any cinematic technique that is used to simulate conditions in that world is acceptable so long as it does not overly tax the audience’s willing suspense of disbelief and thereby “yank them out of the story.” For theme-ride designers, the world is there to support the adventure of the ride. Generally, it is difficult to tell a complex story in the course of the ride (as both Stan Lee and the early designers of Disney’s “Pirates of the Caribbean” discovered); rather, the goal is to provide a global sense of experience (e.g., hang gliding in “Soarin’ over California” or rocket riding in “Space Mountain”) or to immerse the rider in the key scenes of a familiar tale (e.g., Disney’s “Pinocchio’s Daring Journey” and “Peter Pan’s Flight”). So long as the rider enjoys and “gets” the experience, the simulated world has achieved its purpose.²

For the training simulation community, the evaluation bar is set higher. Our goal is to ensure that the skills and strategies acquired in the simulated world will generalize. How do we evaluate the quality of the simulation to determine whether this goal is achieved? Similar to filmmakers and ride designers, our community utilizes subjective evaluations from expert raters. Such subjective ratings still play a role in FAA certification processes (Federal Aviation Administration 1991, 1995).

²As already suggested, designers are not always the best judge of whether the simulation is successful. Disney’s Imagineers originally designed “Peter Pan’s Flight” as a first-person theme ride; that is, the rider was Peter Pan. Audiences loved the ride, but would frequently ask, “Where was Peter? We never saw him!” Finally, park managers conceded that riders were not “getting” the first person perspective, and added an Audio-Animatronic Peter in 1983.

These expert evaluators are typically (and properly) users rather than designers, and they are asked to rate the degree of simulation between the simulation and the actual vehicle or system. On the positive side for our community, a clear standard (i.e., the real-world vehicle or system) can be provided; on the negative, it is still uncertain what criteria the evaluator employs in making a judgment of perceived similarity.

A second class of evaluation employed by the simulation community is based on measuring operational performance. If we can establish performance criteria (and acceptable levels thereof), we can access the similarity of operators' performance in the simulator to their performance with the vehicle or system. Also, we can evaluate the extent to which operational training generalizes from simulator to real-world performance. The limitation of this evaluation methodology is its cost (extensive real-world performance must be conducted and evaluated), and the fact that it can only be applied as a post-hoc evaluation: we need an existing simulator to evaluate.

Ideally, our community would like to be able to establish objective metrics of simulator quality and fidelity. Such metrics can then be applied both to the evaluation of existing systems and to the development of design requirements for future systems. I now consider two areas where such objective metrics could be established: image quality evaluation of physics-based modeling and simulation; and psychophysics-based evaluation of visual systems.

Following my talk at the symposium, Richard Wildes asked for examples where physics-based simulation is inadequate. In principle, a complete physics-based simulation would be fully adequate. But, of course, even with modern computational capabilities, it is not possible to employ complete physical models. Rather, we model those aspects that critically impact the visual simulation. So the question is not whether to start with physics-based modeling; the question is when one can stop the modeling process – when is the image that results from the model perceptually equivalent to the true image?

Assuming that we can assemble a set of true images that adequately samples the visual world, we can examine observers' ability to discriminate between the true and simulated images. Alternatively, we can apply a human vision model such as the Spatial Standard Observer (Watson and Ahumada 2004) to measure this discriminability without the need to perform the tedious psychophysical assessments.

Where psychophysics is needed is to develop the general visual system requirements – and the best psychophysics to use is active psychophysics. Sweet (2006) has developed a useful taxonomy that maps potential visual depth cues (e.g., convergence, binocular disparity, motion parallax, linear perspective, occlusion) to the relevant component of the visual system (i.e., the image generator, the display, or both), and determines whether or not veridical cues are provided. In the case of the display component, the quality of depth cueing can vary as a function of the display type (e.g., collimated screens do not provide proper motion parallax cues due to pilot head movements, but head-tracked stereo displays can).

The next step in the process is to determine, for a given operational task, what visual cues are required for successful execution of the inner-loop control task (see Fig. 1.4). One can then map the operational training needs into simulator design

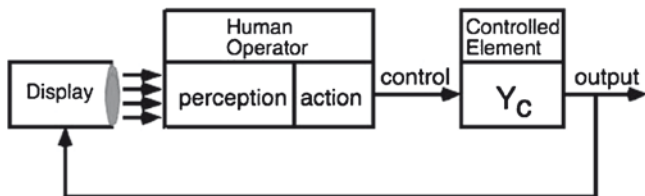


Fig. 1.4 An “active psychophysics” model of inner-loop control (adapted from Sweet and Kaiser 2006)

requirements, recognizing that these requirements may be task specific. For example, mid-air refueling will likely impose different visual requirements (e.g., near-field depth cues) than open-sky target acquisition (e.g., high resolution and wide fields of view). Given our current simulation technology, the goal is not so much to design The Perfect Simulator, but rather to enable a cost-effective match of requirements and design.

At some level, this task-specific approach to simulator design may seem to put us back into the barrel with Antoinette’s apprentices. I would argue, however, that the sophistication of our current simulation technologies allows us to build cost-effective systems that are, in fact, fully sufficient for wide classes of tasks. Still, we must recognize that even FAA-certified full-mission simulators have less-than-desired fidelity for some phases of flight (e.g., taxi and ramp operations).

I believe, therefore, that the best lesson to be learned from the history of flight simulation (and related domains) is one that was taught in Antoinette’s trainer: maintaining a proper balance. But the balance our community seeks is among cost, flexibility, and fidelity: how do we best exploit available technologies to design affordable simulators that can simulate the widest range of task scenarios at the requisite level of fidelity? It is a challenging undertaking, which is why I was pleased to be surrounded by such impressively able colleagues at the ADDNS symposium.



Chapter 2

Development of a DVI-Compatible VGA Projector Engine Based on Flexible Reflective Analog Modulators

Francis Picard, François Duchesne, Michel Jacob,
Carl Larouche, Carl Vachon, and Keith K. Niall

Abstract The development of a Digital Video Interface (DVI) – compatible VGA projector engine based on Flexible Reflective Analog Modulator (FRAM) is reported. The FRAM technology development began a few years ago in response to a need for a new projection technology allowing the achievement of ultrahigh resolution for high fidelity simulations. This technology relies on simple micromirrors produced using typical Micro Opto Electro Mechanical System (MOEMS) manufacturing processes. It has the advantages of offering a simple fabrication process (three masking layers), a quick response time ($5\ \mu\text{s}$) and to be wavelength insensitive over large spectral ranges. Additionally, the light modulation with these microdevices does not require the achievement of a very high quality optically flat state of the micromirrors which is typically difficult to obtain yet necessary for other MOEMS modulation technologies.

Testing and FRAM array selection for packaging: One challenge of the projector engine development was the packaging of the FRAM dies presenting a high number of input signals (480). For the packaging of such dies, the die-on-board approach was selected. 480×1 FRAM arrays appropriate for this packaging approach have been designed, fabricated and tested. Very large arrays (4000×1) were also produced within the fabrication run therefore confirming the feasibility of such arrays. Testing in itself represented a challenge. Actually, it was critical to select a 480×1 FRAM array with a pixel functionality percentage as high as possible to minimize the presence of dead rows in the projected image. Reaching this goal implied that the fabricated FRAM arrays should be carefully inspected and tested. In practice, testing all FRAMs of many 480×1 arrays before packaging is not straightforward. The

F. Picard (✉), F. Duchesne, M. Jacob, C. Larouche, and C. Vachon
INO, Québec (siège social), 2740, rue Einstein, Québec (Québec), Canada, G1P 4S4
e-mail: Francis.Picard@ino.ca

K.K. Niall
Defence Research and Development Canada, Embassy of Canada, 501 Pennsylvania Avenue NW,
Washington, DC, USA

difficulty here is that the time and cost required for such testing must remain reasonable. This requires many FRAMs to be activated and tested at once. Activation of many FRAMs (60) was achieved with a customized probe card combined to a probe station. The response of the activated FRAMs was observed with an optical test bench integrated to the probe station. This test bench was a small and slightly modified version of the optical system used in a projector. A line of light illuminated a number of FRAMs and the light reflected by the FRAMs passed through Schlieren optics to produce pixels in the image plane located at the exit of an optical relay. A camera and a CRT produced a magnified image of the pixels. This provided a convenient method for assessing the functionality of thousands of FRAMs relatively quickly by observing the pixels intensity changes as the voltage applied to the FRAMs was varied.

FRAM array packaging: One 480×1 FRAM array has been selected for final packaging using this method combined with complementary interferometric microscope measurements. A flex board/metallic package assembly has been used to implement the selected die-on-board packaging approach. Within the assembly, each FRAM is electrically connected to the flexible PCB using wire bonding. A housing placed over the FRAM array and clamped to the board provides an hermetic enclosure allowing operation of the FRAM array in an inert atmosphere of dry nitrogen. It also provides mechanical protection to the array. It is equipped with an optical window. Small closable inlet and outlet openings allow regeneration of the inert atmosphere if necessary.

Control electronics, software and final assembly: The link between the die-on-board packaging assembly and the control electronics is insured by standard high pin count connectors. This customized control electronics and the associated control software have been developed to allow the transfer of image data to the modulator array. The control electronics includes a DAC board, a FPGA board and a power supply board. The designed electronics and the associated software are compatible with the Digital Video Interface standard. This allows the projection of a variety of images including still images and animated sequences. With the resulting projector engine, images with VGA resolution can be displayed at a frame rate of 60 Hz. Part of the control electronics and the package assembly are mounted on positioning supports. These supports allow a precise optical alignment of the FRAM array. Full functionality of the engine has been verified by integrating it into an existing optical test bench and projecting still images and animations.

Sommaire Le développement d'un engin de projection de format VGA basé sur les Modulateurs Analogues Réflectifs Flexibles (FRAM) et compatible avec le standard DVI est rapporté. La technologie FRAM a été développée il y a quelques années. Cette technologie repose sur des micro-miroirs simples produits en utilisant des procédés typiques pour la fabrication des MOEMS. Elle a l'avantage d'offrir un procédé de production simple (3 niveaux de masquage), un temps de réponse rapide (moins de $5 \mu\text{s}$) et d'être insensible à la longueur d'onde d'opération sur de grandes plages spectrales. De plus, la modulation de la lumière avec ces micro-dispositifs

ne demande pas de réaliser des micro-miroirs présentant un état plan de très haute qualité optique difficile à obtenir mais nécessaire pour d'autres technologies de modulation basées sur les MOEMS.

Tests et sélection de barrettes de FRAM pour la mise en boîtier: Un défi du développement de l'engin de projection était la mise en boîtier de puces FRAM présentant un grand nombre de signaux d'entrée (480). Pour la mise en boîtier de telles puces, l'approche puce-sur-carte a été sélectionnée. Des barrettes de FRAM 480×1 appropriées pour cette approche de mise en boîtier ont été conçues, fabriquées et testées. De très grandes barrettes (4000×1) ont aussi été produites durant la période de fabrication ce qui a confirmé la faisabilité de telles barrettes. Les tests en eux-mêmes représentaient un défi. En effet, il était critique de sélectionner une barrette de FRAM 480×1 avec un pourcentage de fonctionnement des pixels aussi haut que possible. Atteindre ce but impliquait que les barrettes de FRAM fabriquées devaient être inspectées et testées avec soin. En pratique, tester tous les FRAM de plusieurs barrettes 480×1 avant la mise en boîtier n'est pas simple. La difficulté ici est que le temps et le coût requis pour ces tests doivent demeurer raisonnables. Ceci requiert que plusieurs FRAM soient activés et testés en même temps. L'activation de plusieurs FRAM (60) a été réalisée avec une carte à pointes conçue sur mesure combinée à une station de test. La réponse des FRAM activés était observée à l'aide d'un banc de test optique intégré à la station de test. Ce banc de test était une version réduite et légèrement modifiée du système optique utilisé dans un projecteur. Une ligne de lumière illuminait plusieurs FRAM et la lumière réfléchiée par les FRAM passait à travers une optique Schlieren pour produire des pixels dans le plan image localisé à la sortie d'un relai optique. Une caméra et un moniteur produisaient une image agrandie des pixels. Ceci fournissait une façon commode pour vérifier le fonctionnement de milliers de FRAM relativement vite en observant les changements d'intensité des pixels quand le voltage appliqué aux FRAM était varié.

Mise en boîtier d'une barrette de FRAM: Une barrette de FRAM 480×1 a été sélectionnée pour la mise en boîtier finale en utilisant cette méthode combinée à des mesures complémentaires réalisées avec un microscope interférométrique. Un assemblage carte flexible/boîtier métallique a été utilisé pour implémenter l'approche de mise en boîtier puce-sur-carte sélectionnée. A l'intérieur de l'assemblage, chaque FRAM est connecté électriquement à la carte de circuit imprimé flexible par microcâblage. Un boîtier placé par-dessus la barrette de FRAM et serré sur la carte fournit une enceinte hermétique permettant l'opération de la barrette de FRAM dans une atmosphère inerte d'azote sec. Il procure aussi une protection mécanique à la barrette. Il est équipé d'une fenêtre optique. De petites ouvertures d'entrée et sortie refermables permettent la régénération de l'atmosphère inerte si nécessaire.

Electronique, logiciel de contrôle et assemblage final: Le lien entre l'assemblage puce-sur-carte et l'électronique de contrôle est assuré par des connecteurs standards présentant un nombre élevé de broches de raccordement. Cette électronique de contrôle sur mesure et le logiciel de contrôle associé ont été développés pour permettre le transfert des données des images vers la barrette de modulateurs.

L'électronique de contrôle comprend une carte DAC, une carte FPGA et une carte d'alimentation. L'électronique conçue et le logiciel associé sont compatibles avec le standard Digital Vidéo Interface (DVI). Cela permet la projection d'une variété d'images incluant des images fixes et des séquences d'animation. Avec l'engin de projection résultant, des images de définition VGA peuvent être affichées à une cadence de 60 Hz. Une partie de l'électronique de contrôle et le boîtier sont montés sur des supports de positionnement. Ces supports permettent un alignement optique précis de la barrette de FRAM. Le plein fonctionnement de l'engin a été vérifié en l'intégrant à un banc de test existant et en projetant des images fixes et des animations.

Nowadays, simulations play a role of increasing importance as training tools. No doubt that this is due to the benefits achievable with this training method which are approaching those of field training yet at a lower cost and often at a lower risk for trainees. This situation is causing some pressure on the required simulation quality due to the trainers' and users' expectations. Flight simulators which are amongst the most complex training tools, in particular those developed for jet fighter pilots, do not escape these trends. Therefore, it is no surprise to observe important ongoing efforts for the improvement of all aspects of these simulators. As a crucial component of the flight simulators, spatial light modulators have received a special attention over the past few years resulting in impressive achievements in terms of resolution, contrast and overall image quality. This paper reports efforts in this field related to the Flexible Reflective Analog Modulator (FRAM) technology. In particular, the development of a Digital Video interface (DVI)-compatible VGA projector engine based on this technology will be described.

A FRAM based projector engine consists of a linear array of $25\mu\text{m} \times 25\mu\text{m}$ microbridges which act as flexible micromirrors. Electrostatic actuation allows the control on the curvature of each micromirror independently. When such array is illuminated with a laser and combined with Schlieren optics (van Raalte 1970; see Fig. 2.1), the micromirror curvature variations perform an analog modulation of light intensity. This produces a line of image pixels at the output of an optical relay, each pixel corresponding to one flexible micromirror. This approach allows pixel intensity modulation with switching times in the range of $5\mu\text{s}$. A complete 2-D image is obtained by using a scanning mechanism that displays each image line sequentially. Additionally, projection optics is used to tailor the final image.

The FRAM arrays are produced using simple fabrication processes typical of Micro Opto Electro Mechanical System (MOEMS) manufacturing. Only three photolithographic masks are required to perform the surface micromachining process developed for the FRAM fabrication. Materials common in the microelectronics industry are used in this process: silicon wafers as substrates, aluminum for the actuation electrodes, silicon nitride for electrical insulation, and polyimide for the sacrificial layer which, when removed, will produce the gap between the micromirror and the substrate and finally aluminum alloy for the micromirrors themselves.

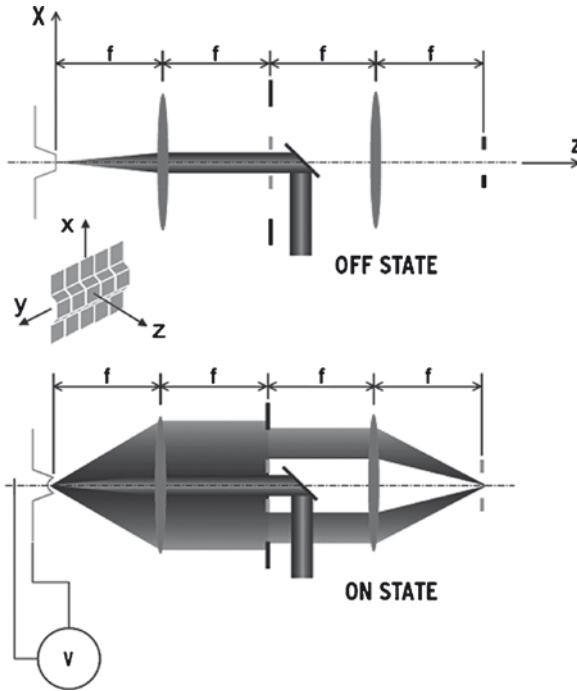


Fig. 2.1 Light modulation approach using a flexible reflective analog modulator (FRAM) array. The few FRAM shown between the OFF and ON state schematic representations indicate the array orientation in the optical setup

The FRAM are typically designed to achieve a maximum curvature corresponding to an f-number of 2. For $25\ \mu\text{m}$ micromirrors, this means producing micromirror sag of about $0.8\ \mu\text{m}$. Such sag is typically obtained for a voltage of less than 230 V for a micromirror thickness of 150 nm. As mentioned above, response times of the order of $5\ \mu\text{s}$ or less are possible with these micromirrors. Additionally, the FRAM are characterized by a high optical damage threshold of more than $8\ \text{kW}/\text{cm}^2$ continuous wave (CW) at 532 nm due to the high reflectivity of these micromirrors (higher than 88% at 532 nm) and to their strong thermal coupling with the substrate. This last feature allows the use of the FRAM arrays with high power laser sources which is not the case for all MEMS based optical modulators.

The development of the FRAM technology began in the year 2000 and since then has progressed to the current state. Following simulations and design, the first FRAM fabrication run was completed and characterized in 2002 (Picard et al. 2002). Projection display using the FRAM technology was demonstrated shortly after (Picard et al. 2003). These achievements have been followed by further studies aimed at improving the micromirror structure (for faster response time, inter alia) and allowing more flexible ways to operate the FRAM arrays for light modulation.

Results of these studies were reported in 2004 (Picard et al. 2004) and 2005 (Doucet et al. 2005). The most recent FRAM technology advances are reported here. They relate the development of a FRAM based DVI-compatible VGA projector engine.

2.1 FRAM Fabrication for VGA Projection Display

The development of a VGA projector engine (640 columns \times 480 lines displayed) involved the fabrication of 480 \times 1 FRAM arrays with individually addressable micromirrors. The high input count of these arrays led to the selection of a slightly modified chip-on-board approach for the packaging. The requirements related to this packaging approach were also important elements affecting the design of the 480 \times 1 micromirror arrays.

As mentioned above the FRAM array fabrication is based on a surface micromachining process involving a polyimide sacrificial layer. The first fabrication step consists of the deposition and patterning of an aluminum thin film on a silicon substrate covered with an electrically insulating silicon nitride layer. This step produces the common electrode located underneath all micromirrors. Second, the polyimide sacrificial layer is deposited and patterned. Of particular importance is the angle of the resulting polyimide structure walls which is a critical parameter affecting the operation of the FRAM. Third, an aluminum alloy film is deposited and patterned to produce the micromirrors. Finally, the sacrificial layer is removed using a plasma ashing step.

Three photomasks were needed for the fabrication of the required FRAM array. They were designed to be used with 6-in. silicon substrates. The mask layout included 288 linear arrays of 480 FRAM, a linear array of 4,000 FRAM, test structures and alignment marks. For the 480 \times 1 FRAM arrays, all micromirrors are individually addressable. For the 4,000 \times 1 FRAM array, 2,000 micromirrors are individually addressable and 2,000 micromirrors are connected in groups of five mirrors. This last array was included on the masks to investigate the feasibility of such large FRAM arrays.

The required 480 \times 1 FRAM arrays have been successfully fabricated using the process described above. Some of the micromirrors from one of the 480 \times 1 FRAM arrays are shown at the top of Fig. 2.2. 4,000 \times 1 FRAM arrays have also been produced (see Fig. 2.2) within this fabrication run. Static characterization of a few micromirrors from the 4,000 \times 1 FRAM arrays has been performed. Typical Deflection vs. Voltage curves (Picard et al. 2003) were generally observed. The voltage for the required maximum deflection (0.8 μ m) varied from 171 to 199 V. This demonstrates some pixel-to-pixel non uniformity. Non uniformity was even more evident in some cases. For these FRAM, atypical asymmetric deformations were observed. Such pixel-to-pixel non uniformity was not unexpected in the context of the first fabrication run involving these very large arrays.

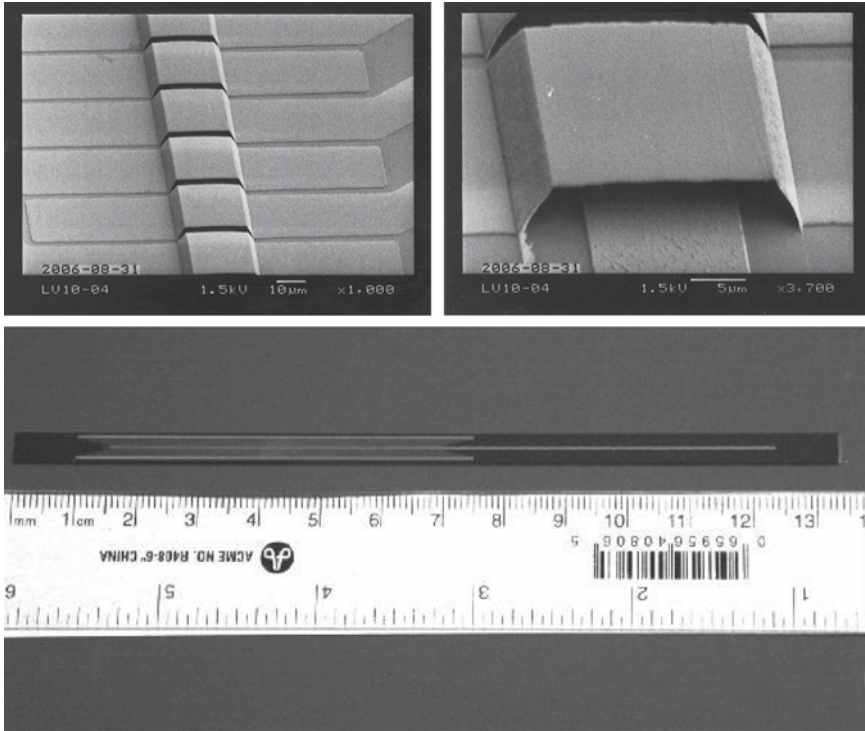


Fig. 2.2 Top FRAM arrays with a membrane thickness of 300 nm. Bottom 4,000×1 FRAM array

2.2 FRAM Array Testing and Selection

For the planned 640×480 pixel projector engine, it was critical to select a 480×1 FRAM array with a pixel functionality percentage as high as possible. This was required to minimize the presence of dead lines in the projected image. Reaching this goal implies that the fabricated FRAM arrays must be carefully inspected and tested. In practice, testing all FRAM of many 480×1 arrays before packaging is not straightforward. The difficulty here is that the time and cost required for such testing must remain reasonable.

2.2.1 Test Setup

The requirement of testing thousands of FRAM implies in practice that many micromirrors from an array must be actuated and tested at once. This required a setup especially designed for massive evaluation of FRAM functionality. Parallel actuation of 60 FRAM has been achieved by using a customized probe card combined

with a probe station (Cascade Microtech, REL-4800). The response of the activated FRAM was observed with an optical test bench integrated to the probe station. This test bench is a small and slightly modified version of the optical system used for light modulation in a FRAM-based projector. A line of light illuminates about 100 FRAM. The light reflected by the FRAM passes through Schlieren optics to produce pixels in the image plane located at the exit of an optical relay. A CCD camera connected to a computer produces a magnified image of the pixels. The operator of the test setup verifies visually that the pixel intensity profile changes normally as the voltage applied to the FRAM is varied. Once a set of 60 FRAM is checked, the optical setup is removed, the probe card is positioned to actuate the next 60 FRAM and the optical setup is put back in place to validate the response of this new group of micromirrors. The process is repeated until a full array has been tested. This approach provides a convenient way to observe thousands of FRAM relatively quickly and assess qualitatively their functionality.

2.2.2 FRAM Array Screening Procedure

The screening process for the FRAM arrays comprised an initial selection based on a visual inspection with a microscope. Then, the retained arrays went through a partial testing where the Deflection vs. Voltage curve was measured for a few FRAMs on each selected array. The Deflection vs. Voltage curve was obtained using an interferometric microscope. This step allowed further screening of the arrays. The remaining arrays were finally tested on the probe station with the setup described above (Sect. 3.1) to assess, at least qualitatively, the percentage of functional FRAM for a given array.

The developed procedure allowed the testing of FRAM arrays to an extent never approached before. This provided new insights about FRAM behaviour, in particular about their failure mechanism. For standard FRAM (that is, fabricated with the same membrane thickness as previously, i.e. 150 nm) it was relatively easy to find, by visual inspection, arrays with no defective mirror. The same arrays showed a typical Deflection vs. Voltage mirror response when characterized with the interferometric microscope. Ten mirrors were characterized with the interferometric microscope for each array. The maximum voltage to be applied to all mirrors with the probe station setup was based on these measurements. Relatively large variations of the voltage required to reach maximum deflection (0.8 μm) were observed among the characterized mirrors (a few tens of volts). The maximum voltage to apply to the array was set as the minimum voltage observed for maximum deflection within the ten measurements performed minus 15 V to account for the possibility of mirrors with even lower actuation voltages. When this voltage was applied to the whole array (60 micromirrors at once), FRAM were sometimes observed to collapse just beside mirrors which withstood the applied voltage.

The failure in these cases is not due to the sacrificial layer thickness non-uniformity or to the aluminum alloy film thickness non-uniformity. As was observed in previous

fabrication runs, the side walls of the polyimide sacrificial layer are not smooth but present a network of fibre-like structures generally oriented perpendicularly to the FRAM array. This pattern is replicated in the aluminum alloy deposited as the structural layer for the FRAM. These fibre-like structures are somewhat different from one micromirror to another. They are most probably due to the interaction of the oxygen with the silane groups within the polyimide during the RIE etching which produces silicon oxide residues. The resulting wall pattern is unique to each mirror and produces a unique stress distribution within the walls. This unique wall structure can explain either a shift of the Deflection vs. Voltage curve toward low voltages or defective walls with a reduced overall structural strength. Therefore, the stress concentration within the walls can lead to the destruction of one mirror upon actuation while the adjacent mirror withstands the same actuation voltage. Either the pull-in phenomena typical of electrostatic microdevices (Hung and Senturia 1999) or simple structural breakdown can be at the origin of the failure.

This does not necessarily exclude the use of such mirrors for projection display. If the failure mechanism is related to the pull-in phenomena, implementing a pixel-to-pixel non-uniformity correction procedure would solve the problem. If the failure mechanism is rather a structural breakdown occurring before the pull-in voltage is reached then it would be difficult to find arrays offering 100% functionality. In either case, pre-packaging testing with the probe station is not easily applicable and would most probably lead to the destruction of some FRAM within the tested array.

FRAM with thicker walls (300 nm), such as those shown in Fig. 2.2, were also fabricated. These micromirror arrays showed a better mirror-to-mirror Deflection vs. Voltage response uniformity. As expected, the actuation voltage was higher (around 260 V) than for the standard FRAM discussed above. However, the influence of the fibre-like structures in the walls was much lower, as demonstrated by the improved FRAM uniformity. For these arrays, it was possible to select an actuation voltage corresponding to a deflection close to the required maximum for most FRAM. One of these arrays was selected for final packaging after which testing with the probe station setup confirmed a functionality of 100% for this array.

As mentioned above, these arrays show a higher actuation voltage than the standard FRAM. However, they still offer a large useful deflection range but require a higher bias on the common electrode of the array for proper operation. This is compatible with the electronic driver designed for the DVI-compatible VGA projector engine and described in the following section.

2.3 Projector Engine Control

Control software has been developed for the image data transfer from a computer to the projector engine electronics. This Microsoft Windows application allows the loading and transfer of fixed images (.bmp format) and animations (.tmi format).

The image data are transferred to the projector engine through a DVI link. A built-in look-up table (LUT) defined by the user allows individual modification of the Pixel Intensity vs. Grey Level Input relation for each FRAM. This feature allows correction of the pixel-to-pixel non-uniformity when relevant data are available. Alternatively, this LUT can be replaced by a 14-bit linear LUT for which the voltage applied to the FRAM is proportional to the grey level data for all FRAM of the array. Additionally, contrast, brightness and gamma correction can be performed globally on the images before the user-defined LUT (or the 14-bit linear LUT) is applied to the image data. Other features of this software application allow the user to manually adjust frame synchronization and the number of displayed columns.

A built-in tool has also been added to the control software. This tool allows the quick and easy generation of user-defined test images. The tool can be opened as a new window from the main software window. In this window, the user can compose a program of simple instructions. When this program is run, the corresponding image is generated and displayed in the same window. This image can subsequently be saved or saved and loaded directly into the main software application. The programming script is relatively simple but still supports typical features such as basic arithmetical operations, loops, logical operations and conditions.

The Windows application is run on a host PC that feeds the control electronics with a monochrome video stream of size 640×480 pixels at the rate of 60 frames/s. The Windows application requires that a standard DVI interface board be installed in the host PC in order to send the video stream to the electronics through a DVI cable. The control electronics acts as a slave with respect to the host PC. It receives the video stream and forwards the image flow to the 480 micromirrors of the FRAM array. No information is transferred back from the electronics to the host PC. These control electronics consist of an FPGA board, a high-voltage digital-to-analog converter (HV DAC) board and a high-voltage DC power supply board.

The main components located on the FPGA board are a DVI input port for the video stream received from the host PC, a double-buffer of random access memory (RAM), and a CameraLink port for output of the video stream to the HV DAC board. These components are controlled by a VHDL application programmed in a Xilinx Virtex-2 Pro FPGA.

Image data are received as 640×480 pixels/frame, 60 frames/s. The pixel data are 14-bit binary words that are forwarded to the DAC board. Images are received in the standard 640×480 VGA format from the DVI port installed on the host PC. Since the FRAM arrays comprise 480 micromirrors, images must first be stored line-wise as 640×480 in a buffer memory, and then read column-wise as 480×640 during the display process. Such procedure requires a double-buffer structure. A steady video stream between input and output is achieved by writing the first incoming frame to a first buffer, then writing the second incoming frame to a second buffer while the preceding frame is being read simultaneously from the first buffer and so on. This procedure avoids interruptions in the data flow toward the DAC board and the FRAM array.

As its main functions, the VHDL application installed in the FPGA locks to the incoming video stream, manages the proper line-wise/column-wise addressing scheme for read/write accesses in the double-buffer memory structure and issues

frame synchronization to the DAC board. Images are sent from the FPGA board to the DAC board through a CameraLink port. Sending a full image frame nominally takes $1/60$ of a second. This period is partitioned as $13 \frac{2}{3}$ ms for writing the frame in the digital-to-analog converters and 3 ms to allow an eventual scanning device to return to its initial position before it can display a new frame. All micromirrors are driven as blanked lines (0 V applied) during this return interval.

The CameraLink port provides 24 independent data channels and each channel is used as a serial data link to access the serial input port of each of 24 multi-DAC chips installed on the DAC board. Each multi-DAC chip contains 32 individual 14-bit DACs, but the application timings permit using only 20 of the 32 available DACs on each chip. Since the VHDL application in the FPGA accesses the 24 channels in parallel, writing a single 480-pixel column requires 20 consecutive write operations to the 24-bit Camera Link port and this procedure must be repeated 640 times for a complete 480×640 frame to be written in the DAC. The number of 480-pixel columns that are written into the DAC is defined by the user and can vary from 480 to 640. If desired, this allows the adjustment of the time available for the HV DAC outputs to reach the required voltage values and stabilize.

The Digital-to-Analog Converter Board comprises 480 individual 14-bits DAC grouped into 24 AD5535 Analog Devices chips assembled on the printed circuit board. A common voltage reference signal is connected to each AD5535 chip. This signal allows the setting of the maximum output voltage from 50 to 200 V. The so called FRAM PC board further discussed below and comprising the packaged FRAM array is connected directly to the DAC board. Since images are displayed column-by-column, the DAC board also provides a signal available to drive a galvanometer which is the scanning device typically used to produce a 2-D image. This driving signal is a periodic triangular waveform that can be synchronized with the frame rate of the incoming video stream. In addition, the rising time, falling time and amplitude of the triangular waveform can be set with potentiometers as part of the galvanometer driver circuit.

2.4 480×1 FRAM Array Packaging

The packaging of microdevices exhibiting hundreds of I/O pads is generally a challenge. Typically no off-the-shelf packages are available to fulfill all requirements for such packaging tasks. The FRAM arrays used to produce the VGA projector engine, with their 481 bonding pads to be connected, are no exception. Therefore, a custom package and packaging procedure have been developed to link all micromirrors to the control electronics. The selected packaging approach is a modified die-on-board packaging method. In this approach, the FRAM array silicon die is glued to a metallic package base plate with a low outgas rate epoxy. This assembly is then secured to the FRAM PC board with screws. The FRAM PC board is a high density two-layer printed circuit. Its central region is rigid and comprises all bonding pads for the FRAM connection. The resolution of the metal traces in this region is very high, close to the current limit of the technology.

A hole in the center of this region allows the positioning of the FRAM die in close proximity with the PC board bonding pads by passing it through the board, the base plate of the package remaining at the back of the FRAM PC board. Both external sides of the FRAM PC board are also rigid. This is where high pin count standard connectors are mounted. These connectors provide the link between the FRAM PC board and the DAC board. Finally, the part of the FRAM PC board linking the connectors to the central region is flexible and exhibits a high density connection routing. A large hole and slits in this flexible part of the board increase the compliance of the mechanical link between the connectors and the FRAM array. At this point of the packaging process, the package base plate being attached to the FRAM PC board, wire bonding is performed between the FRAM array bonding pads and the FRAM PC board pads.

In the next step of the packaging procedure, a fused silica window with AR coatings is placed in a groove machined in the package cover and secured with a window holder and screws. Silicone is then injected in a groove surrounding the window through holes in the window holder. This assembly provides a hermetic joint between the window and the package cover and can be reworked if required. This is an important advantage for such an experimental packaging approach.

The package cover is then attached to the package base plate. In this process, both package parts are pressed against the central portion of the FRAM PC board which is in between. Silicone o-rings placed in grooves machined in the cover and the base plate ensure a hermetic assembly. At this point, the package enclosure must be purged with dry nitrogen. This step allows the operation of the FRAM array in an inert atmosphere which is generally required for Micro-electrical Mechanical Systems (MEMS) devices. Two holes in the package base plate allow this purge. A small tube is inserted in one of these holes and a low flow of dry nitrogen is circulated in the package for 24 h. After this period, the two holes are plugged with flat headed screws equipped with small o-rings which hermetically close the package cavity. All the purging operations described above are performed in a glove box purged with dry nitrogen to insure that no contamination of the package atmosphere occurs during the final closure of the purge holes. The purging manipulation can be repeated to regenerate the package atmosphere after a given operation period, if required. Four holes located at the outer margin of the package allow its attachment to positioning stages or supports, if needed. The assembly obtained with the packaging procedure described above is shown in Fig. 2.3.

2.5 Projector Engine Integration

All components of the projector engine have been assembled within an existing test bench. This test bench provides the required optical setup to produce images with the fabricated projector engine. In its current state, this test bench can illuminate up to 250 pixels which translates to 640 columns \times 250 lines images on the screen. The resulting projection device is shown in Fig. 2.4. Part of the positioning stages is

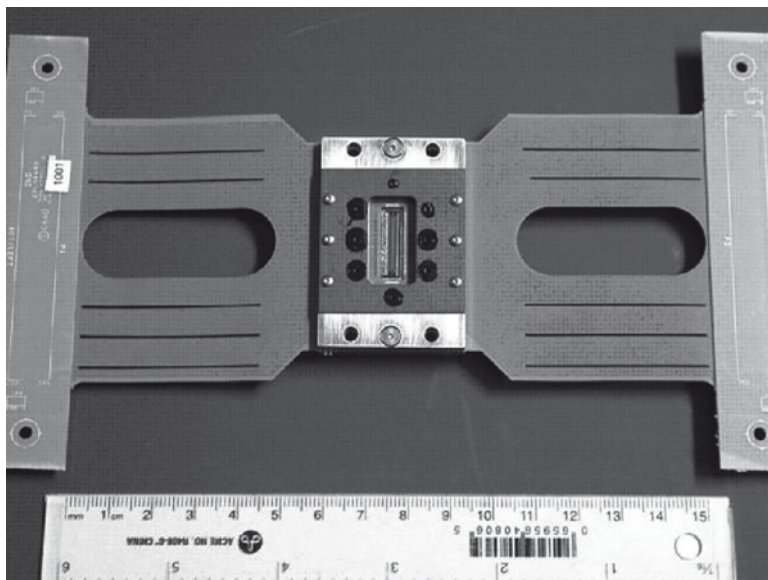


Fig. 2.3 Packaged 480×1 FRAM array

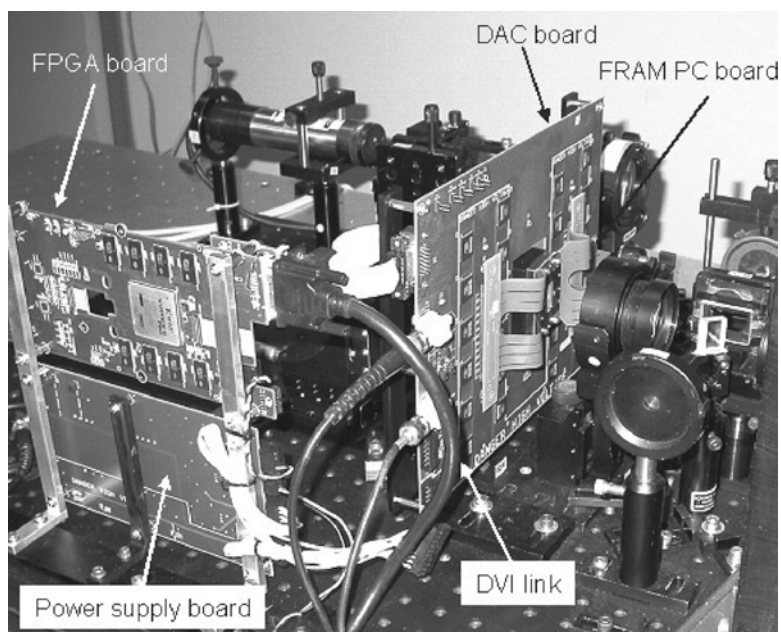


Fig. 2.4 The projector engine and part of the optical test bench

attached to the DAC board and allows coarse positioning of the DAC board/ FRAM PC board assembly. Part of the positioning stages is attached directly to the FRAM array package through a hole in the DAC board. This configuration is advantageous since it allows fine alignment of the micromirror array by moving only the small mass of the FRAM array package. The flexible links between the FRAM array package and the DAC board make this arrangement possible.

Final tests confirmed that the projection engine allows the projection of VGA still images or animations at a frame rate of 60 Hz. All features of the control software and of the control electronics are functional.

2.6 Conclusion

The goal of the work was the demonstration of a projector engine based on the FRAM technology and capable of displaying 640×480 pixels at a frame rate of 60 Hz. To meet this objective, a new data/FRAM interface for a 480×1 FRAM array has been designed and fabricated. Customized control software allows the transfer of image data to the interface. The control electronics includes a DAC board, a FPGA board and a power supply board. The designed electronics and the associated software are compatible with the Digital Video Interface (DVI) standard. This allows the projection of a variety of images including still images and animation sequences. In addition, 480×1 FRAM arrays appropriate for the selected packaging approach (modified chip-on-board) have been designed, fabricated and tested. Full testing of 480×1 arrays before packaging has been successfully achieved by combining a specialized optical setup with a probe station. Very large arrays ($4,000 \times 1$) were also produced within this fabrication run therefore confirming the feasibility of such arrays.

One 480×1 FRAM array has been selected and packaged using a PC board/ metallic package assembly. This assembly has been connected to the data/FRAM interface. Integration of the projector engine into an existing test bench has been performed. Tests confirmed that the projector engine allows the projection of VGA format still images or animations at 60 Hz. All features of the control software and of the control electronics are functional.



Chapter 3

Brightness and Contrast of Images with Laser-Based Video Projectors

Michel Doucet, Mélanie Leclerc, Francis Picard,
and Keith K. Niall

Abstract Laser based video projectors (LBVP) are relatively new on the scene of video display. LBVPs consist of laser beams that are scanned and synchronously modulated. An important characteristic of the LBVP is the very short lifetime of the pixels. Each pixel produces light only for the very short time during which the laser beam passes on it. The short transients make the measurement of a LBVP image's properties more difficult to obtain in comparison to conventional display devices. The validity of the results is even questionable when conventional methods are used to measure the characteristics of LBVP images. This paper presents an alternative method for the characterization of brightness and contrast of LBVP images that is expected to provide more reliable results.

Proposed alternative method: A rapidly varying optical signal can be measured accurately provided that it is strong enough. In order to facilitate the measurements of the characteristics of the LBVP images, the detection paradigm must be changed in order to ensure a stronger optical signal on the photodetector. The proposed alternative method aims at achieving this goal. This alternative method involves measuring the irradiance (optical power per unit surface) in the plane before the screen. The radiance or luminance is then determined by computation using the measured reflectivity properties of the screen.

In the proposed method, the irradiance is measured at different points in the plane or surface located immediately in front of the surface normally occupied by the screen. The test pattern proposed for the measurement is a 10 by 10 chequer-board. The irradiance measurement is performed in the middle of each bright square. For each sampling position, a first measurement is performed for the

M. Doucet (✉), M. Leclerc, and F. Picard
INO, Québec (siège social), 2740, rue Einstein, Québec (Québec), Canada, G1P 4S4
e-mail: Michel.Doucet@ino.ca

K.K. Niall
Defence Research and Development Canada, Embassy of Canada, 501 Pennsylvania Avenue NW,
Washington, DC, USA
e-mail: Keith.Niall@forces.gc.ca

maximum value of irradiance, and a second one is performed for the minimum value with a reverse video test pattern.

Some modern screens possess microstructures allowing control of the directional scattering properties. For such screens, the scattering properties generally depend on the illumination direction. The mathematical representation of the screen's reflectivity properties that is used for the luminance computation is the Bidirectional Reflectance Distribution Function (BRDF). The BRDF gives the reflected radiance per unit of incident irradiance for any reflection direction and any illumination direction. Specialized devices exist that allow the efficient measurement of the BRDF spectral distribution. The method involves that a part of screen be sacrificed for the BRDF measurement. Having the spectral BRDF of the screen and a sampling of the screen irradiance, the luminance may be determined for any observer position and any point on the screen using a dedicated computer program.

Practical considerations: Parasitic light is generated both by the projector itself and by the environment. The veiling glare comes from the undesirable reflections on the lens surfaces and on mechanical parts of the projector. It increases with the total amount of optical power contained in the projected image. The ambient illumination and the back reflection on walls and objects are the two main sources of parasitic light generated by the environment. Such parasitic light is environment dependent. Parasitic back reflection occurs when light coming from the screen is reflected back toward it by walls and objects. The parasitic light considerably affects the image contrast since it adds background intensity on the screen. The 10 by 10 chequerboard test pattern was chosen in order to generate the unavoidable veiling glare that is representative of a normal image. Such test pattern contains about half of the maximum available optical power and generates about half of the maximum possible veiling glare. Choosing a pattern with less optical power would reduce the amount of veiling glare and result in an estimated contrast not representative of the contrast for a normal image. The signal deterioration, due to environment parasitic light, changes with the configuration of the projection room and this is a problem for standard measurement.

The sampling process required for any measurement must be adapted to take into account the spatial distribution as well as the temporal fluctuations of the phenomenon to be characterized. In the case of a LBVP image, care must be taken to avoid problems related to the pixelization of the image and fast variations due to scanning. Depending on the LBVP technology, there can be dead zones in between pixels or inside each of them. Such periodic structures may induce measurement errors when the detector dimensions do not include an integer number of structure periods. Depending on the relative position of the photodetector with respect to the pixel array, more or less dead zones are included in the effective area of the detector. This means that the signal is dependent on the detector position with respect to the pixel lattice. These potential spatial sampling problems can be avoided if simple rules are respected. These rules concern the number of pixels contained within the detector's effective surface, the shape of the detector, and its orientation with respect to the pixels.

Sommaire Les projecteurs vidéo à laser (LBVP) sont relativement nouveaux sur la scène de l'affichage vidéo. De tels projecteurs utilisent des faisceaux laser qui sont balayés et modulés en synchronisme avec le balayage. Une particularité du LBVP est la durée de vie très courte des pixels. Chaque pixel produit de la lumière seulement durant le temps très court correspondant au passage du rayon laser sur le pixel. Les variations très rapides du signal optique font en sorte qu'il est plus difficile de mesurer les propriétés des images produites par LBVP en comparaison avec le cas des dispositifs d'affichage conventionnels. On peut même douter de la validité des résultats lorsque des méthodes conventionnelles sont employées pour mesurer les caractéristiques des images produites par un LBVP. Cet article présente une méthode alternative pour la caractérisation de la brillance et du contraste des images de LBVP qui devrait fournir des résultats plus fiables.

Méthode alternative proposée: Un signal optique à variation rapide peut être mesuré et caractérisé avec précision à condition d'être assez fort. Afin de faciliter la mesure des caractéristiques des images de LBVP, il y aurait avantage à changer le paradigme de détection de manière à assurer un plus fort signal optique sur le détecteur de l'appareil de mesure. La méthode alternative proposée vise à atteindre ce but. Cette méthode alternative implique de mesurer l'éclairement (puissance optique par unité de surface) dans le plan situé juste devant l'écran. La radiance ou la luminance est alors déterminée par calcul en utilisant les mesures des propriétés de réflectivité de l'écran.

Dans la méthode proposée, l'éclairement est mesuré à différents points dans le plan ou la surface située immédiatement devant la surface normalement occupée par l'écran. La figure de test proposée pour la mesure est un échiquier de 10 carreaux (5 éclairés, 5 noirs) sur 10. La mesure de l'éclairement est effectuée au milieu de chacun des carreaux brillants. Pour chaque position d'échantillonnage, une première mesure est effectuée pour la valeur maximum de l'éclairement, et une seconde est effectuée pour la valeur minimum avec la figure de test en inverse vidéo.

Certains types d'écrans modernes possèdent des microstructures permettant le contrôle des propriétés directionnelles de dispersion de la lumière. Pour de tels écrans, les propriétés de dispersion dépendent généralement de la direction incidente de l'illumination. La représentation mathématique des propriétés de la réflectivité de l'écran qui est employée pour le calcul de luminance est la fonction de distribution bidirectionnelle de la réflectivité (BRDF). La BRDF donne la radiance réfléchie par unité d'éclairement incident pour n'importe quelle direction de réflexion et n'importe quelle direction d'illumination. Il existe des appareils spécialisés permettant la mesure de la distribution spectrale de la BRDF de façon efficace. La méthode implique qu'une portion de l'écran soit sacrifiée pour la mesure de la BRDF. Disposant de la BRDF spectrale de l'écran et d'un échantillonnage de l'éclairement d'écran, la luminance peut être déterminée pour n'importe quelle position d'observateur et n'importe quel point sur l'écran en utilisant un logiciel dédié.

Considérations pratiques: La lumière parasite est produite par le projecteur lui-même et par l'environnement ambiant. Le voile lumineux (veiling glare) est généré

par des réflexions indésirables sur les surfaces optiques de l'objectif et sur les pièces mécaniques du projecteur. Il augmente avec la quantité totale de puissance optique contenue dans l'image projetée. L'illumination ambiante et la rétro-réflexion sur les murs et les objets sont les deux principales sources de lumière parasite causée par l'environnement. Une telle lumière parasite est spécifique au lieu où se fait la projection. La rétro-réflexion parasite se produit quand la lumière venant de l'écran est rétro-réfléchi vers l'écran par les murs et les objets présents dans la pièce. La lumière parasite affecte considérablement le contraste de l'image projetée puisqu'elle ajoute une intensité de fond sur l'écran. L'échiquier de 10 par 10 a été choisi de manière à produire l'inévitable voile lumineux dans une proportion représentative de ce que l'on retrouve dans le cas de la projection d'une image normale.

Une telle figure de test contient environ la moitié de la puissance optique maximum disponible et produirait donc environ la moitié de la quantité de voile lumineux possible. Le choix d'une figure de test avec moins de puissance optique réduirait la quantité de voile lumineux produit et résulterait en une estimation du contraste qui serait non représentative du contraste pour une image type. La détérioration du signal par la lumière parasite de cause environnementale change d'un endroit à l'autre et constitue donc un problème pour la standardisation des mesures.

Le procédé d'échantillonnage requis pour la caractérisation d'un phénomène doit pouvoir tenir compte de la distribution spatiale aussi bien que des fluctuations temporelles du phénomène en question. Dans le cas de la caractérisation d'une image LBVP, une attention particulière doit être portée aux problèmes liés au tramage de l'image et aux variations rapides du signal optique dues au balayage. Selon la technologie de LBVP, il peut y avoir des zones mortes entre les pixels ou à l'intérieur de chacun d'eux. De telles structures périodiques peuvent induire des erreurs de mesure lorsque les dimensions du détecteur optique ne comprennent pas un nombre entier de périodes de la structure. Selon la position du détecteur optique relativement à la matrice de pixels, un nombre plus ou moins important de zones mortes sont incluses dans la surface active du détecteur. Ceci signifie que le signal dépend de la position relative du détecteur optique relativement au treillis de pixels. Ces éventuels problèmes d'échantillonnage spatial peuvent être évités si certaines règles simples sont respectées. Ces règles concernent le nombre de pixel contenus dans la surface efficace du détecteur optique, la forme du détecteur et son orientation par rapport à l'agencement des pixels de l'image.

Laser-based video projectors (LBVP) are relatively new on the scene of video display. LBVPs consist of laser beams that are scanned and synchronously modulated. An important characteristic of the LBVP is the very short lifetime of the pixels. Each pixel produces light only for the very short time during which the laser beam passes on it. Depending on the type of LBVP, the pixel lifetime decreases with the number of pixels or the number of image lines or columns.

The short transients make the measurement of LBVP image properties more difficult to obtain in comparison to conventional display devices. There are some doubts on the validity of the characterization of LBVP images obtained by conventional

methods. The mandate to propose an alternative method better adapted to the particular case of the LBVP was given to INO. This paper presents the proposed alternative method specifically defined for the characterization of brightness and contrast of LBVP images. It is expected that this method will provide more reliable results. However, this remains to be tested since this paper is focused on theoretical considerations.

3.1 LBVP Displaying Mechanisms

The manner by which the image is produced dictates the methods to use for its characterization. The two principal mechanisms used to generate dynamic images from laser sources are the flying spot and the linear SLM.

- **Flying spot:** This method consists of a laser beam that is scanned in two orthogonal directions. The intensity of the laser beam is modulated as the spot moves on the screen. Depending on the type of laser used, the beam intensity is modulated by acting directly on the electric power supplied to the laser or by means of an external modulator. The modulation with an external optical modulator is the most common for high power laser and consequently for projection display.
- **Linear SLM:** Many of the laser-based projector technologies make use of the deformation of a reflective microstructure to modulate the light. The deformation of the microstructure causes deflection¹ of the illumination beam and modulation is accomplished by means of a blocking stop. The deformable microstructure is replicated to form a linear array of light modulators or a linear Spatial Light Modulator (linear SLM). When illuminated with a thin laser line, a linear SLM allows the display of a line of pixels. A complete image is produced by scanning the line of pixels by means of a scanning mirror.

The flying spot displaying mechanism is similar to a conventional cathode ray tube (CRT) mechanism except that the electron beam is replaced by a laser beam. Another distinction with CRT is the lifetime of the pixel which is significantly longer in the case of the CRT since the pixel light is generated by phosphorescence under the control of the electron beam energy.

3.2 Conventional Image Characterization

Conventionally, the brightness and contrast of a display are measured using a telephotometer. As its name indicates, the telephotometer allows the measurement of light from a distant position. The telephotometer has an imaging lens that projects the image of the scene on a detector. Two stops are used to isolate a beam of light

¹By deviation of the beam on a tilted mirror or by diffraction grating effect.

that comes from a small part of the observed scene. The flux of the isolated beam is entirely collected on the photodetector which is equipped with a special filter. The spectral response of the set consisting of the photodetector and the filter emulates the spectral response of the human visual system. The filter/photodetector spectral response together with the optical configuration guarantees that the electrical output signal of the detector is proportional to the luminance of the observed zone on the scene. The proportionality constant is determined by a calibration and modern digital telephotometers directly give the calibrated luminance value.

In the context of display characterization, the telephotometer is pointed on a small part of the display screen and it collects rays coming from this observation zone. Only a fraction of the optical power reaches the detector of the telephotometer. In the case of LBVP, the optical signal is difficult to measure since it is relatively weak and contains rapid transients. Telephotometers are generally designed and optimized for relatively slow phenomena. Hence, such devices do not perform at their best for LBVP characterization and it would not be surprising to have erroneous results.

3.3 Highlight of the Proposed Alternative Method

A rapidly varying optical signal can be measured accurately provided that it is strong enough. In order to facilitate the measurement of the characteristics of the LBVP images, there is an advantage to changing the conventional detection paradigm in such a manner to ensure a larger optical signal on the photodetector. The proposed alternative method aims at achieving this goal. This alternative method involves measuring the irradiance (optical power per unit surface area) in the plane before the screen (Halsted 1993). The radiance or luminance is then determined by computation using the measured reflectivity properties of the screen.

The mathematical representation of the screen's reflectivity properties that is used for the luminance computation is the Bidirectional Reflectance Distribution Function (BRDF) (Commission internationale de l'éclairage 1983). The BRDF gives the reflected radiance per unit of incident irradiance for any reflection direction and any illumination direction. Specialized devices exist that allow the efficient measurement of the BRDF spectral distribution. The method involves that a part of screen be sacrificed for the BRDF measurement. Having the spectral BRDF of the screen and a sampling of the screen irradiance, the luminance may be determined for any observer position and any point on the screen using a dedicated computer program. The method is schematically presented in Fig. 3.1.

3.4 Optical Properties of Screens

The screen is a very important component of a projection system. The screen is a surface that receives light from the projector and redistributes it toward the audience. The rays generated by the projector and reaching a particular pixel on the

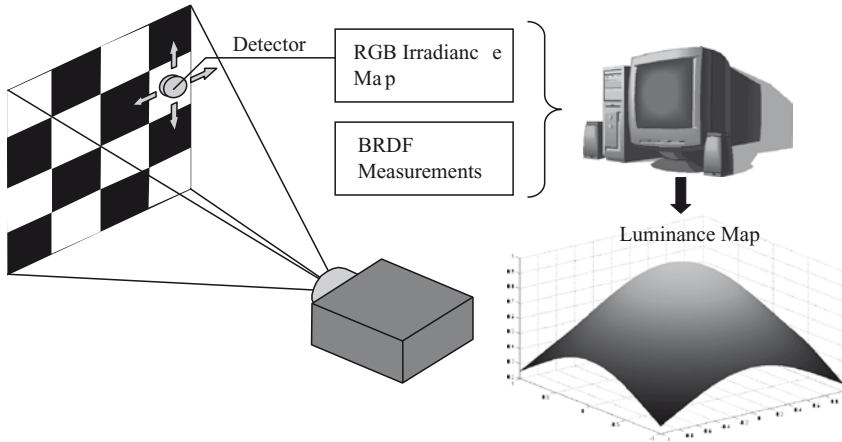


Fig. 3.1 Schematic description of the proposed alternative (BDRF) method

screen can be considered as nearly collimated. The screen redistributes the directions of these rays and each pixel becomes a secondary source that can be seen from all positions in front of the screen. Depending on the type of screen, which can be reflective or transmissive, various physical principles can be used for the redistribution of the rays' direction.

The proposed method involves the knowledge of the optical properties of the screen. The next sub-sections are dedicated to the description of the optical properties of screens using BRDF.

3.4.1 Reflectivity Characteristics of a General Screen

Screen reflectivity characteristics can be represented using the concept of the screen gain, or as a Bidirectional Reflectance Distribution Function (BRDF). The BRDF is a more precise and more complex representation of the reflectivity properties of materials. The screen gain is a valid description only for particular types of materials.

The BRDF gives the differential reflected radiance produced by an oriented incident beam. The geometry is illustrated in Fig. 3.2. The oriented beam is defined by the element source δA_s and the reflecting element δA_d . This beam produces a spectral irradiance $dE(\theta_i, \varphi_i, \lambda)$ on δA_d . The differential spectral radiance $dB_r(\theta_r, \varphi_r, \lambda)$ of the reflected radiation is measured by receptor element δA_r .

The BRDF is defined by:

$$f_r(\theta_i, \varphi_i, \theta_r, \varphi_r, \lambda) = \frac{dB_r(\theta_r, \varphi_r, \lambda)}{dE(\theta_i, \varphi_i, \lambda)} \quad (3.1)$$

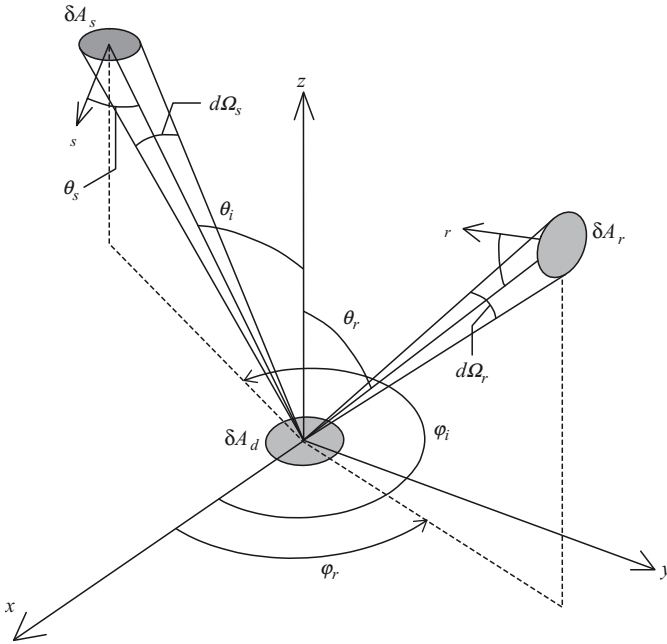


Fig. 3.2 Geometry of the BRDF

It is a function that characterizes the geometrical and spectral-reflectivity properties of the reflecting surface for points around the centre of the small element δA_d . For elements δA_s , δA_d and δA_r , small in comparison to the distances between them, the reflected radiance varies proportionally with the irradiance, and the ratio of these quantities varies only with the reflective properties of the small element δA_d . This reflective element must be small enough to avoid non-uniform properties.

The angular distribution of the radiance of some types of highly diffusing materials does not change significantly with the incident direction. For those materials, the BRDF varies only with the reflection angular coordinates, $f_r(\theta_i, \varphi_i, \theta_r, \varphi_r, \lambda) = f_r(\theta_r, \varphi_r, \lambda)$. The Lambertian diffuser belongs to this family of materials, because its radiance is constant. For Lambertian materials, the BRDF is constant, $f_r(\theta_i, \varphi_i, \theta_r, \varphi_r, \lambda) = f_r(\lambda)$. For screens made of highly diffusing materials, the gain $g(\theta_r, \varphi_r, \lambda)$ is defined as the ratio of their incident direction-independent BRDF over the BRDF of a perfect Lambertian screen:

$$g(\theta_r, \varphi_r, \lambda) = \frac{f_r(\theta_r, \varphi_r, \lambda)}{f_{Lambertian}} \tag{3.2}$$

3.4.2 Determining Spectral Radiance of a Screen from its BRDF

A projector that illuminates a screen can be considered as a small element source because the screen is normally located at a large distance from the projector. Thus, we have the same geometry as in Fig. 3.2, with the exit pupil P_E of the projector objective acting as the source. This is represented in Fig. 3.3. A small part, δA_{screen} , of the screen is considered and the observer is replaced by a receiving element δA_r . The element δA_{screen} is large enough to contain a large number of image pixels, but is small in comparison to the screen dimensions and the distance D_{PS} between the screen and the projector. The receiving surface has to be located at a large distance from the screen, in comparison to the size of the element δA_{screen} .

Figure 3.3 illustrates the general case of a curved screen. The element δA_{screen} is localized with the position vector \vec{r} , while the vector \vec{s} gives the position of the receiving element δA_r . These two vectors have their origin at O and their ending points are B and C , respectively. The angular coordinates used for the measurement of the BRDF of the screen reflective material are referenced with respect to the local normal \hat{n} of the surface, and with respect to a preferential direction defined by the unit vector \hat{d} . This preferential direction may be chosen, for example, along oriented microstructures devised to reflect light preferentially in a given direction. Usually, \hat{d} will be along the vertical or the horizontal direction.

For simplicity, the optical axis of the projector points toward the center of the screen, where the origin O of the coordinate system is located. The optical axis

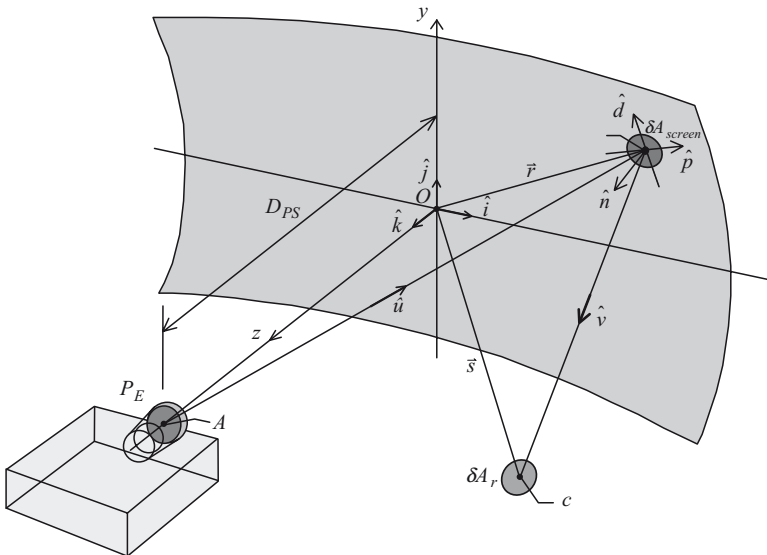


Fig. 3.3 Geometry for computing the radiance reflected by a screen

coincides with the z -axis, which is along the normal to the screen at point O. The y -axis is oriented in the vertical direction.

The spectral radiance $B_r(\theta_r, \varphi_r, \lambda)$ perceived by the receiver element δA_r is obtained by an integral sum over the source. The source is divided into a set of non-overlapping small source elements δA_s with uniform radiant properties. Using (3.1), the spectral radiance is given by:

$$B_r(\theta_r, \varphi_r, \lambda) = \int_{P_E} dB_r(\theta_r, \varphi_r, \lambda) \big|_{\delta A_s} = \int_{P_E} f_r(\theta_i, \varphi_i, \theta_r, \varphi_r, \lambda) dE(\theta_i, \varphi_i, \lambda) \quad (3.3)$$

where $dB_r(\theta_r, \varphi_r, \lambda) \big|_{\delta A_s}$ is the spectral radiance reflected by δA_{screen} toward δA_r , but contributed by the small source element δA_s alone. The angular coordinates θ_i , φ_i , θ_r and φ_r are defined as in Fig. 3.2, except that the φ are measured with respect to the preferential direction defined by the unit vector \hat{d} . Because the exit pupil is small in comparison to the distance D_{PS} , the angles θ_i and φ_i do not vary much over the elements δA_s of the exit pupil. Moreover, the BRDF varies slowly with the incidence angular coordinates, and all elements of the source are associated with approximately the same value of BRDF.² In these conditions, the BRDF within the integral may be replaced by a mean value:

$$B_r(\theta_r, \varphi_r, \lambda) = f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_r, \varphi_r, \lambda) \int_{P_E} dE(\theta_i, \varphi_i, \lambda) = f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_r, \varphi_r, \lambda) E(x_B, y_B, \lambda) \quad (3.4)$$

where $\bar{\theta}_i$ and $\bar{\varphi}_i$ are the angular coordinates corresponding to the central point A of the exit pupil, and $E(x_B, y_B, \lambda)$ is the total spectral irradiance at the central point B of the reflecting element δA_{screen} . The radiance is obtained by a simple multiplication of the BRDF and the irradiance produced by the projector.

To complete the computation of the radiance, general expressions for the angular coordinates of lines AB and BC must be determined. Let (x_B, y_B, z_B) and (x_C, y_C, z_C) be the coordinates of points B and C respectively. The unit vectors \hat{u} and \hat{v} along lines AB and BC respectively are given by:

$$\begin{aligned} \hat{u} &= \frac{x_B \hat{i} + y_B \hat{j} + (z_B - D_{PS}) \hat{k}}{\sqrt{x_B^2 + y_B^2 + (z_B - D_{PS})^2}} \equiv \hat{u}_x \hat{i} + \hat{u}_y \hat{j} + \hat{u}_z \hat{k} \\ \hat{v} &= \frac{(x_C - x_B) \hat{i} + (y_C - y_B) \hat{j} + (z_C - z_B) \hat{k}}{\sqrt{(x_C - x_B)^2 + (y_C - y_B)^2 + (z_C - z_B)^2}} \equiv \hat{v}_x \hat{i} + \hat{v}_y \hat{j} + \hat{v}_z \hat{k} \end{aligned} \quad (3.5)$$

The components \hat{u}_\perp and \hat{u}_\parallel , of the unit vector \hat{u} , that are respectively perpendicular and parallel to the plane of element δA_{screen} , are given by:

$$\begin{aligned} \hat{u}_\perp &= (\hat{u} \bullet \hat{n}) \hat{n} \\ \hat{u}_\parallel &= \hat{u} - \hat{u}_\perp = \hat{u} - (\hat{u} \bullet \hat{n}) \hat{n} \end{aligned} \quad (3.6)$$

² A material with rapid BRDF variation is not a good material for a screen because it will produce undesirable variations in brightness from one region of the screen to the next.

where $\hat{u} \cdot \hat{n}$ is the scalar product of the two unit vectors. This gives the equations for the angular coordinates for incident direction:

$$B_r(\theta_r, \varphi_r, \lambda) = \int_{F_E} dB_r(\theta_r, \varphi_r, \lambda) \big|_{\delta\Omega_s} = \int_{F_E} f_r(\theta_i, \varphi_i, \theta_r, \varphi_r) d\Omega_i$$

$$\rightarrow \bar{\varphi}_i = \arccos\left(\frac{\hat{u}_{\parallel} \cdot \hat{d}}{|\hat{u}_{\parallel}|}\right) = \arccos\left(\frac{\hat{u} \cdot \hat{d} - (\hat{u} \cdot \hat{n})(\hat{n} \cdot \hat{d})}{|\hat{u}_{\parallel}|}\right) = \arccos\left(\frac{\hat{u} \cdot \hat{d}}{|\hat{u}_{\parallel}|}\right) \quad (3.7)$$

where \hat{d} is the unit vector along the preferential direction, and $|\hat{u}_{\parallel}|$ is the module of the component \hat{u}_{\parallel} . Similarly, we have for the reflection angular coordinates:

$$\theta_r = \arccos(\hat{v} \cdot \hat{n})$$

$$\varphi_r = \arccos\left(\frac{\hat{v} \cdot \hat{n}}{|\hat{v} - (\hat{v} \cdot \hat{n})\hat{n}|}\right) \quad (3.8)$$

3.4.3 Determining Reflected Luminance

The luminance of a light stimulus can be determined from the integral sum of the spectral radiance weighted by the spectral response of the human eye. The spectral reflected radiance is given by (3.4). Using those two equations, the luminance of point (x_B, y_B) of the screen, for an observer located at point (x_C, y_C, z_C) , is given by:

$$L(x_B, y_B, x_C, y_C, z_C) = K \int_{360nm}^{830nm} V(\lambda) B_r(x_B, y_B, \theta_r, \varphi_r, \lambda) d\lambda =$$

$$= K \int_{360nm}^{830nm} V(\lambda) f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_r, \varphi_r, \lambda) E(x_B, y_B, \lambda) d\lambda \quad (3.9)$$

where $V(\lambda)$ is the normalized spectral-efficiency function of the human visual system and K is a scaling factor. In the case of a laser-based projector, the spectral content of the source consists of a finite number of very narrow bands corresponding to the individual narrow laser lines.³ For such laser projectors, (3.9) is reduced to a simple sum:

$$L(x_B, y_B, x_C, y_C, z_C) = K \sum_{q=1}^Q V(\lambda_q) f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_r, \varphi_r, \lambda_q) E_q(x_B, y_B) \quad (3.10)$$

³Contrary to popular belief, the radiation produced by a laser is not absolutely monochromatic. Many phenomena are responsible for this fact. For example, fast molecular movements in a gas laser produce spectral broadening of the laser radiation due to the Doppler effect.

where $E_q(x_B, y_B)$ is the irradiance at point (x_B, y_B) of the screen due to the q^{th} laser with central wavelength λ_q .

3.4.4 Determining Contrast

For any point on the screen, the contrast is defined as the ratio between the maximum and the minimum values of the range of luminance that can be displayed at the observed point. The contrast ratio C_r of a point (x_B, y_B) on the screen is given by:

$$C_r(x_s, y_s, \hat{u}) = \frac{L_{\max}(x_B, y_B, \hat{v})}{L_{\min}(x_B, y_B, \hat{v})} \quad (3.11)$$

This definition takes into consideration the directional nature of the luminance. Equation (3.11) gives the contrast ratio for a specific point on the screen observed in a specific direction defined by the unit vector \hat{v} .

The contrast ratio can be measured directly with a telephotometer. This process requires many measurements to be taken to ensure a good sampling of different points on the screen and of different observation positions (directions).

The contrast ratio can also be determined with the BRDF. Using (3.10) in (3.11), the contrast ratio can be expressed as:

$$C_r(x_B, y_B, \hat{v}) = \frac{L_{\max}(x_B, y_B, \hat{v})}{L_{\min}(x_B, y_B, \hat{v})} = \frac{\sum_{q=1}^Q V(\lambda_q) f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_{\hat{v}_r}, \varphi_{\hat{v}_r}, \lambda_q) E_{q_{\max}}(x_B, y_B)}{\sum_{q=1}^Q V(\lambda_q) f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_{\hat{v}_r}, \varphi_{\hat{v}_r}, \lambda_q) E_{q_{\min}}(x_B, y_B)} \quad (3.12)$$

where $E_{q_{\max}}(x_B, y_B)$ and $E_{q_{\min}}(x_B, y_B)$ are respectively the maximum and minimum values of the range of irradiance that can be produced at point (x_B, y_B) by the q^{th} laser line of the projector. In the case where the reflection properties of the screen do not change significantly with the wavelength (a perfectly white, or grey, screen), the equation is greatly simplified:

$$\begin{aligned} C_r(x_B, y_B, \hat{v}) &= \frac{f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_{\hat{v}_r}, \varphi_{\hat{v}_r}) \sum_{q=1}^Q V(\lambda_q) E_{q_{\max}}(x_B, y_B)}{f_r(\bar{\theta}_i, \bar{\varphi}_i, \theta_{\hat{v}_r}, \varphi_{\hat{v}_r}) \sum_{q=1}^Q V(\lambda_q) E_{q_{\min}}(x_B, y_B)} = \\ &= \frac{\sum_{q=1}^Q V(\lambda_q) E_{q_{\max}}(x_B, y_B)}{\sum_{q=1}^Q V(\lambda_q) E_{q_{\min}}(x_B, y_B)} = \frac{I_{\max}^a(x_B, y_B)}{I_{\min}^a(x_B, y_B)} = C_r(x_B, y_B) \end{aligned} \quad (3.13)$$

where $I_{max}^a(x_{B'},y_{B'})$ and $I_{min}^a(x_{B'},y_{B'})$ designate, respectively, the maximum and minimum values of the irradiance at point $(x_{B'},y_{B'})$.

3.5 Practical Considerations

Many practical considerations should be addressed for the implementation of the proposed method. Parasitic light generally has significant impact on the results. In regard to parasitic light and for purpose of measurement standardization, the test patterns displayed during the measurements must be selected carefully. Test patterns must be chosen to ensure that a representative amount of image-dependent parasitic light be generated. Moreover, both spatial and temporal samplings of the optical signal may also introduce measurement error if they are not done adequately.

3.5.1 *Parasitic Light*

Parasitic light is generated both by the projector itself and by the environment. Veiling glare belongs to the first type of parasitic light while the back reflection and the ambient illumination are the principal sources of parasitic light due to the environment.

3.5.1.1 Veiling Glare

Despite technical progress in optical anti-reflection (AR) coating, there is still non-null reflection at the glass interfaces. The most severe problem occurs in complex optical systems that contain large numbers of lenses, as a part of the light reflected on lens interfaces is transmitted toward the screen. This parasitic light is called veiling glare. Veiling-glare light also includes specular and diffuse reflections of light by the mechanical parts holding the lens.

The amount of veiling-glare light increases with the intensity of the beam passing through the optical channel. It can be reasonably supposed that the principle of superposition applies; the veiling-glare phenomenon can thus be represented by a linear equation, according to Badano and Flynn (2000). Moreover, if the properties of the AR coating are uniform over lens surfaces, it is reasonable to presume that veiling glare is not greatly dependent on image-intensity distribution.

Veiling glare is intrinsic to the projector, but depends on the amount of light passing through the optics and the amount of light projected on the screen. It is important to select the image used in measurements to include a representative amount of parasitic light from the veiling-glare process. Illuminating only the small portion of the screen where measurements are made has the advantage of reducing parasitic light, such as that caused by veiling glare. However, such a low level of veiling-glare light is not representative of normal operating conditions, and leads to overestimations of contrast value.

3.5.1.2 Environment Dependent Parasitic Light

The purpose of the projection screen is to diffuse light coming from the projector toward the observers. Unfortunately, the light diffused by the projection screen not only illuminates observers' eyes, but also the projection-room walls and objects within the room. These objects, in turn, diffuse the light throughout the room, including back to the screen. This back-reflected light can significantly reduce the contrast of the projected image by producing background illumination. The amount of light reflected back to the screen depends on the configuration of the room, and the absorption and reflection characteristics of the walls and objects. The resulting image brightness, and especially the contrast, can differ significantly from one room to another. In a manner similar to back reflection, ambient illumination reduces the contrast of a projected image by way of background illumination. With a high level of ambient illumination, the screen can appear white even when the projector displays a black image. Contrast reduction due to ambient illumination is less severe for high-powered projectors. Ambient illumination may come from light passing through curtains or under doors. It may also come from any source of light inside the room, such as computer screens or illuminated signs.

The signal deterioration, due to environment parasitic light, changes with the configuration of the projection room and this is a problem for standard measurement. Fortunately, simple techniques exist to overcome these problems. The contribution of the environment dependent parasitic light may be measured and then subtracted from the useful signal (Zalewski 1995), which is accomplished by placing the photodetector in the shadow of a black opaque mask. The light from the projector (including the veiling glare) is blocked and only the ambient illumination and the parasitic back reflection can reach the photodetector.

3.5.2 Spatial Sampling

The image produced by a video projector generally consists of a periodic arrangement of a very large number of small discrete elements called pixels. Depending on the technology used, there can be dead zones between pixels or inside each individual pixel. This periodic structure may induce measurement errors. As well, speckle noise produces a random spatial redistribution of energy that may also induce measurement errors if the spatial sampling is inadequate.

3.5.2.1 Pixel Filling, Pixel Overlap, Pixel Shape

This section concerns all structured non-uniformity of the spatial distribution of energy within each pixel at the screen. Figure 3.4 illustrates the problem of spatial sampling in the presence of structured non-uniformities, in the particular case of horizontal dead stripes between successive rows of pixels. Such a pattern is characteristic of laser-based video projectors with a linear SLM.

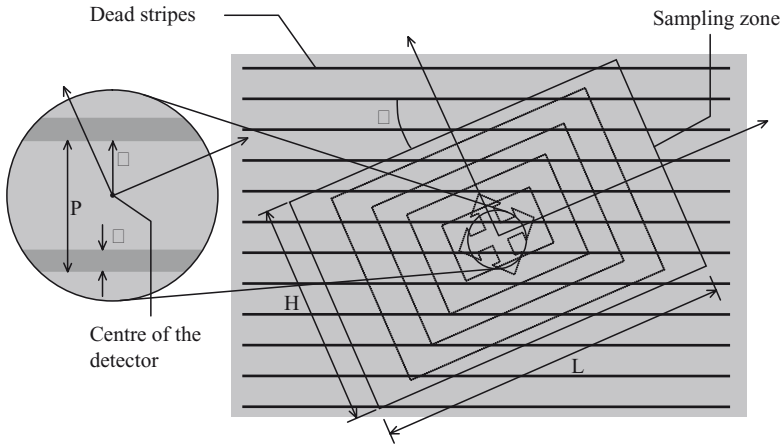


Fig. 3.4 Effect of dead stripes between pixel rows on the energy measurement for different sizes of the sampled zone

Measurement of the image's irradiance, or its radiance or luminance, involves spatial integration over a group of pixels. The amount of energy collected by the measurement device depends on the zone in the image sampled by the detector's active area. Figure 3.5 shows the variation in the proportion of the area occupied by dead stripes as the size of the detector is increased. The dimensions of the detector L and H , the stripes width δ and the centering parameter Δ are all normalized by the height P of the pixel. The curve in Fig. 3.5d corresponds to a round detector, while Fig. 3.5a–c represent the effect obtained with a square detector. The detector is perfectly oriented with respect to the stripes ($\theta=0$) in the case of the first curve, while the second and the third correspond to the rotation angles θ of 45° and 40° respectively. In all cases, the normalized width δ of the stripes is 0.1. This means that 10% of each pixel is occupied by dead area. The best results are achieved with a square detector rotated by 45° with respect to the stripes. The fluctuations of the area ratio are reduced very quickly, and the curve reaches its converging value for a detector size of about 10 pixels. The result for a rotation angle of 40° is nearly as good as its counterpart with a rotation angle of 45° . This means that a square detector having 10 pixels high, and rotated by about 45° with respect to the dead stripes, gives a measure of irradiance (power per unit area) very close to the ideal case of a detector which would measure the total irradiance for the zone of interest.

These results are also valid for any type of systematic non-uniformity in the energy distribution within the pixel. This includes the case of partial overlapping (flying-spot projector) and pixels with unusual shapes. The fast-convergence behaviour for the square detector rotated by 45° is valid even for a filling factor (active area/total area) as poor as 55% (curve with an asymptotic value of 0.45). The amount of energy collected by the device, assuming a square detector rotated

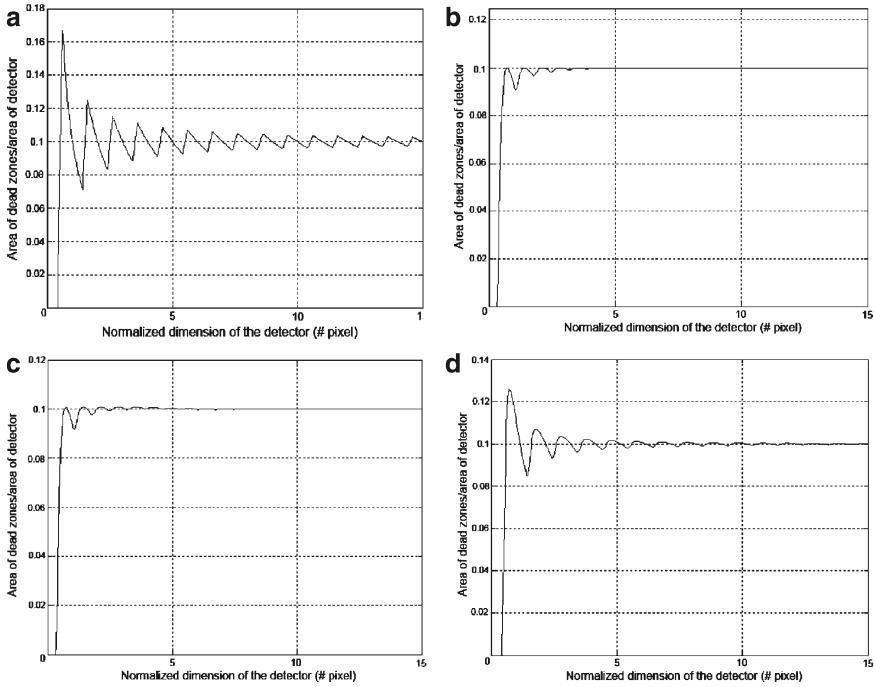


Fig. 3.5 Variation in the proportion of the dead zone as a function of the detector size

by 45° , is also insensitive to the centering of the detector with respect to the pixels. Centering errors of up to \pm half of a period ($\Delta=P/2$) do not alter the asymptotic convergence. This is true for horizontal patterns as well as vertical patterns, due to the symmetries. The good behaviour of the tilted-square detector is explained by the fact that dead zones enter or exit the detector at the detector's corner tips, where its width is small enough to not significantly modify the energy budget.

The size of the image produced by a video projector is generally quite large. It is not certain that large and fast detectors, with sizes around 10 image pixels, exist. For a detector smaller than 10 image pixels, the mean value of the measurements at different positions in the image should be taken. The detector can be moved using a motorized linear displacement stage. Figure 3.6 shows the curve corresponding to the mean value for 50 displacements of the detector as a function of the detector size. The detector is moved on a distance of 10 image pixels in 50 equidistant steps; the results are accurate even for detectors with a height as small as one image pixel.

3.5.2.2 Speckle

Speckle occurs when the wavefront of a coherent beam is randomly modified. For example, this can happen when a laser beam is reflected on, or transmitted through,

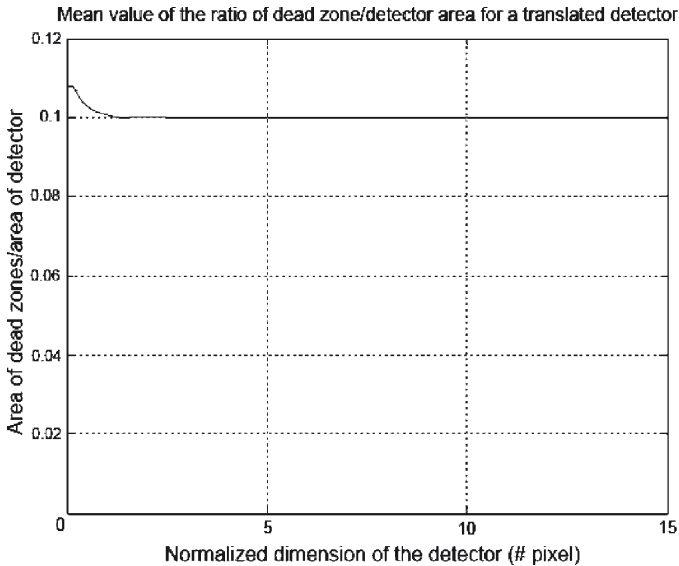


Fig. 3.6 Mean over a large set of detector positions for the ratio of the dead-zone area/detector area of 0.1

an uneven surface. These conditions are met when an image is projected by a LBVP and reflected by a diffusing screen. Due to time averaging, the speckle is greatly attenuated because the scanning process produces fast modifications of the speckle pattern. However, setting aside its effects on image quality, speckle can cause measurement error if the speckle grains are not small enough in comparison to the size of the detector.

As for other interference phenomena, there is no creation or annihilation of energy, only a redistribution of the incident energy. The energy integration in a portion of the image containing a large number of speckle grains is probably representative of the mean energy incident on the corresponding part in the object plane.

In an image produced by a LBVP, the speckle is generated when the light is diffusely reflected by the screen. Speckle patterns are unlikely to occur in the optical path preceding the screen, because the optical surfaces that comprise the projector do not have scattering properties. The problem of speckle is significant only for the conventional measurements done with a telephotometer after reflection on the screen. For such measurements, one must ascertain that the observed zone contains a large number of speckle grains. In measuring image irradiance before reflection from the screen as in the proposed alternative method, the speckle phenomenon does not need to be taken into consideration.

3.5.3 *Detector Temporal Response*

As mentioned in the previous paragraphs, the lifetime of a particular pixel is very brief, but intense. A detector receiving the light of a given pixel perceives a rapid and substantial increase in luminous energy, followed by a similar, sudden decrease. The shape of the signal depends on the shape and orientation of the detector and also on its relative size with respect to the pixel size.

Real detectors do not react instantaneously to input signals. The response associated to a very fast input signal is generally only a more or less distorted replica of the input signal. The capability of a detector to respond rapidly is evaluated in the frequency domain and the cut-off frequency is the parameter that characterizes this capability. A cut-off frequency of about:

$$f_c \geq \frac{5NF_r}{1+n_d} \quad (3.14)$$

is probably enough for the detection system (photodiode plus amplifier), where N is the number of pixels in the scan direction, F_r is the frame rate in Hz and n_d is the size of the detector measured in number of pixels along the scanning direction.

Faithful reproduction of the input signal is useful because it permits a better control of the measurement process, but this is not absolutely required. A detector generating a signal proportional to the received optical energy can also provide valid information for the evaluation of the image brightness and contrast. In the case of the flying-spot projector which involves optical phenomena thousands of times faster than those for a linear SLM, the existence of a detector fast enough to follow the optical signal is doubtful and a specific measurement strategy might be recommended. For flying spot LBVP, the optical signal corresponding to a pixel is a very short pulse followed by a long, slow varying plateau. The pulse is similar in duration and intensity to the signal produced by a pulsed laser. Special detectors exist for such pulsed lasers. These detectors directly measure the energy contained in the laser short pulse. However, such detectors cannot measure a long pulse, or a slowly varying optical signal. The idea is to combine a pulse detector with a photodiode. The two detectors are placed, one after the other, along the scanning direction. The pulse detector measures the energy of the pulse (useful signal) while the photodiode measures the mean energy (parasitic signal) between pulses.

3.6 **Description of the Proposed Method**

The proposed method should allow the measurement of the brightness and contrast of a screen image projected by a LBVP. The previous sections dealt with theoretical and practical considerations on a general basis; the following explore the subject in greater detail. The proposed method involves three characterization steps. First, the irradiance is measured at a plane before the screen. Second, the screen's BRDF is

characterized. Finally, the luminance is computed for any observer position, using the measured irradiance pattern produced by the projector at the screen position and the screen BRDF.

3.6.1 Step 1: Irradiance Measurement

In the proposed method, the irradiance is measured at different points in the plane or the surface located immediately in front of the screen. The selection of the test pattern used for the measurement is important since it must provide a representative amount of the intrinsic parasitic light. The selection of the detector system is also important to avoid problems related to spatial sampling. The spectral contents of the optical signal must be characterized to allow the computation of the luminance of the beams reflected by the screen. All those subjects will be summarized in the following subsections together with a simple method to manage the problem of the environment-dependent parasitic light.

3.6.1.1 Test Pattern

The proposed test pattern consists of a 10 by 10 chequerboard. Half of the screen is at the lowest value of the grey scale, while the rest is at the maximum value. This provides about half the total optical power that can be produced by the projector and a representative amount of intensity-dependent intrinsic parasitic light (veiling glare, etc.). Measurements are done at the center of the bright squares. The magnification factor should be adjusted so that each chequerboard square is at least twice as large as the detector to avoid edge effects. The maximum values of the optical irradiance are measured at the center of the bright squares, while the minimum values are measured at the same positions, but with the video inverse pattern. Using an image with smooth transitions between adjacent dark and bright regions is recommended to eliminate high-spatial frequencies and avoid image aliasing. To minimize back reflection from the surrounding environment, replacing the screen with a black absorbing curtain is recommended.

3.6.1.2 Detector Shapes and Sizes

As mentioned previously, a tilted square detector presents very interesting behaviours with regard to spatial sampling. A detector with side dimensions measuring at least 10 pixels is recommended. The magnification factor of the test pattern may be adjusted appropriately to fulfill the 10-pixel criteria. However, one must keep in mind that the distance between the projector and the screen has to be large in comparison to the diameter of the exit pupil of the projector. This is required for the validity of the calculation with the BRDF. If a square detector is not available, a

round detector is also acceptable, provided that its diameter and magnification factor are such that the area perceived by the sensor is at least 15-pixels wide. For both types of detectors, the centering with respect to the pixel is not critical. For the tilted square detector, the tolerance range for the orientation is about $\pm 5^\circ$; this degree of accuracy can be achieved without the help of an instrument.

If it is not possible to find a detector that is both fast enough and large enough to fulfill the 10- or 15-pixels criteria, spatial averaging must be done. In that case, the detector must be shifted over a minimum 10-pixel size area of the image and the mean value over a large set of positions (approximately 50) must be taken.

3.6.1.3 Detector-System Cut-Off Frequency and Sampling Rate

In the case of the linear SLM projector it is better to have a detector that can follow the input optical signal. To achieve that goal, a detector cut-off frequency of about:

$$\frac{5NF_r}{1+n_d} \quad (3.15)$$

is probably enough, N being the number of pixels in the scan direction, F_r the frame rate and n_d the size in pixels of the detector along the scanning direction. For this type of projector, a sampling rate of the detector signal given by:

$$\frac{100NF_r}{1+n_d} \quad (3.16)$$

would probably be enough.

In the case of the flying-spot projector, a cascade of two detectors would be more appropriate for the measurement. The two detectors are placed close to each other along a common scanning line (in the middle of the same square of the checkerboard pattern). The first detector is a pulse detector used to measure the energy contained in the pulse part of the optical signal, while the other one is a photodetector used to measure the mean energy in the rest of the signal. The photodiode is used to measure the veiling glare parasitic light. Photodetector cut-off frequencies in the 100–1,000 Hz range and sampling rates in the 10–20 kHz range would probably be enough.

3.6.1.4 Spectral Measurements

Luminance measurement requires each spectral component to be weighted by the corresponding spectral response of the human visual system. For the proposed method, this involves isolating and analyzing each individual laser line separately. For this purpose, interferential filters optimized for each laser line are recommended. The transmission of each filter must be measured using a precision

spectrometer. It is important that the filters have high rejection properties (optical density ≥ 4) for the spectral regions outside their respective waveband. The filters must be placed near the detector, but oriented to reflect light toward the projector (normal incidence). For maximum efficiency, placing the filters in a rotating wheel enables rapid filter substitutions.

3.6.1.5 Reduction of Measurement Errors Due to Parasitic Light

As mentioned previously, back reflection and ambient illumination can affect image brightness and contrast. Such signal contamination changes from place to place, posing a problem for standard measurements. Fortunately, simple techniques exist to overcome such problems. The contribution of parasitic light may be measured, and then subtracted from the useful signal. This is done by putting the detector in the shadow of a black opaque mask. The light from the projector (including veiling-glare light) is blocked, and only parasitic light can reach it, as explained in Boynton and Kelley (2001).

3.6.2 Step 2: Measurement of the BRDF

The measurement of the BRDF of a surface requires specialized equipment. The development of such equipment would require significant effort, and the use of existing services is recommended. BRDF is difficult to measure, since it is a four-parameter function or even a five-parameter function if spectral properties are required. The measurements are generally done using a gonireflectometer that allows rotation of the detection direction of regard and the rotation of the sample itself or the incident direction of the light source. The measurement is time consuming because the four or five dimensions space has to be sampled and the raw data requires large storage space. Moreover, the BRDF measurement services offered by few specialized labs are relatively expensive. Fortunately, the interest on BRDF goes beyond scientific domains and reaches other domains such as animation cinema where it is used to make objects appear to be more realistic. Motivated by increasing demand, developments in BRDF measurement methods aim at making measurement more affordable.

3.6.3 Step 3: Data Processing

Luminance is computed by inserting the measurements for each individual projector laser line into (3.10). The contrast is computed with (3.12) or with (3.13) in the case of a perfectly white, or grey, screen. A computer program can be useful in calculating luminance and contrast, and performing fits and interpolations of the

BRDF data. It can also be used to easily and rapidly explore the results for many different observer positions, and produce result maps and figures.

3.7 Conclusion

This paper presented a method for the characterization of the brightness and contrast of images projected by LBVP. It was proposed with an aim of circumventing the reliability problems of the conventional methods. The proposed method implies the measurement of the irradiance of the LBVP image followed by the computing of the radiance or the luminance of the reflected beams by using the measured screen reflectivity properties. The procedure described in this document is based on theoretical considerations and has not been tested. It is essential to submit the proposed method to a battery of tests before attempting to use it for the evaluation of projector performance. If the proposed method works, it should be valid not only for LBVP but also for other projector types that can be adequately characterized with conventional methods. An interesting test that could be undertaken is to characterize a conventional projector (CRT or LCD projector) with both the proposed method and an appropriate conventional method.



Chapter 4

Physics Based Simulation of Light Sources

**Jeff Clark, Brad Colbert, Karl Mathia,
and Brett Chladny**

Abstract Light sources inherent in urban populations, battlefield effects, and vehicles are the dominant features in a simulated night scene. The effectiveness of optical sensors in these environments is positively or negatively affected depending upon the quantity, type, power, placement, housing, and orientation of these sources. Yet, traditional simulation techniques have focused little on the accurate representation of light sources and their interaction with other scene elements. Traditional lightpoint rendering techniques, while addressing modest out-the-window training requirements, fall well short of recreating the challenges associated with the employment of image intensifying sensors for situational awareness in today's complex mission environments. The elevated priority of low-altitude, urban missions such as close air support (CAS) and combat search and rescue (CSAR) will require richer, more accurate, and more dense representations of light sources if the sensor-related elements of these missions are to be trained in simulators.

Leveraging related research and development, ADDNS Phase II development efforts will integrate point source modeling techniques pioneered under collaborative research and development. The presented paper will summarize the core concepts employed to model urban lighting with the density and accuracy desired for sensor training applications. Secondly, a proof-of-concept, open architecture XML storage structure for point source modeling parameters will be summarized. Finally, the results of these techniques will be presented in an application example demonstrating the offline prediction of point source types, power, and location from remote sensing data as well as the production of physics-based terrain irradiance maps for simulation.

J. Clark (✉), B. Colbert, K. Mathia, and B. Chladny
Renaissance Sciences Corporation, 1351 North Alma School Road, Suite 265, Chandler,
AZ 85224, USA
e-mail: jeffclark@rscusa.com

Sommaire Selon les circonstances, les principales sources de lumière d'une scène nocturne sont l'éclairage urbain, les effets des combats et les véhicules. Dans les environnements visés, l'efficacité des capteurs optiques est considérablement et diversement affectée, positivement ou négativement, par la quantité, le type, l'intensité, la position, l'habillage et l'orientation de ces sources. Or, les techniques classiques de simulations reproduisent avec peu de précision les sources de lumière et de leur interaction avec les autres éléments de la scène. Bien que, dans le passé, elles aient été suffisantes pour les besoins modestes de formation aux scènes hors de l'aéronef, les techniques classiques de rendu des sources ponctuelles laissent beaucoup désirer au plan de la reconstitution des difficultés d'utilisation des capteurs à amplification d'image aux fins de la connaissance de la situation, dans les environnements complexes des missions actuelles. Actuellement, on accorde une forte priorité aux missions à basse altitude en terrain urbain comme l'appui aérien rapproché (AAR) ou la recherche et sauvetage de combat (RESCO). Or, si nous voulons assurer avec nos simulateurs, une formation aux éléments de ces missions qui requièrent l'utilisation de ces capteurs dans ces environnements très complexes, nous devons disposer de représentations des sources de lumière beaucoup plus riches, précises et denses.

Tirant parti des travaux connexes dans ce domaine, les travaux de développement de la phase II intégreront les techniques de modélisation des sources ponctuelles, mises au point par la collaboration en R et D. Nous décrirons les concepts centraux de la modélisation des sources ponctuelles avec des détails suffisamment fins pour produire des modèles physiques de rayonnement nécessaires à la reproduction de la complexité de la scène, avec la précision nécessaire pour son utilisation lors de la formation avec les capteurs. Ensuite, nous présenterons la structure de stockage en architecture ouverte XML produite aux fins de la validation de principe. Pour finir, nous présenterons les résultats de l'utilisation de ces techniques dans un sommaire de travaux connexes, démontrant les prédictions hors-ligne du type, de l'intensité et de la position des sources ponctuelles dans des champs denses d'éclairage artificiel ponctuel, ainsi que la production de cartes de l'éclairage du terrain, fondées sur la physique, pour les simulations.

Early aircrew training systems emerged as devices chiefly intended to train the operation of the aircraft and its flight systems. In these devices, the fidelity of the visual systems carried a low priority. In fact, the earliest devices were fielded without any visual representation and were found to be rather useful for their design purpose. As training systems technology advanced, so did the scope of training objectives that influenced the design of flight simulators. In time, more advanced generations of training systems were employed to train advanced mission skills in addition to aviation skills. In contrast to early systems which simply trained basic operation of the aircraft, mission training systems were required to train mission-specific flight skills and the employment of those weapon systems critical to mission execution. Today, the technological foundation of our most advanced mission training systems reflects decades of iterative improvements aimed at better recreating those

missions which were of the highest priorities in their day. For very nearly the entire history of mission training systems technology development, the mission that has dominated flight simulator design and development has been air-to-air; day air-to-air to be more precise.

The change in shape of contemporary conflicts and opposing forces, however, has seen the diminishing priority of the day air-to-air mission in favor of air-to-ground engagements, low-altitude flight profiles, and increasingly complex and urban environments. Moreover, night missions are increasingly becoming the operational norm across many mission types. The elevated priority of night air-to-ground, close air support (CAS), combat search and rescue (CSAR) and similar missions has created capability gaps in our training systems which correspond to technology gaps across a broad range of the modeling and simulation discipline.

Within this context, this article focuses on the synthetic environment elements of these technology gaps and, more specifically, on the modeling and simulation of cultural lighting, arguably the dominant element of a synthetic night urban environment. It is reasonable to anticipate that the ability to represent synthetic urban night environments with accuracy, density, and detail will be a key differentiator among next-generation training systems which seek to train night air-to-ground mission skills and weapon systems rather than foundational aircraft skills.

4.1 Background: The State of Fielded Training Systems Technology

The premise put forward in the introduction suggests that the cultural lighting representations in a typical fielded training system lack the realism required to replicate the real-world challenges associated with operating among complex urban lighting scene elements and backgrounds in NVG-aided and unaided scenarios. As background, the following sections will summarize the typical fidelity among fielded systems and suggest areas for improvement. The discussion of cultural lighting realism can be divided into three primary scope areas: point-source modeling, modeling of reflected light, and population of cultural lights.

4.2 Modeling of Point Sources

The physical modeling of the quantity, directionality, and spectral makeup of energy radiated from simulated point sources is the foundation of an accurate cultural lighting simulation. That is, sophisticated radiative models or high density placement models will be of limited utility if the radiation created by each single type of light cannot be quantified. While there are certainly exceptions, the modeling

fidelity of the typical visual system, in this regard, has been limited to a normalized tri-color stimulus model (i.e. RGB) consistent with the capability of the typical display system. Modeling of directionality of light types can generally be expected among typical visual systems. In those cases where the scope of the visual system is limited to direct viewing of out-the-window (OTW) representations, it can be reasonably argued that this fidelity of point source modeling matches the fidelity of the display system and is, therefore, sufficient. After all, the complete dynamic range of luminance values which would be produced by a quantitative point source model cannot be accurately represented on training systems displays. However, most mission simulators would benefit from a quantitative representation of point sources that lie within the luminance ranges supported by OTW displays. As examples, truly accurate modeling of the visibility of aircraft wingtip lights, vehicle lights, and urban lighting at specific distances would be a useful enhancement to many fielded training systems. The most compelling justification, however, for physics-based simulation of point sources becomes apparent when the requirement for a correlated image intensified scene (i.e. NVG) is assumed. NVGs can respond to emissive sources at great distances and can be overwhelmed by bright sources. It is desirable to offer accurate and OTW-correlated NVG representations in both cases. To do so requires a quantitative and multi-spectral point source radiation model.

4.2.1 Modeling of Reflected Light

In the preceding section, we discussed the value of accurately simulating the quantities of light that reach the observer eye point via a direct, unobstructed path. Here, we are interested in the light that originates with the cultural light sources (points) but would, in the real world, reach the eye point from many additional paths. While traditional flight simulators have put little emphasis on accurately modeling reflected light, it is a significant source of light. In many real world scenes, in fact, a greater quantity of light reaches the eye point from reflections than from direct viewing of the sources.

When we intend to use simulated scenes as input to physics-based NVG simulations, we are particularly interested in the accuracy of the simulated light levels (from all paths) across the entire scene. This, because of the fact that many electro-optical (EO) sensors, NVGs especially, perform differently as a function of their total radiance input. In this regard, the light “lost” by not accurately estimating reflected light meaningfully diminishes the realism of an NVG simulation. The impact of the realism of reflected light carries additional significance with respect to the visual realism of the scene as a whole. As is illustrated by the two comparative night OTW scenes, the addition of a ground plane cultural lighting representation provides a substantial amount of additional information about the simulated environment that would be available in the real world (Fig. 4.1).

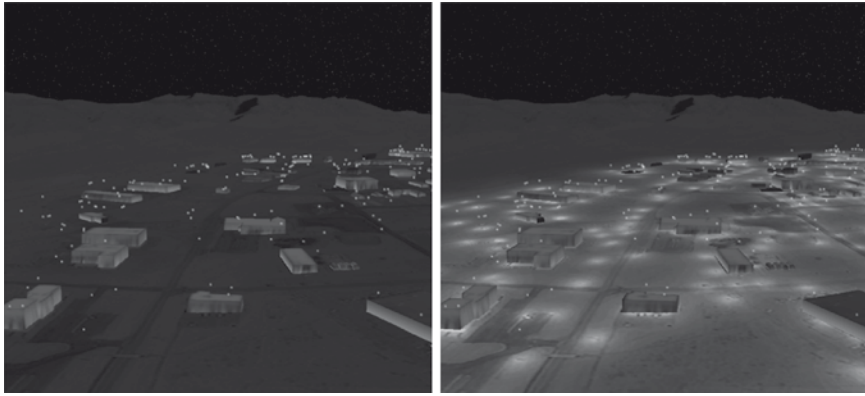


Fig. 4.1 Out-the-window scenes with (*right*) and without (*left*) ground plane lighting effects

The second image exploits a ground-plane cultural irradiance map produced using some of the techniques summarized later in the article. The first image, unfortunately, represents a level of fidelity typical of many fielded aircrew training systems. The scene detail benefits of the second image are apparent in this OTW scenario, but the quantitative benefits of the cultural irradiance information is, in fact, essential if a physics-based NVG simulation is to be employed with meaningful realism.

4.2.2 Modeling of the Placement of Cultural Lights

Having summarized the impacts of point-source modeling and reflected-light modeling, the accurate modeling of the placement of simulated cultural lights can be identified as a third goal independent of the modeling of the light they produce. The random scattering of graphical lightpoints across an urban area has been a popular approach in the past but does not result in a particularly believable result. In the past, the computing power traditionally available to image generation has limited the density of simulated lights, so a more realistic placement of lights would have exceeded the rendering budget available to most simulations. The advancements of modern GPUs, however, allow us to consider the value of far more realistic placement and density of simulated cultural lights.

It can be argued that, in the night environment, the magnitude and boundaries of urban light fields should serve as useful aids to situational awareness in much the same way that the accurate placement of natural landmarks (such as rivers or terrain features) contribute to the validity of the synthetic environment. In fact, aviation navigation charts include urban lighting outlines that are readily recognizable in flight but rarely present in simulation. Beyond fundamental visual realism and situational awareness, the complexity of many urban mission tasks, especially on NVGs, is difficult to represent without more closely matching the density and placement of real-world cultural lighting.

4.2.3 Physical Cultural Lighting Data in the Public Domain

Initial research focused on the collection of data from the public domain that might provide a foundation for a physics-based approach for the modeling of cultural lighting in real time visual simulation applications. As is often the case, it was found that richer descriptive data can be found among those communities that most directly deal with the subject technology. In this case, we found that architectural, manufacturing, traffic regulation, and GIS communities maintain physical data that would be useful to the modeling and simulation community. A summary of those findings appears in the subsections that follow.

4.2.4 Appearance of Cultural Lighting Objects

Cultural lighting objects here are referred to as “luminaires,” a term adopted by the manufacturing community. A luminaire is defined (Ode 2003) as a complete lighting unit consisting of a lamp or lamps together with the parts designed to distribute the light, to position and protect the lamps and ballast (where applicable), and to connect the lamps to the power supply. Noted key parameters for the appearance of luminaires include the light source technology, optics and light distribution pattern, rated power and spectral power density (SPD), and the luminaire’s location with respect to its environment.

The “light source technology” here refers to the physical process which generates the electromagnetic waves. In this context, the visible and near-infrared ranges of wavelengths are of particular interest. The common technologies that were considered in the research included incandescent, fluorescent and high intensity discharge (HID) light sources. HID sources include low and high pressure sodium lamps. For outdoor lighting, fluorescent and low pressure sodium are predominant. The power consumption and SPD together allow the prediction of the multi-spectral radiance produced as well as the visible color appearance of a light source. The availability of power and SPD data meaningfully enhances the physics-based modeling accuracy that is possible.

4.2.5 Radiative Properties of Cultural Lighting Objects

The combination of distinct elements of the luminaire, including its lamp, optics, and housing, determine its light distribution pattern. For streetlights there are five types of distribution patterns specified by the Illuminating Engineering Society of North America (IESNA) and the American National Standards Institute (ANSI). (See IESNA 2000a,b). The classes determine the shape of a light pattern on the ground. For example, Type V is circular. The other types are ellipses, with Type III being the most commonly available pattern.

To enable a physics-based simulation, however, magnitude and spectral distribution of radiated energy, in all directions, is required. The lighting industry has various

standards for the electronic dissemination of photometric information for lighting fixtures and other light sources. This data largely contains the visible-wavelength radiative patterns of specific light fixtures (luminaires) and is used in, for example, producing highly detailed renderings of architectural structures and roads. There are a number of software packages popular in these industries for rendering detailed scenes with light models that are parametrically defined by this data. The formats used to package this data vary from region to region but essentially describe similar information. The following list summarizes a selection of the most widely used formats.

- BSR/IESNA LM-63-2002 (IESNA standard file format for electronic transfer of photometric data). Used extensively in North America; accepted as an approved ANSI standard in August 2002.
- CIBSE TM-14:1988 (CIBSE standard file format for the electronic transfer of luminaire photometric data). Used extensively by British lighting manufacturers.
- CIE 102-1993 (recommended file format for electronic transfer of luminaire photometric data). A world-wide standard developed by the Commission internationale de l'éclairage (CIE) which is not widely used.
- EULUMDAT. The de facto standard for European lighting manufacturers.

4.3 Placement of Cultural Lighting Objects

For our initial investigation, we focused on the research of public domain placement data for street lighting as the dominant light source in many urban scenes. It was found that the type and placement of street lighting objects, conveniently, is heavily regulated in North America and Europe largely due to an increasing emphasis on safety, energy management, and light pollution. Most municipal and county ordinances specify many of the parameters that would be useful to a physics-based lighting simulation, and those parameters are specific as a function of street category and zoning. For example, streetlights are often placed at a height of 30 feet for arterials, 20 feet for collectors, and 17 feet for local streets and pedestrian areas. The associated distance between streetlights can be approximated at twice the height: 60 feet for arterials, 40 feet for collectors, and 34 feet for local streets and pedestrian areas. Such detailed guidelines in conjunction with GIS data, for example street categories, facilitate the modeling of street lights and other cultural lighting objects. As an example, Table 4.1 summarizes rules for street light type and placement taken from an ordinance associated with El Paso, TX.

4.3.1 A Proof of Concept

The subject research, performed in 2006, sought to leverage techniques used in other technical disciplines and extend and combine those techniques to pursue higher levels of scene complexity and accuracy in multi-sensor visual simulation

Table 4.1 Cultural lighting guidelines for urban street hierarchy

Street category	Light technology	Power (watts)	Height (feet)	Spacing (feet)	Angle (deg)
Arterials	LPS (50%)HPS (50%)	200	30	60	45
Collectors	LPS (50%)HPS (50%)	150	20	40	45
Local streets, pedestrian	Mercury vapor	80	17	34	45
Arterial intersections	HPS (50%)	310	36	N/A	45
Collector intersections	HPS (50%)	200	30	N/A	45

LPS low-pressure sodium lamps, *HPS* high-pressure sodium lamps

applications. A primary focus of the research was the investigation of the application of physically accurate cultural lighting content and its impacts to visual fidelity and accuracy. The following three sections summarize, across the same three modeling domains offered in the Background, approaches applied to the resulting technology demonstration.

4.3.2 Modeling of Point Sources

The simulation software read a custom data file which contained the light positions in latitude, longitude, height above ground, and wattage integrated over the wavelengths viewable by the simulated observer. This was used at run-time to place custom light nodes in the scene graph so that a point representation was rendered. The cutoff for each light was simulated by comparing the cutoff angle of the point source to the view angle. As a simplification, no 3D model was loaded to represent the luminaire housing or mast.

Using a GPU shader-based computational implementation, at-aperture radiance values were calculated, with 16-bit floating point numerical precision, for the NVG or visible bands in real-time for each visible light point and rendered as a function of light type, wattage and distance.

4.3.3 Modeling of Reflected Light

The modeling of the reflected energy was performed in two steps. First, a ground plane irradiance image was created that contained quantitative irradiance values produced, at the ground plane, by each geospecific light in the dataset. Each value was calculated using a simple radiative model modulated by a ground-plane lobe distribution function. These images were stored as geotiffs at a resolution equivalent to the source reflectivity and out-the-window imagery (about 1 m per pixel). The irradiance values were encoded as normalized logarithmic values to reduce the memory footprint (from 32 to 8 bits) while still maintaining a useful dynamic range ($10E-11$ to $10E-4$ W/cm²). This image set was stored as a texture layer in the

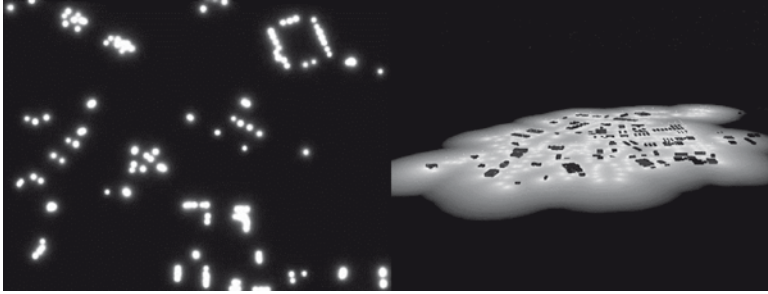


Fig. 4.2 Irradiance maps: raw image (*left*), logarithmic encoding and perspective projection (*right*)

run-time database. At run-time, the texture was decoded by a GPU shader which modulated the value with the underlying reflectivity layer (NVG or OTW, as appropriate) and then calculated an at-aperture radiance value from this. Figure 4.2 shows a 2D orthographic rendering of the stored irradiance map (without terrain texture for illustrative purposes) followed by a 3D perspective rendering of the same irradiance map rendered with geospecific building models.

4.3.4 Modeling of the Placement of Cultural Lights

In this experiment, the placement and specification of the light types was digitized, by hand, using day aerial imagery with resolution sufficient to discern the positions of the light sources. The light types were subjectively determined. Approximately 700 lights were present in the resulting dataset. While the potential for large-scale prediction of light types and positions from other available geospatial data appeared obvious, this step was not taken under the subject research in favor of focusing on the development of lighting representations. Follow-on efforts have yielded successful results in this regard.

4.3.5 Application Description

A C++ application was developed which comprised the simulation of the point light sources and the associated ground-plane reflected light using the techniques described in the preceding subsections. The scene graph used for the test application was Open Scene Graph. One benefit of using this API was that it facilitated multi-texture terrain generation which was required by the GPU shader-based computational techniques employed. The run-time database was stored and loaded using the Open Scene Graph native database format. In the run-time application, a shader was

placed on the terrain to interpret and combine irradiance and reflectivity texture layers to solve for at-aperture radiance. Each frame of simulated at-aperture radiance was then processed by a human-visible or NVG observer model, as applicable.

4.3.6 Results

The employment of these technical approaches and the utilization of the previously summarized public domain data yielded compelling results. The results of the experimental work showed that physics-based cultural lighting and effects could be produced with existing GPU and scene graph technologies. Subjective assessments showed that visual fidelity and NVG realism were meaningfully enhanced as compared to traditional real-time techniques. Figure 4.3 shows OTW and simulated NVG images captured from the resulting interactive application.

4.3.7 Next Steps

The realism benefits of the featured techniques are apparent, but additional steps must be taken if these techniques are to be implemented broadly within traditional visual simulation production pipelines. The following sections outline the some fundamental steps that may be taken to advance these concepts into broad application.

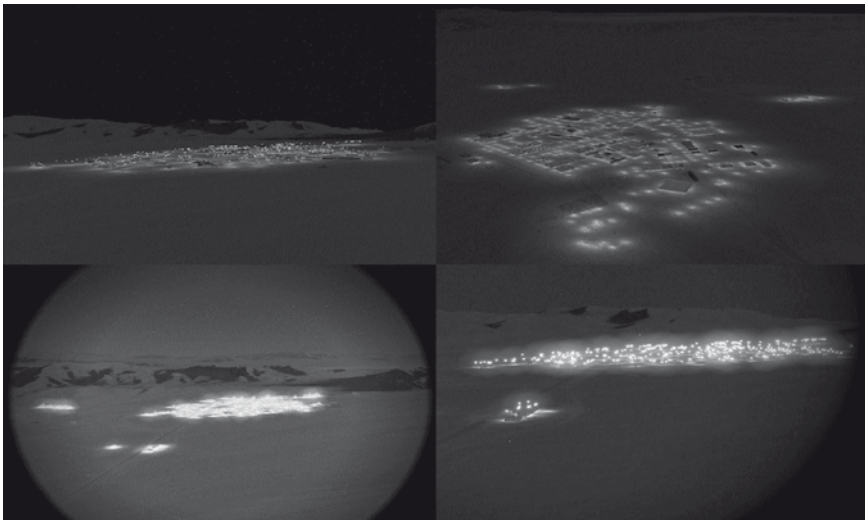


Fig. 4.3 OTW and NVG experimental results

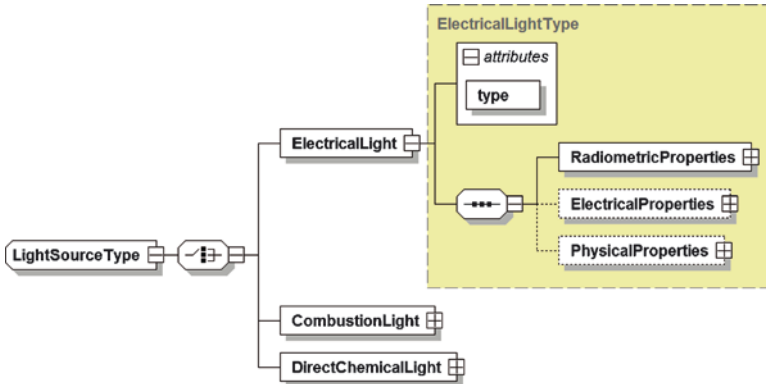


Fig. 4.4 Sample of XML schema for describing cultural lighting systems

4.3.8 Integration of Visual Simulation and Lighting Formats and Standards

The ability to reference a machine-readable physical luminaire definition from machine-readable simulation database formats is a first fundamental step required in order to make broad use of detailed lighting data in database modeling tools and simulation runtime applications. Such a mechanism would enable the automated production of irradiance maps within popular database modeling tools and the utilization of physics-based point source representations within runtime applications.

As an example, it is not necessarily unusual to encounter the population of light-point type attributions within Openflight using either the FID (Feature Identifier), SMC (Surface Material Code) or even in the user comment fields. The attributed ID, however, rarely references physical descriptions with the depth required to satisfy the inputs to models described in this article. Such attribution, when paired with an open standard for robust description of luminaire physics would yield more realistic and more consistent simulation results. Adopting the XML conventions specified by the NASMP Portable Source Initiative (NPSI) metadata specification, a physics-based cultural lighting metadata schema was drafted to illustrate this concept. Figure 4.4 shows a graphical representation of portions of the schema produced during the research. Such an XML-based approach could be used to aggregate the various sources of public domain and manufacturing data into a single, machine-readable format for consumption by visual simulation applications.

4.3.9 Aggregation of Detailed Cultural Lighting Data Dictionary

Having integrated popular formats and standards typical in the visual simulation industry and architectural disciplines, users would benefit from the collection of

high-quality data into a common and consistent archive. The availability of an XML-based archival approach similar to that summarized in the previous section would facilitate the development of such a dictionary.

4.3.10 Simulation Database Tools

The availability of physics-based descriptions of cultural lighting in machine-readable formats would enable more automated attribution techniques in database modeling tools. Visual simulation databases of greater accuracy and complexity could be supported by new modeling tools that employ regional rule sets and detailed cultural lighting data to identify the likely type and placement of cultural lighting objects. Additionally, in some cases, it will be possible to directly employ municipal GIS datasets to populate geospecific cultural lighting as the research showed the availability of such data to be increasingly available.

4.3.11 Runtime Graphics Tools

Once the light sources have been attributed, run-time simulation software is needed to process robust varieties of light source type IDs. Such a software module should support processing all attributed light source for all wavebands. Inputs to this module should likely include the type ID, distance, environmental conditions, and desired waveband. The output should be color intensities for out-the-window and irradiance/radiance values for sensor applications.



Chapter 5

Integration of a Deployable CIGI-Based Image Generator in an Existing Simulation

Robert Wolfe and Tony Ghoman

Abstract An image generation system delivers the principal imagery in visual simulations. It may be split into its constituent software (image generator) and hardware (display) components. The image generator supplies both image rendering functions as well as informational mission functions. The system is often characterized by its ability to provide a certain frame rate for a given image resolution, scene complexity, and latency. The ADDNS project has produced an advanced image generator (IG) for flight simulation known as the Eagle IG. The large screen and visual acuity resolution of the Evans and Sutherland laser-based projector display make this system one of the first of its kind. The Eagle IG is based upon Multigen FlightIG image generation software, providing a network-based interface based upon the Common Image Generator Interface (CIGI). Here we discuss issues of the demonstration integration of the new IG, using CIGI, with the existing Advanced Distributed Mission Training (ADMT) Multi-Task Trainer (MTT) for the CF18, located at DRDC Toronto. In normal operations, the MTT employs a multi-channel MetaVR image generation system which display over eight flat projector screens in a dome arrangement.

CIGI is a network-based protocol created by Boeing that defines the format and content of CIGI packets but does not define how they are transmitted across a network. Boeing publicly released its original CIGI SDK in 2001 to promote the use of an open image generator interface for visual simulation. Development of CIGI support software continues today in the open source community. CIGI originates in the IG, perhaps based on the philosophy that absolute frame borders in the IG are fixed or very difficult to change given features such as genlocking. Ideally, the Host must be able to respond

R. Wolfe (✉)

Serenica Technologies Corporation, #2-7 Meadowbrook Rd., Toronto, ON, Canada M6B 2S3
e-mail: rwolfe@vif.com; Robert.Wolfe@drdc-rddc.gc.ca

T. Ghoman

Defence Research and Development Canada, Toronto, ON, Canada

to the IG with the data appropriate to the IG frequency. There are a myriad of issues that complicate the integration of any host simulation with an IG. These issues are complex when integrating with a host that is changing from one IG to another and more so when integrating side-by-side in a heterogeneous IG system. The approach taken for integrating the Eagle IG with the MTT is to create a side-by-side demonstration of the usual MTT visual with a replicated front channel on the Eagle IG using CIGI protocol to control the viewpoint only. A generic analytical approach for facilitating the integration includes scoping the existing IG installation, comparing with the new IG protocol, and substituting appropriate host information and messages in the new protocol. The complete integration of the MTT with the Eagle IG is far more complex than viewpoint integration and well beyond the scope of this work. The process of integration follows the steps of comparison of native MTT graphics messages with equivalent CIGI messages (which could be used to scope complete MTT integration), the assembly of well-ordered CIGI API source code calls, and the inclusion of MTT data in the CIGI messages.

To facilitate the integration, a side-by-side of CIGI equivalents to the ChannelSurfer messages used by the MTT was prepared. The important parameters of each of the CIGI Advanced Deployable Day Night Simulation Symposium message types are indicated. For ChannelSurfer messages, it is indicated if the message is of a generic type (header only), a minimum requirement for ChannelSurfer communications, or a corresponding message for completion of the protocol exchange. Pseudo-source code fragments demonstrate the sequence of CIGI calls required in communication with any IG. The calls are based on the freely available Boeing CIGI software, the same code exploited by Multigen to create their own host emulation application for FlightIG software. Additional pseudo-code snippets demonstrate the integration of MTT values with the properly sequenced CIGI calls within the MTT software. Different issues that may have complicated or facilitated the integration process are also discussed.

Sommaire Un système de génération d'images fournit les images principales pour les simulations visuelles. Il peut se répartir entre son logiciel constitutif (générateur d'images) et ses composants matériels (affichage). Le générateur d'images assure les fonctions de représentation d'images ainsi que les fonctions d'information de mission. Le système est souvent caractérisé par sa capacité à offrir une certaine cadence d'images complètes en fonction d'une définition d'image, d'une complexité visuelle et d'un temps de latence donnés. Le projet ADDNS a produit un générateur d'images avancé pour la simulation de vol appelé générateur d'images Eagle. Le grand écran et la définition visuelle du projecteur laser Evans and Sutherland connexe fait de ce système un des premiers du genre. Le générateur d'images Eagle se fonde sur le logiciel de génération d'images Multigen FlightIG, lequel offre une interface réseau tributaire de l'interface commune de générateur d'images (CIGI). Le présent document porte sur la démonstration de l'intégration du nouveau générateur d'images utilisant le CIGI de concert avec le simulateur multitâches d'instruction

de mission avancée à distance pour le CF-18 situé à RDDC Toronto. En opérations normales, le simulateur multitâche fait appel à un système de génération d'images MetaVR à canaux multiples avec affichage sur huit écrans plats de projecteur à l'intérieur d'un dôme.

Le CIGI est un protocole réseau créé par Boeing qui définit le format et le contenu des paquets CIGI, mais non la façon selon laquelle ils sont communiqués sur le réseau. Boeing a rendu publique sa trousse de développement de logiciels pour le CIGI en 2001 afin de promouvoir l'utilisation d'une interface de générateur d'images ouverte en simulation visuelle. La mise au point du logiciel de soutien du CIGI se poursuit aujourd'hui dans le milieu des sources ouvertes. Le CIGI découle du générateur d'images, sans doute fondé sur le principe que le cadre absolu des images complètes du générateur d'images est fixe ou très difficile à modifier, compte tenu de caractéristiques comme le verrouillage de synchronisation. Idéalement, l'hôte doit être en mesure de répondre au générateur d'images au moyen des données convenant à la fréquence du générateur.

Une foule d'éléments compliquent l'intégration de toute simulation d'hôte au sein d'un générateur d'images. Ces éléments sont complexes dès lors qu'il s'agit d'intégration à un hôte qui change d'un générateur d'images à l'autre, à plus forte raison lors de l'intégration juxtaposée à un générateur d'images hétérogène. L'approche adoptée pour intégrer le générateur d'images Eagle au simulateur multitâches consiste à juxtaposer la démonstration de la visualisation habituellement utilisée par le simulateur multitâche à une réplique du premier canal du générateur d'images Eagle au moyen du protocole CIGI pour contrôler le point de vue seulement. Une approche analytique générique visant à faciliter l'intégration consiste à déterminer la portée du générateur d'images existant, à la comparer avec le nouveau protocole du générateur d'images et à substituer l'information et les messages appropriés de l'hôte dans le nouveau protocole. L'intégration complète du simulateur multitâche au générateur d'images Eagle est beaucoup plus complexe que l'intégration du point de vue, et elle se situe bien au-delà de la portée de la présente étude. Le processus d'intégration suit les étapes de la comparaison des messages graphiques d'origine du simulateur multitâches par rapport aux messages CIGI équivalents (qui pourraient être utilisés pour déterminer la portée d'une intégration complète du simulateur multitâches), l'assemblage d'appels de code de source API bien ordonnés du CIGI et l'inclusion des données du simulateur multitâches dans les messages CIGI.

Afin de faciliter l'intégration, une juxtaposition des équivalents CIGI aux messages ChannelSurfer utilisés par le simulateur multitâches a été préparée. Les paramètres importants de chaque type de message CIGI sont indiqués. Pour les messages ChannelSurfer, on indique si le message est générique (en-tête seulement), une exigence minimale pour les communications ChannelSurfer ou un message correspondant destiné à la réalisation du protocole d'échange.

Les fragments de codes de pseudo-source fournis montrent quelle séquence d'appels CIGI est nécessaire pour communiquer avec n'importe quel générateur d'images. Les appels sont basés sur le logiciel libre CIGI de Boeing, le même

code exploité par Multigen pour créer sa propre application d'émulation d'hôte pour le logiciel FlightIG. D'autres bribes de pseudo-code montrent comment sont intégrées les valeurs du simulateur multitâches aux appels CIGI convenablement séquencés au sein du logiciel du simulateur multitâches. Différentes questions peuvent avoir compliqué ou facilité le processus d'intégration dont il est question ici.

DRDC Toronto is providing support to Canada's CF-18 training by offering options for future improvements to scene generation in combat simulators through the ADDNS project. One such demonstrable option is the system consisting of an advanced software image generator (IG) with host hardware for flight simulation (Array Eagle IG), and high definition Evans & Sutherland laser projector display (ESLP). The Eagle IG supplies both image rendering functions as well as informational mission functions, providing a network-based interface based upon the Common Image Generator Interface (CIGI). The visual acuity resolution of the ESLP and corresponding informational content of the Eagle IG promises to provide a more natural virtual environment. That is an environment where detection of distant detail or objects is not dependent upon display pixel size. Furthermore, the deployability of this system in a network-based environment provides modular flexibility in a distributed simulation.

We discuss the issues of the demonstration integration of the new IG, using CIGI, with the existing Advanced Distributed Mission Training (ADMT) CF-18 Multi-Task Trainer (MTT), located at DRDC Toronto. [Note that the reader will find a list of initialisms at the end of this chapter.] The creation of a side-by-side demonstration of the usual MTT MetaVR IG visual is described, with a replicated front channel on the Eagle IG using CIGI protocol to control only the viewpoint. Native MTT graphics messages (ChannelSurfer) are compared with equivalent CIGI messages (which could be used to scope complete MTT integration). Well-ordered CIGI API source code calls are assembled, and MTT data included in the CIGI messages. Pseudo-source code fragments based on the freely available Boeing CIGI software demonstrate the sequence of CIGI calls required.

The ADDNS project has produced an advanced image generator (IG) for flight simulation known as the Eagle IG, consisting of several compute nodes each producing one tile of the aggregate display output. The large screen and visual acuity resolution of the associated Evans and Sutherland laser-based projector display make this system one of the first of its kind. The Eagle IG is based upon Multigen FlightIG image generation software, providing a network-based interface based upon the Common Image Generator Interface (CIGI). Here we discuss issues of the demonstration integration of the new IG, using CIGI, with the existing Advanced Distributed Mission Training (ADMT) Multi-Task Trainer (MTT) for C-18 located at DRDC Toronto. In normal operations, the MTT employs a multi-channel MetaVR image generation system with display over eight flat projector screens in a dome arrangement.

5.1 Background

An image generation system in simulation is often characterized by its ability to deliver a certain frame rate for a given image resolution, scene complexity, and latency. It may be split into its constituent software (image generator) and hardware (display) components.

The image generator supplies both image rendering functions as well as informational mission functions. Image rendering functions may include different terrain, model, and environmental features, and hardware and performance support features. Mission functions provide time-critical information that logically should reside with the image generator given its detailed knowledge of the terrain topology and all dynamic models. These may include functions for collision detection, distance ranging, or intervisibility.

The display system is the final stage in generating the light intercepted by the human eye. It is often characterized by such parameters as screen size, resolution, brightness, refresh frequency, and contrast ratio, although there may be many others depending upon the particular technology employed (LCD, plasma, laser, etc.). The many IG features complicate the integration of any host simulator with an IG. In particular, differences in IG and host frame rates and coordinate systems exacerbate the problems.

5.2 MTT Visual Functions

The MTT relies on several components to provide all the information required to appropriately render the virtual environment. The central component is the CF18 host emulation system but this also relies on the Mission Function Server (MFS) based at least in part on the MDX read API, the Next Generation Threat System (NGTS), and the MetaVR IG subsystems through the ChannelSurfer protocol. The NGTS only serves weapon information to the host emulation system. The MFS provides intervisibility and collision detection functions which are used by the host's interface to the IG. Communication between the host and MetaVR IG subsystems is managed in the PCIG program of the MTT VME host computer.

5.3 MetaVR Image Generator

The MetaVR IG is a DirectX-based renderer which supports DIS and its own native network-based ChannelSurfer Protocol, simultaneously if desired. The MetaVR MDX read application programming interface (API) provides time-critical mission function information quickly within one frame period. The MetaVR Virtual Reality Scene Generator (VRSG) Dashboard provides configurability of the renderer. The dashboard allows the user to specify the display hardware employed for each

channel, the DIS exercise messages that will be interpreted, and a host of parameter values that be used as defaults by the interacting host application.

5.4 Flight IG

Multigen FlightIG software runs in the Eagle IG within a Linux (CentOS) and OpenGL-based environment. It uses the VegaPrime scene graph with large area database management and texture paging algorithms with support for virtual texture. It provides real-time mission functions for line of sight, height of terrain, height above terrain, and collision detection. The host communicates with the renderer using CIGI or FlightICD protocol. In the case of multiple rendering nodes, CIGI commands are synchronized through a master/slave arrangement.

5.5 Common Image Generator Interface (CIGI)

Although a complete description of the CIGI protocol is beyond the scope of this document, a brief discussion here should help clarify its deployment of the protocol in the MTT.

CIGI is a network-based protocol which defines the format and content of CIGI packets but does not define how they are transmitted across a network except to say that the content format is big-endian. CIGI originates in the IG, perhaps based on the philosophy that absolute frame borders in the IG are fixed or very difficult to change given features such as genlocking. Ideally, the Host must be able to respond to the IG with the data appropriate to the IG frequency. Asynchronous communication arises if the Host communications occur at a non-adjustable or inflexible rate. CIGI employs a geodetic (such as WGS84) entity positioning system with entity coordinate axes in the Local North, East, and Down (x, y, z with respect to a local geodetic tangential plane) directions and rotations defined in the same order (heading, pitch, roll) about those axes.

CIGI defines four principal modes of operation of the IG: Reset/Standby, Operate, Debug, Offline maintenance. One sequence of events leading to changes in mode of operation is described below:

- IG initialization including loading default terrain/models
- IG sends SOF packet with IG Mode=Reset/Standby (IG is mission-ready)
- Host responds with IG Mode=Operate in the IG Control data packet
- IG responds with SOF packet with IG Mode=Operate
- Host sends packets in addition to IG Control
- Host eventually sends IG Mode=Reset/Standby
- IG removes any entities created during the previous session

Table 5.1 CIGI IG-to-Host packets

CIGI IG-to-Host Packet (x = partial FlightIG support)	
Start of frame (SOF) (database, status, mode, counter, timing value)	
Image generator message (message)	
Sensor response (view, status, sensor, target offset and size)	x
Collision detection segment response (entity, segment, entity/non-entity collision, contacted entity, material, segment collision location)	
Collision detection volume response (entity, volume, entity/non-entity collision, contacted entity)	x
Height above terrain response (valid, material, altitude)	
Height of terrain response (valid, material, altitude)	
Line of sight response (valid, occulted, material, range, intersection location)	
User-definable data (packet ID 236–255, packet size, user-defined data)	

The following lists the different CIGI-defined packets with exceptions to the standard as implemented by the FlightIG software. Other important restrictions of the specific FlightIG implementation of the CIGI protocol include the limitation per frame to one UDP packet of maximum 64 KB size containing several CIGI packets but no more than one IG Control packet. The basic parameters of the CIGI messages are described in parentheses (Tables 5.1 and 5.2).

5.6 Integration of the ADDNS Image Generation System with the Multi-Task Trainer

The approach taken for integrating the Eagle IG with the MTT is to create a side-by-side demonstration of the usual MTT visual with a replicated front channel on the Eagle IG using CIGI protocol to control the viewpoint only. This provides a good demonstration and verification tool. It is important to realize that the side-by-side installation of multiple heterogeneous IGs in a same host simulation is the worst case scenario for integration. The complete integration of the MTT with the Eagle IG is far more complex than viewpoint integration and well beyond the scope of this work. The process of integration follows the steps of comparison of native MTT graphics messages with equivalent CIGI messages (which could be used to scope complete MTT integration), the assembly of well-ordered CIGI API source code calls, and the inclusion of MTT data in the CIGI messages.

5.7 IG Message Comparison

To facilitate the integration, a side-by-side of CIGI equivalents to the ChannelSurfer messages used by the MTT has been prepared below. The items marked * indicate IG-to-host messages. Information in parentheses for CIGI messages indicates the

Table 5.2 CIGI Host-to-IG packets

CIGI Host-to-IG Packet (x=partial FlightIG support only, X=none)	
IG control (database load, mode change, tracker enable and boresight, counter and timing value) – timing value not supported in FlightIG; therefore, no asynchronous mode is available	x
Entity control (load, hide, attitude/position, animation, opacity, temperature, children, collision)	
Component control (requires IG-specific component identification scheme)	
Articulated part control (part relative position and orientation enables/values)	
Rate control (entity, child, or articulated part position and orientation velocities)	
Environment control (ephemeris/MODTRAN enables, date, humidity, air temperature and pressure, wind direction and speed, aerosol)	
Weather control (weather/scud/winds aloft random enables, entity and phenomenon air temperature, opacity, runway visibility range, scud frequency, coverage, elevation, thickness, transition bands, winds aloft speed and direction)	
View definition (view group, view type, pixel replication mode, mirror mode, tracker assign, FOV near/far, left/right, up/down) – view group not supported in FlightIG	x
View control (view group, entity relative position and orientation enables/values) – view group not supported in FlightIG	x
Sensor control (on/off, polarity, line-by-line dropout, track mode, black/white tracking, autogain on/off, gain, level, AC coupling, and noise, used in coordination with View control, View definition, Component Control, LOS Range Request)	
Trajectory definition (acceleration factor (gravity), retardation rate (drag), terminal velocity, used in coordination with Entity control, Special effects, and Rate control)	x
Special effect definition (backwards/forwards animation, color override, new color, size scaling, time scaling, burst count, intraburst separation distance, burst rate, effect or burst sequence duration) – size scaling in FlightIG may be used to scale all entities	
Collision detection segment definition (low speed with contact location and material type) (enable, environment features mask, segment start/end)	
Collision detection volume definition (high speed with no contact location nor material type) (enable, volume centroid, size)	x
Height above terrain request (altitude, latitude, longitude)	
Height of terrain request (latitude, longitude)	
Line of sight occult request (source and destination altitude, latitude, longitude)	
Line of sight range request (heading, pitch, min/max distance, source altitude, latitude, longitude)	
User-definable data (packet ID 236–255, packet size, user-defined data)	x

important parameters of each of the message types. For ChannelSurfer messages, the parenthetic information may indicate if the message is of a generic type (header only), a minimum requirement for ChannelSurfer communications, or a corresponding message for completion of the protocol exchange (Table 5.3).

Table 5.3 CIGI–MetaVR comparison

CIGI (*IG-to-host, **actually used)	MetaVR ChannelSurfer (*IG-to-host, ***not used)
*Line of sight response	*LaserRangeMsg (LosRequestMsg or ClientUpdateMsg, lrfClient option)
*Height of terrain response	*TerrainHeightMsg (ClientUpdateMsg, aglClient option)
*Start of frame (SOF), **IG control	StartClientMsg (minimum requirement)
**Entity control	ClientUpdateMsg (minimum requirement)
*Start of frame (SOF)	*ReplyMsg (generic, minimum requirement)
IG control – mode change to standby/reset	ShutdownMsg (generic, minimum requirement)
IG control – mode change to standby/reset	RestartMsg (generic, minimum requirement)
**View definition, **View control	Viewport2UpdateMsg
Special effect definition, Entity control	ShowEffectMsg
Entity control	RemoveModelMsg
Entity control	ModelDataMsg
Environment control	*TimeOfDayMsg, ***WindDirectionMsg, **VisibilityMsg
Weather control	StartClientMsg (fogDensity, cloudHeight), *FogMsg, ***TimeOfDayMsg, ***WindDirectionMsg, ***VisibilityMsg
**Component control	various including ShowEffectMsg, ClientUpdateMsg, environment and weather-related

5.8 CIGI API Calls

The following pseudo-source code fragment (it will not compile as such) demonstrates the sequence of CIGI calls required in communication with any IG. The calls are based on the freely available Boeing CIGI software, the same code exploited by Multigen to create their own host emulation application for FlightIG software.

```
// Boeing CIGI API Host Order of Declarations/Calls for VxWorks PPC
// possible target architecture requirement
#define MOST_SIGNIFICANT_BIT_FIRST
// possible requirement for correct alignment of double field in a structure to double
// word boundary
#define CIGI_FORCE_DWORD_ALIGN
typedef long (* CIGICBPROC)(const int sessionid, void * packet);
int CigiInit(const int maxsessions, const int minver);
int CigiCreateSession(const int type, const int numbuffers, const int buffersize);
int CigiSetCallback(const int opcode, const CIGICBPROC proc);
int CigiSetIncomingMsgBuffer(const int sessionid, const unsigned char *buffer,
const int size);
int CigiProcessIncomingMsgBuffer(const int sessionid);
int CigiStartMessage(const int sessionid);
```

```

void CigiInitXXXPacket(CIGI_YYY* packet);
int CigiAddPacketXXX(const int sessionid, const CIGI_YYY* packet);
int CigiEndMessage(const int sessionid);
int CigiSyncFrameCounter(const int sessionid);
int CigiGetOutgoingMsgBuffer(const int sessionid, unsigned char **buffer, int
*size);
int CigiUninit(void);

```

5.9 Integration Logic of CIGI Calls with MTT

The following describes the logical leaving of CIGI-related calls within the MTT simulation code in the file `pcig.c` where the calls above originate.

- send and receive ChannelSurfer sockets are initialized
- send and receive CIGI sockets are initialized
- start of loop
- receive CIGI packets, if any
- process CIGI reply, if any received, including start-of-frame replies confirming IG mode of operation
- receive SurferChannel packets, if any
- process SurferChannel packets, if any received, including ReplyMsg confirming database load from each node, laser range reply, and terrain height reply indicating that the MTT MetaVR renderer has been initialized
- if nothing was received via either protocol or the IG was reset, go back to start of loop
- otherwise, assemble and send SurferChannel messages:
- startClientMsg on first time through the loop
- any other messages if the MetaVR renderer has been initialized
- assemble and send CIGI messages:
- configuration messages if IG is in operate mode and not already configured
- ownership entity control packet if the CIGI renderer has been configured and the MetaVR renderer has been initialized
- end of loop

5.10 Eagle IG/MTT Integration Issues

A network bridge and traffic sniffing facility was established between the MTT and Eagle IG by running a Knoppix V5.0 Linux rescue CD on a computer with dual network interfaces. This easily implemented arrangement tends to guarantee that the bridging computer is not tainted by any sensitive data (if it should actually exist) transmitted by the host or the IG. The built-in bridging modules of Knoppix were used to establish the network connection with the only hiccup elaborated below.

WireShark software was used to examine CIGI packets. This helped clarify byte and bit field orders for proper software compilation (Fig. 5.1).

Several issues were encountered that hampered the direct adoption of the Boeing CIGI software for use in the MTT.

- MTT code is compiled using an old unsupported Green Hills C language compiler on the Sun Solaris Operating System platform.
- Much of the associated Green Hills Multi Development Environment is not available and virtually none of the debugging features are available (C printf is essential).
- The MTT compilation platform (Solaris) and target processor(s) (PPC VxWorks 5.4 WIND Version 2.5) require cross-compilation and its associated complications.
- Boeing CIGI C source code is only available for the basic CIGI API.
- Boeing CIGI API C source code required modification to function on the target VxWorks platform with little available guidance on compilation macros.
- Some inefficient code in the Boeing CIGI API C source code was improved.
- Boeing CIGI host emulation software ported from C++ to C in order to exploit the script configuration file parsing still does not compile on the target VxWorks platform due to suspected header file problems (probably related to cross-compilation).
- C structures intended to directly mimic CIGI packet structure employ bit fields, requiring macros for cross-platform support of compiler-specific bit fields.

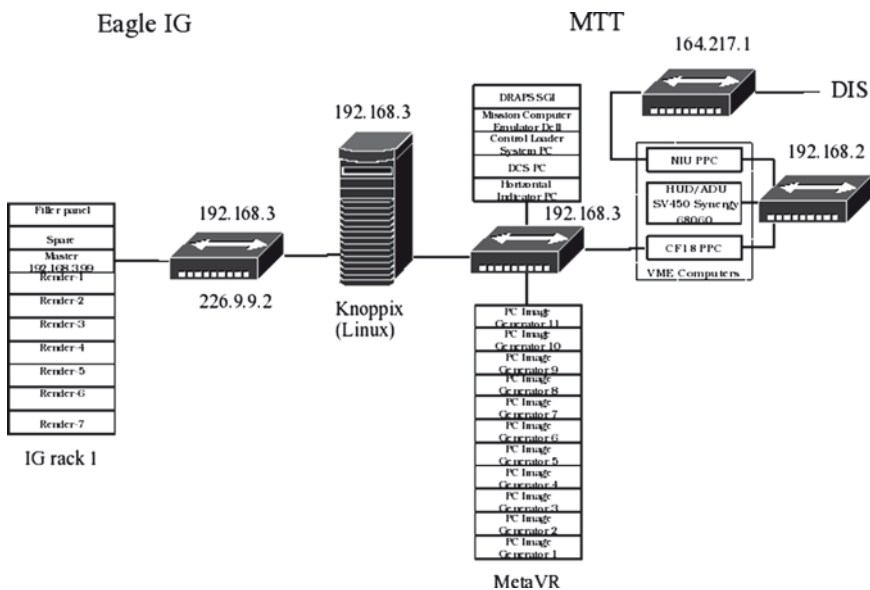


Fig. 5.1 Connectivity between the Eagle IG and MTT simulator

- The hardware configurable byte order of the PPC processors used under VxWorks was unknown.
- Inflexible MTT IP addressing required change to IP in the Eagle IG instead.
- Autonegotiation of Ethernet parameters between the bridge computer and MTT failed to provide the maximum possible 100 Mbps bandwidth. This was corrected manually each time the bridge was run.

Some compatible items that facilitated integration of the MTT with the Eagle IG include:

- Boeing CIGI C source code only covers CIGI major revisions 1 and 2 but ADDNS FlightIG only supports CIGI version 2 anyway.
- The C language is supported in the basic Boeing CIGI API source code and MTT.
- MTT host frame rate and IG frame rate match for all intents and purposes.
- Networking facilities were already in place since the ChannelSurfer protocol also relies on network communication for IG control.
- CIGI-compatible coordinates (geodetic) were available in the MTT code.

5.11 Conclusion

There are a myriad of issues which complicate the integration of any host with an IG. These issues are complex when integrating with a host that is changing from one IG to another and more so when integrating side-by-side in a heterogeneous IG system. A generic analytical approach for facilitating the integration includes scoping the existing IG installation, comparing with the new IG protocol, and substituting appropriate host information and messages in the new protocol. Specifically with the MTT, the process of integration of the Eagle IG follows the steps of comparison of native MTT graphics messages with equivalent CIGI messages, the assembly of well-ordered CIGI API source code calls, and the inclusion of MTT data in the CIGI messages.

The variety of IG features and implementations, and the continual push and pull between centralized and distributed computing and graphics hardware means that continual work is required to maintain a visual simulation on the leading edge of technology, even if an open IG interface standard can be maintained. For example, in the MTT, the adoption of the NGTS for weapons serving may lead to standardized weapon serving across multiple platforms but may also risk loss of terrain database correlation for intervisibility with the IG.

The consideration of hardware driven IG synchronization issues lead to the independent consideration of each of the software image renderer and hardware display. In the prototype laser projector system, the interface is restricted to digital video input (DVI) channels of 1,280 x 1,024 pixels each which offers less opportunity to exploit newer compute nodes and graphics cards. However, by examining the MTT display specifications, it was realized that there was opportunity to upgrade the MetaVR compute nodes and graphics cards to drive the existing display projectors at higher resolutions.

List of Abbreviations

ADCTS	Advanced Distributed Combat Training System
ADDNS	Advanced Deployable Day/Night Simulation
ADMT	Advanced Distributed Mission Trainer
AFRL	Air Force Research Lab
AMD	Advanced Micro Devices
CF	Canadian Forces
CIGI	Common Image Generator Interface
CPU	Central Processing Units
DB	Database
dB	Decibel
DIS	Distributed Interactive Simulation
DR	Distributed Rendering
DVI	Digital Video Interface
EIA	Electronics Industry Alliance
FOV	Field of View
GB	Gigabytes
GB/s	Gigabytes/sec
GDDR3	Check
GHz	Gigahertz
GNU	GNUs not UNIX (recursive)
GPL	Gnu Public License
GPU	Graphics Processor Units
GSP	Generic Signal Processor
HE	Host Emulator
HLA	High-Level Architecture
HW	Hardware
Hz	Hertz
ID	Identification
IG	Image Generator
IRIS	Integrated Raster Imaging System (from SGI)
IRIX	IRIS Unix (from SGI)
LADBM	Large-Area Database Module (Vega-Prime)
LCD	Liquid Crystal Display
LWH	Length, Width, Height
MB	Megabytes
MHz	Megahertz
M2DART	Mobile, Modular Display for Advanced Research and Training
MPI	Multigen-Paradigm Inc. (now part of Presagis)
n.d.	Not dated
OS	Operating System
PC	Personal Computer
PC-IG	Personal Computer Image Generator

PTT	Part-Task Trainers
ROPs	Raster Operations
RPM	Revolutions Per Minute
SGI	Silicon Graphics Incorporated
SW	Software
TDP	Technology Demonstration Project
U	Units (form factor from EIA standards)
UHR	Ultra-High Resolution
VA	Volt-Amps
VIDS	Visual Integrated Display System
W	Watts
WADS	Wide Area Distributed Simulation
WOL	Wake-On-LAN



Chapter 6

Advances in Scalable Generic Image Generator Technology for the Advanced Deployable Day/Night Simulation Project

Ray M. Obidowski and Rajesh Jha

Abstract The Advanced Deployable Day/Night Simulation (ADDNS) Technology Demonstration Project (TDP) contained a critical requirement for at least 16 synchronized Image Generators (IGs) to render scenes at 60 Hz for a greater than 20 million pixel Ultra-High Resolution laser projection system. Array Systems Computing Inc. (Array) adapted its advanced ScalableGSP (Generic Signal Processor) architecture to satisfy the ADDNS requirements. The Array ADDNS-IG solution is composed of a 16-node Personal Computer (PC) Linux cluster configuration, with each PC consisting of integrated Commercial Off-the-Shelf (COTS) hardware, featuring multiple Opteron Central Processing Units (CPUs), NVIDIA graphics cards, and supporting components.

To operate the hardware and use distributed rendering, Array integrated Multigen-Paradigm Inc. (MPI)'s Flight-IG runtime software and the Common Image Generator Interface (CIGI) under Centos Linux. During the development of the Array ADDNS-IG, a number of engineering and design challenges were overcome.

Additional enhancements currently being developed for the IG system include boot-off-network diskless slaves, a fully integrated network backup server, and an updated operating system build (Centos 5). These enhancements enable a reduction in the number of hard drives in the cluster from 18 to 3, simplifying the configuration and backup of the system, while also reducing the number of hard drives to be secured in the future use of classified data. Array has developed and deployed a modern, COTS based, open architecture, scalable generic image generator system for ADDNS. This IG system can be easily upgraded and expanded to address more demanding future applications, such as cockpit flight simulator displays, and higher resolution laser projectors. Future challenges include further reducing network latency, and increased graphics card resolution to support next generation high resolution laser projectors.

R.M. Obidowski (✉) and R. Jha
Array Systems Computing Inc., 1120 Finch Avenue, 7th floor, Toronto, ON, Canada M3J 3H7
e-mail: robidowski@gmail.com

Sommaire Le projet de démonstration de la technologie du simulateur déployable avancé pour la vision de jour et de nuit (ADDNS TDP) exigeait à titre essentiel au moins 16 générateurs d'images (GI) afin de rendre des scènes à une fréquence de trame de 60Hz pour un système de projection laser à très haute résolution, soit plus de 20 millions de pixels. Array Systems Computing Inc. (Array) a adapté son architecture de processeur évolutif générique (PEG) avancé pour répondre aux exigences de l'ADDNS. Cette solution comprend une configuration en grappes Linux d'ordinateurs personnels (PC) à 16 noeuds, chaque PC étant composé de matériel commercial intégré avec de multiples processeurs Opteron, cartes graphiques NVIDIA et composants de soutien. Pour exploiter le matériel et utiliser le système rendu réparti, Array a intégré le logiciel d'exécution Flight-IG de Multigen-Paradigm Inc. (MPI) et l'interface CIGI (Common Image Generator Interface) sous Linux Centos. Lors de ce processus de développement, un certain nombre de défis d'ingénierie et de conceptions nouvelles ont été relevés afin de produire une solution commerciale évolutive et déployable faisant appel à des PC.

Le système GI a été conçu pour pouvoir communiquer avec d'autres systèmes utilisant l'architecture ouverte et les normes associées. À cette fin, on a fait appel au protocole ouvert CIGI (Common Image Generator Interface), à des composants commerciaux et à la distribution Centos du système d'exploitation Linux. On prévoit aussi d'assurer l'interface avec d'autres réseaux en utilisant les protocoles de simulation interactive répartie (SIR) et d'architecture évoluée (HLA).

Le «pavage» obligatoire 4×4 de 16 entrées d'interface vidéo numérique (DVI) pour la construction d'une scène entière par le système d'affichage au moment de l'exécution nécessitait le rendu réparti utilisant des techniques de réseautage avancé dans Linux, y compris la multidiffusion, la synchronisation hôte-entrée et une configuration de noeuds maître-esclave. Ces techniques sont présentées ici en détail. Les exigences en matière de coût et d'exploitation du système réclamaient l'utilisation de composants génériques commerciaux facilement remplaçables, y compris des cartes graphiques modernes avec processeurs graphiques (GPU) évolués. Ces spécifications et les données de performances sont exposées dans le présent document. Array a développé et déployé pour l'ADDNS un système moderne de générateurs d'images génériques évolutifs, à architecture ouverte, utilisant des composants commerciaux. Le système peut facilement être étendu à des applications plus exigeantes; il est en outre souple et suffisamment reconfigurable pour être utilisé dans d'autres applications de génération d'images.

The Advanced Deployable Day/Night Simulation (ADDNS) Technology Demonstration Project (TDP) contained a critical requirement for 16 synchronized Image Generators (IGs) to render scenes at 60 Hz for a >20 million pixel Ultra-High Resolution (UHR) laser projection system. Array Systems Computing Inc. (Array) adapted a scalable GSP (Generic Signal Processor) architecture to meet these requirements in an open, scalable, generic IG system. The PC-IG system (Eagle-IG) contains 16 Personal Computers (PC) in a parallel Linux cluster, each PC with Commercial Off-The-Shelf (COTS) hardware, multiple AMD Opteron

Central Processing Units (CPUs), an NVIDIA 7800 graphics card, and supporting components. Multigen-Paradigm Inc. (MPI)'s Flight-IG software implements the Common Image Generator Interface (CIGI) in a Centos-based Linux platform. Additional enhancements being developed for the system include boot-off-network diskless render-IGs, fully integrated network backup server, and updated OS build (Centos 5). This enables a reduction in hard-drives from 17 to 2 – simplifying the configuration, backup, and securing of classified data. Future challenges include network latency minimization and increased graphics card resolution.

The Advanced Deployable Day/Night Simulation (ADDNS) Technology Demonstration Project (TDP) requires advanced Image Generator (IG) hardware and software for driving the Ultra-High Resolution (UHR) laser projection system. An objective of the IG design is to leverage advances in Commercial Off-The-Shelf (COTS) computer software and hardware technology, procuring a cost-effective and high-performance solution to facilitate the integration of ADDNS into the Advanced Distributed Mission Trainer (ADMT) and also have direct compatibility with the Advanced Deployable Combat Training System (ADCTS) system

The ADDNS project shares objectives with a research program of the United States Air Force Research Lab (AFRL), of producing an ultra-high resolution (UHR) projector capable of displaying an unprecedented resolution of 5,120 pixels by 4,096 lines at 60 Hz frame rate.

6.1 UHR Projector IG Interface Requirements

The main purpose of the IGs for ADDNS is to provide real-time, seamless video signals to the UHR projector. The primary interface uses single-link digital video interface (DVI) connections. IG interface specifications are the primary consideration for the IG design and implementation. The main IG requirements are as follows: in tiled Standard Raster approach, 16 progressively scanned 1,280×1,024 at 60 Hz video inputs are used, with 16 independent PC-IG channels to drive a UHR projector. The 16 IG video outputs are frame-synchronized, and each output fills a specified section of the projectors double-buffered frame buffer. These frame buffers are large enough to support the desired output resolution at 12 bits per color.

The final projected image is thus composed of real-time stitching of the individual input tiles as shown in Fig. 6.1. Together, this arrangement of tiles creates the >20 MP laser display.

The most cost-effective overall IG solution to meet these video interface specifications and the ADDNS environmental requirements guided the deployment of a synchronized cluster of COTS personal computers (PCs) running MPI's Flight-IG runtime in the Centos Linux Operating System (OS). This modern and cost-effective design fully utilizes the advanced technology and price/performance of modern consumer-grade PC graphics cards with advanced Graphics Processor Units (GPUs). Additional key technology for these COTS PC-IGs is special genlock hardware (HW).

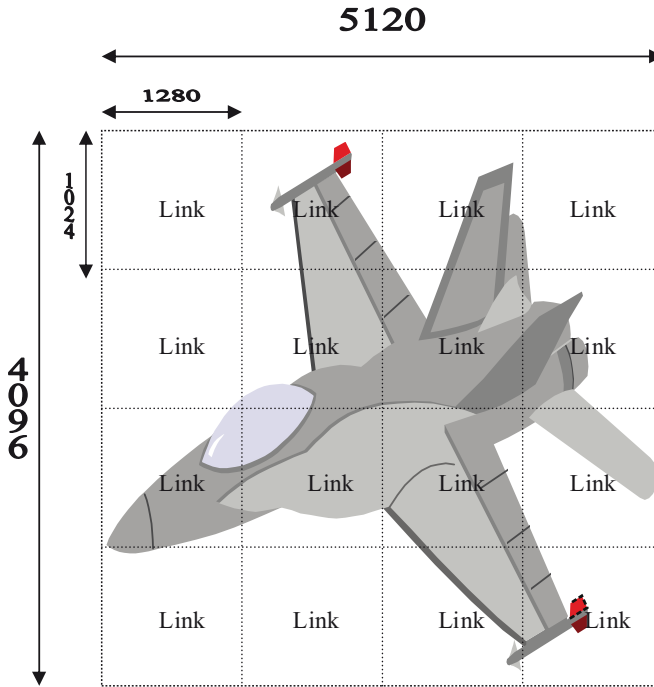


Fig. 6.1 Standard raster DVI links

6.2 Additional IG Design Considerations

In addition to projector requirements, IG design considerations are:

1. The system must be field-deployable for training; consequently, it must occupy as little physical space as possible.
2. The system must be deployable in locations with limited power availability; thus total power consumption is to be minimized.
3. The system must operate as quietly as possible, since it is to be located within 50 ft of the laser projector and its operator(s).
4. The system may have to deploy to arid climates, thus it must have a built-in, effective cooling system.
5. The system must be easily maintainable with replacement COTS components or stocked spares for 5 years.
6. Upgrades to the system are to be able to take advantage of technology improvements during the 5-year life span.
7. The system and architecture must be scalable in terms of graphics resolution and number of nodes.
8. The system must be able to use both classified and unclassified data; this requires that all hard-drives be removable for lock-up.
9. For security, the system must run unconnected to any external networks.

We were able to accommodate all of the above requirements in its system design by adapting the scalable generic signal processor (GSP) architecture to this PC-IG application.

6.3 Scalable GSP Software Architecture

We have developed software (SW) architecture for real-time data collection and processing, called the Scalable GSP. The Scalable GSP architecture is machine-independent and can be readily ported from one platform to another, as illustrated in Fig. 6.2.

Scalable GSP technology includes the following characteristics:

- *Open System.* Third parties can independently develop and integrate modules with new functions
- *Open Source.* Developed under open source Linux OS and open source standard libraries
- *Open Standard.* Non-proprietary standards used throughout the design process
- *Open Architecture.* Only COTS non-proprietary components and interconnects are used
- *Scalable.* Capability to add/remove processors without re-writing software
- *Flexible.* Ability for user to build and modify the processing chain at run-time

The Scalable GSP design was used as a basis for the open, scalable, generic, PC-IG solution for the ADDNS UHR projector.

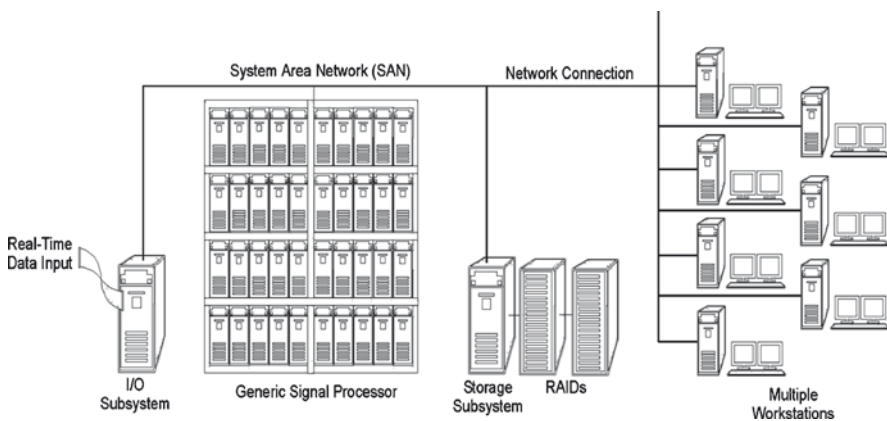


Fig. 6.2 Scalable GSP system overview

6.4 Overview of PC-IG Hardware and Software Selections

The PC-IG HW solution considered two critical design considerations (1) the UHR projector video signal interface specifications, and (2) the ability to interface with the ADCTS CF-18 simulators. First, selection of a suitable graphics card was essential. Research and testing showed an NVIDIA 7800GTX (GeForce) graphics card combined with a HW genlock system would meet video requirements with the best price/performance. Second, integrating the MPI Flight-IG run-time graphics system (built on Open GL) provided the same graphics interface as the ADCTS simulators.

Two other important factors to consider for building an open and scalable solution were to select the OS and Central Processing Units (CPUs). The Flight-IG SW runs under both Windows and Linux OS, as both are widely used in PC-IGs. Linux OS was chosen over Windows for the following reasons:

1. Linux provides closer compatibility than Windows to the ADCTS platform, which uses both IRIX (SGI workstations) and Linux (for sensors and Part-Task Trainers (PTTs)).
2. A smaller memory footprint using Linux can be implemented for some configurations, reducing the required memory on the PC-IGs.
3. FlightIG in Linux offers improved vertical synchronization compared to using Windows, according to the manufacturer.
4. Linux permits the addition and use of real-time kernel interrupts and processor scheduling.
5. Linux eliminates node-locked license keys (as required in Windows XP), thus enabling hard-drives to be interchanged among the PC-IGs.
6. Open-source Linux OS is more compatible with the open, scalable GSP architecture.

The specific OS for the IGs is Linux CentOS 3.6. Flight-IG mainly supports Red-Hat-based Linux distributions, and CentOS 3.6 is an open source Linux distribution that is designed to be Red-Hat compatible. In ADDNS, it is used under a GNU Public License (GPL).

PC hardware for IGs is mounted on standard 19-in. racks. The Electronics Industry Alliance (EIA) 2U form-factor (3.5 in.) allows IGs to be compactly assembled for deployment. The selection of the CPU was carefully evaluated. After review of power usage, memory interface, and overall price/performance (at the time), the AMD Opteron 250 CPU was chosen, with two CPUs per IG. FlightIG software is a multi-threaded application that allows user control of process threads. The dual CPU solution enables optimal performance on the IGs by allocating and distributing the draw, cull, and app threads as needed.

The video card in the IGs is the NVIDIA GeForce 7800 GTX, manufactured by XFX Corporation. It has 256 MB of memory, allowing large quantities of textures and geometry to be loaded directly into the video card. Another important advantage of this card is that it has a narrow profile, enabling it to easily fit into a single slot. This is critical for fitting the required graphics and genlock cards into the IGs'

Table 6.1 GeForce 7800 GTX – GPU specifications

Process technology	0.11 μm
Number of transistors	302 Million
Memory architecture	256-bit GDDR3
Core clock speed	430 MHz
Memory clock speed	600 MHz (1.2 GHz effective)
Vertex shading units (#)	8
Pixel shader pipelines (#)	24
Peak pixel fill rate	6.9 GB/s (16 ROPs \times 430 MHz)
Bilinear-filtered texel fill rate	10 GB/s (24 pipes \times 430 MHz)
Vertices/s	860 million

riser card assembly and compact case. The key specifications of the XFX NVIDIA GeForce 7800GTX graphics card are summarized in Table 6.1.

6.5 Physical IG Characteristics and Operating Considerations

It was desired to minimize the physical footprint of the IGs for deployment of the system. The main factor affecting the footprint size is the total physical size of the required PCI cards, such as the video card. The cards required in the IGs are the NVIDIA 7800 graphics card and two genlock system cards. Fitting the graphics card into the IG's 2U form factor required using a PCI Express riser card and mounting the graphics card sideways. The riser card was also able to accommodate the mounting of the genlock cards. In this arrangement, an important consideration is cooling the internal components. The built-in CPU, case, and graphics card fans were found to sufficiently cool the units. The IGs use a high velocity wind-tunnel design, drawing air from front to back, across the components. The final 2U PC-IGs have the following physical characteristics (per unit): weight: 37 lb (17 kg), size (LWH): 28 in. \times 19 in. \times 3.5 (2U) in.

6.6 Genlock System and Testing

The genlock architecture uses a master/slave relationship between the IG-nodes, with the master IG-node signaling the slave(s) IG-nodes. Genlock is achieved when a slave device's vertical sync is aligned to within 10 ms of the master node's vertical sync signal. A 100 MHz oscilloscope test verified signal-level genlock of a two-unit Master/Slave configuration within the time specification of less than 10 ms. In addition to the lab verification shown above, the UHR projector system contains built-in frequency, stability, and synchronization monitoring which were used to verify that the full set of 16 DVI input signals remained within the required tolerance during IG operation.

6.7 Power Investigation and Measurement

As a deployable system, reducing power requirements is an important consideration. To assess the operational power required with the laser projector, IG power measurements were taken using an in-line digital power meter. All measurements were taken at steady-state operation, running a typical flight script scenario. The measurements show a fairly modest total power requirement of approximately 1,024 W, or under 4,000 VA.

6.8 Acoustic Noise Level Investigation and Reduction Strategy

The small form-factor (2U) of the IG-nodes requires smaller, faster cooling fans than larger systems. The high revolutions per minute (RPM) that the internal fans must maintain to cool the ADDNS PC-IG created a noise concern. By measuring and extrapolating a single IG unit's noise, the full noise level for 16 nodes at 1 m away is predicted to be between 74–77 dB. For comparison, a 50 dB sound is approximately equal to conversational speech at 1 m. The predicted 77 dB sound is comparable to ambient noise found in a typical department store or busy office setting. To reduce the overall noise in the system, new case cooling fans with thermal monitors and auto speed control for internal temperature were adopted. In addition, special new power supplies with quieter fans were used. These measures delivered a sustained 30% reduction in operating noise to the predicted 77 dB.

6.9 Multiple-Channel Integration and Distributed Rendering Using CIGI Protocol

The PC-IG architecture uses Distributed Rendering (DR) to optimize the parallel performance of the cluster. DR is achieved by using the CIGI protocol to limit Field Of View (FOV) settings for each PC-IG, along with the simultaneous running of 16 copies of Flight-IG. DR uses network multicasting on a dedicated high-speed (gigabit) Ethernet switch. Multicasting as a network transmission technique is described in Deering and Cheriton (1990). Normally, the CIGI protocol interfaces one host (transmitter) and one IG (receiver) (Phelps, 2002). However, Flight-IG uses a CIGI relay to multicast packet data to all IGs, each with a unique render number. Although all IGs receive the full CIGI packets, each one renders data only in its pre-set FOV, established during the initialization. The recommended FOVs for the UHR projector were provided by its manufacturer; typically each IG renders FOV 20° horizontally by 16° vertically. Operating a host-based multi-channel IG cluster required the ability to start each channel remotely and to synchronize all channels to the host. Remote Linux shell scripts start each unit, along with a multicast relay

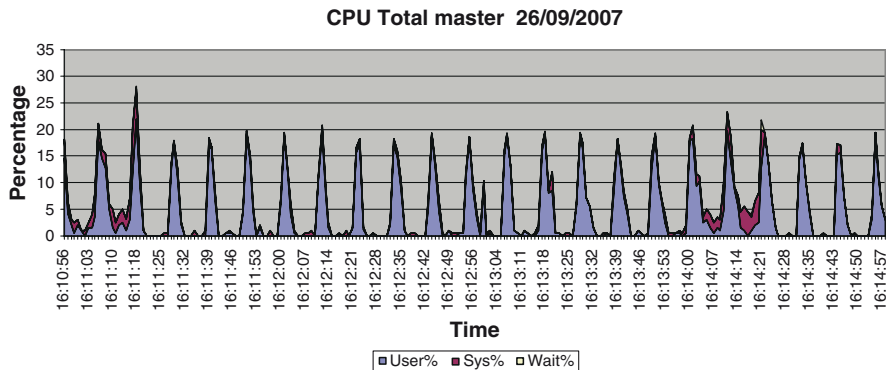


Fig. 6.3 Measured total CPU usage on the Master-IG

on the Master IG. The host communicates with the relay, which communicates with the other IGs over the network switch. A multi-system setup contains the render IGs subscribing to the multicast address, with the first IG (Master) acting as arbitrator, collecting all the responses. This is illustrated in Fig. 6.3.

Another critical part of the overall run-time system is the integrated geographical terrain database (DB). The terrain DB supplied for the initial evaluation is a small geographical terrain database that covers the Nellis Air Force Base in Nevada. The Nellis-DB is 476 MB in size, including entities and terrain data. In our initial build of the SW for the PC-IG system, identical local copies of each database were installed on each IG. This is accepted practice in wide area distributed simulation (WADS) exercises, such as those using high-level architecture (HLA) and/or distributed interactive simulation (DIS) protocols, as, for example, described in Simtech (1998). However, during the system deployment, it was noted that the Meta-Flight DB format and Flight-IG's built-in large-area database module (LADBM) allows network-mounted databases to be used without excessive network traffic and latencies. Initial testing of a remotely mounted database on the master IG node was successful, and this configuration is planned for adoption in the next release build of the system.

6.10 SW Host Emulator and Scripts for System Demonstration

The use of CIGI protocol to operate and synchronize the IGs required a dedicated host computer and control software. This system operates open source Host Emulator (HE) software provided by Boeing and described in Durham (2006). For demonstrations of the system, a number of scripts were programmed using an MPI-supplied CF18 graphics model and the Nellis geographical terrain DB. One real-time script demonstrated the improved visual acuity of the UHR laser projector compared to a conventional liquid crystal display (LCD) projector.

6.11 IG System Validation and HW Performance Analysis

In addition to our informal live visual validation of system performance and repeated live inspections of video frame rates displayed on the system, real-time performance data for CPU loading, free memory, and network speeds were also collected. This approach validates the selection of HW, SW, and OS, and evaluates the amount of computational reserve in the system.

The data was collected for approximately 6 min, while operating the IGs with the laser projector running a simulated CF18 flight pattern around Nellis Air Force Base. Performance data was logged simultaneously on the Master IG and the last render node (render-15). The NMON performance-monitoring tool was used to collect the monitoring data, along with the N-Perf analysis tool. Measurement points were collected every 2 s, with a total of 240 samples used in the analysis below. Figure 6.3 shows the CPU utilization on the IG master node during steady-state operation, while running a flight script as described above. Notice in Fig. 6.3 that the maximum CPU load on both processors on the IG master node is only 30% of total CPU capacity. This demonstrates that there are sufficient computing power and reserves on the IG. The free memory on the Master IG was also measured – a maximum of approximately 1,250 MB RAM free, around 60% of the 2 GB memory capacity.

Figure 6.4 shows network data transmission rates during the tests on the Master IG node, which is typical for all the nodes. As shown, a small fraction of the available band-width is used. This is because the databases are locally mounted on each IG rather than being network-mounted. The large amount of remaining network bandwidth supports the feasibility of network loading both the OS and databases for the system for the render nodes. These performance measurements demonstrate the capacity of the hardware in the PC-IGs, showing that the system has adequate reserves for adding moving entities and larger terrain databases.

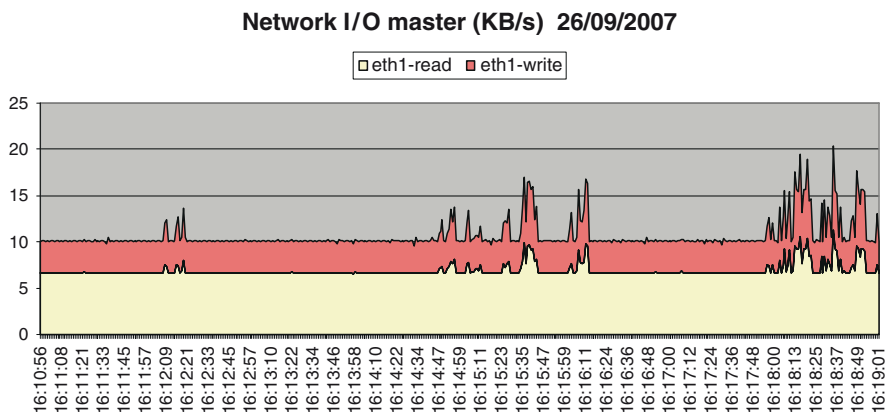


Fig. 6.4 Measured network transmission rates on Master IG node

6.12 Current Development and Status of the Second Build of the IG Software System

The initial build of the PC-IG system was shown to operate the UHR laser projector as required for ADDNS. The following upgraded build is currently planned:

1. Update Centos 3.6 to Centos 5.0, with V2.6.18 Kernel – providing numerous bug-fixes and enhancements to the OS, such as improved cluster support
2. Use diskless (stateless) render IGs by a boot-off network thick client approach – allowing fewer hard-drives (two instead of 17)
3. Network-mount geographic terrain database on the master IG
4. Support full backup and restore on a dedicated backup server
5. Use network Wake-On-LAN (WOL) for remote power sequencing.

Running render-IGs stateless (without hard-drives) has the following advantages:

1. Simpler maintenance and upgrades, as only one hard-drive and OS image needs to be updated
2. Less data to back-up, allowing faster backups and restoring of the system
3. Fewer (2 instead of 17) hard-drives to manage and secure for classified data use
4. Reduced heat, noise, and power consumption due to reduction in number of hard-drives

6.13 Conclusions

We have designed, developed, and validated a COTS, scalable generic IG system with parallel distributed rendering. This CIGI-based cluster system can easily be upgraded and expanded (by adding nodes or reconfiguring graphics channels) for more demanding future applications, such as cockpit flight simulator displays and higher resolution laser projectors, thus opening new avenues for realistic, deployable field rehearsal and training simulators.



Chapter 7

Detection Threshold of Visual Displacement in a Networked Flight Simulator

Christine M. Covas, James P. Gaska, George A. Geri,
Megan E. Shamp, and Byron J. Pierce

Abstract Networked flight simulators facilitate team training by creating exercise environments and simulations representative of real-world operations. However, networked simulators are often plagued by connectivity issues, such as constant and variable network delay. Network delay can cause positional discrepancies or visual jitter in the flight path of a moving model which could effect pilots overall perception of fidelity of motion. To provide a realistic team-training environment, simulators are networked enabling trainees to perform exercises emulating real-world conditions. When simulators are networked, especially over large geographical areas, interactions between them become more complicated. Three important issues related to the realism of time must be considered when networking simulators: the timing consistency of the virtual world, delays, and the overall representation of time in a simulator.

The level of realism of an entity in a networked simulator depends upon factors such as network delay, thresholds used to determine when a simulated position error should be corrected, and algorithms used to correct position errors. These factors can interact and produce positional discrepancies between the position provided by a local model of an entity, which is used to render the entity, and the actual position of the entity. The correction of positional discrepancies via the use of smoothing within the Image Generator (IG) results in a positional displacement of the rendered entity [Lin et al. (Proceedings of the 1995 American Institute of Aeronautics and Astronautics Flight Simulation Technologies Conference 83–87, 1995)]. The purpose of this experiment was to assess the detection threshold of positional error or displacement threshold as a function of distance from eyepoint, and positional smoothing performed by the IG. This data will contribute to a larger

K.M. Covas (✉), M.E. Shamp, and B.J. Pierce
Air Force Research Laboratory, Immersive Environments Branch, 6030 South Kent Street,
Mesa, AZ 85212-6061, USA
e-mail: Christine.Covas@mesa.afmc.af.mil ; Christine.Covas-Smith@us.af.mil

J.P. Gaska and G.A. Geri
Link Simulation and Training, 6030 South Kent Street, Mesa, AZ 85212-6061, USA

research paradigm to assess the impact of network parameters on task performance in a flight simulation environment.

The results of this experiment show that increasing simulated distance increases the amount of displacement, in simulation space, required to detect target aircraft motion. However, when the displacement was expressed in units of visual angle, the effect of simulated distance was not significant, indicating that the thresholds were determined by angular displacement. Additionally, under the conditions used in this study, smoothing approximately doubled angular displacement thresholds.

Under all conditions, displacement thresholds were small – approximately 2 arcmin for the unsmoothed condition and 4 arcmin for the smoothed condition. This translates into displacement thresholds of $\frac{1}{2}$ pixel and 1 pixel for the unsmoothed and smoothed conditions, respectively. In fact, studies performed under optimal conditions report displacement thresholds of 2 arcsec.

The finding that smoothing, which decreases target velocity, increased the displacement threshold is not surprising. Smoothing is used in a networked environment to reduce the visual saliency of the displacement and provide a more realistic looking flight path to an observer. However, the smoothed flight path will provide a less accurate representation of the true flight path. The high sensitivity of the visual system in detecting displacement shown in this study and others suggests that the positional error of the smoothed flight path can be encoded by the visual system and could, therefore result in performance degradation. Furthermore, because the smoothed flight path looks realistic it may alter a pilot's performance whereas the unrealistic jump with smoothing disabled can be easily discounted by the pilot. Therefore the benefits and advantages of smoothing are still subject to debate and to future research.

This experiment is the first in a series of experiments designed to determine the effect of network conditions on flight tasks in a networked simulator training environment. At the Air Force Research Laboratory in Mesa, AZ we are conducting a detailed evaluation of the actual physical positional discrepancies that occur as a function of the overall load of the network. This evaluation will allow us to compare our results from the present experiment to actual positional discrepancies that would manifest as a result of network load.

Sommaire Pour fournir un environnement réaliste de formation au travail en équipe, on utilise des simulateurs maillés en réseau permettant aux stagiaires d'effectuer des exercices reproduisant des conditions qui prévalent dans la réalité. Lorsque des simulateurs sont maillés en réseau, en particulier sur de vastes régions géographiques, les interactions entre eux deviennent plus complexes. Lorsque l'on procède au maillage en réseau de simulateurs, on doit tenir compte de trois problèmes importants reliés au réalisme temporel: la cohérence temporelle du monde virtuel, les délais et la représentation globale du temps dans un simulateur.

Le réalisme d'une entité faisant partie d'un simulateur maillé en réseau dépend de facteurs comme le délai du réseau, les seuils utilisés pour déterminer quand une

erreur de position simulée doit être corrigée ainsi que les algorithmes utilisés pour la correction de ces erreurs de position. Ces facteurs peuvent interagir et produire des différences de position entre la position fournie par un modèle local d'une entité qui est utilisé pour représenter cette entité et la position réelle de cette entité. La correction de différences de position par l'utilisation du lissage à l'intérieur du générateur d'images (GI) se traduit par une variation de position de l'entité représentée. Cette expérience visait à évaluer les seuils de variation en fonction de la distance du cercle oculaire ainsi que le lissage de position effectué par le GI. Ces données contribueront à un paradigme de recherche plus important pour l'évaluation de l'impact qu'ont les paramètres de réseau sur l'exécution des tâches dans un environnement de simulation de vol.

Les résultats de cette expérience indiquent qu'une augmentation de la distance simulée provoque une augmentation de la variation, en espace de simulation, nécessaire à la détection du mouvement de l'aéronef cible. Cependant, lorsque la variation était exprimée en unités d'angle visuel, l'effet de la distance simulée n'était pas important, ce qui indique que les seuils étaient déterminés par la variation angulaire. De plus, dans les conditions utilisées dans le cadre de cette étude, le lissage double presque les seuils de variation angulaire.

Dans toutes les conditions, les seuils de variation étaient petits – environ 2 minutes d'arc pour la condition sans lissage et 4 minutes d'arc pour la condition avec lissage, ce qui se traduit par des seuils de variation de $\frac{1}{2}$ et 1 pixel pour les conditions sans lissage et avec lissage, respectivement. En fait, dans le cadre d'études effectuées dans des conditions optimales, on signale des seuils de variation de 2 secondes d'arc.

Il n'est pas étonnant de constater que le lissage, qui réduit la vitesse de la cible, a provoqué une augmentation du seuil de variation. Dans un environnement maillé en réseau, on utilise le lissage pour réduire la prégnance visuelle de la variation et fournir à un observateur une trajectoire de vol qui a l'air plus réaliste. Cependant, la trajectoire de vol avec lissage fournira une représentation moins précise de la trajectoire de vol réelle. La sensibilité élevée du système visuel en matière de détection de la variation que démontre cette étude et d'autres études laisse croire que l'erreur de position de la trajectoire de vol avec lissage peut être codée par le système visuel et pourrait donc provoquer une détérioration des performances. De plus, comme la trajectoire de vol avec lissage a l'air réaliste, elle peut altérer les performances d'un pilote, tandis que le passage brusque non réaliste à un lissage désactivé peut facilement être ignoré par le pilote. Les bénéfices et les avantages du lissage laissent donc toujours place à la discussion.

Cette expérience est la première d'une série d'expériences conçues pour déterminer l'effet des conditions du réseau sur les tâches de pilotage dans un environnement d'entraînement sur simulateur maillé en réseau. Au Air Force Research Laboratory de Mesa (Arizona), nous procédons à l'évaluation approfondie des différences de position physiques réelles qui surviennent en fonction de la charge globale du réseau. Cette évaluation nous permettra de comparer nos résultats obtenus lors de la présente expérience aux différences de position réelles provoquées par la charge du réseau.

Entity realism in a networked simulator depends upon factors such as thresholds used to determine when to correct errors, algorithms used to correct errors, and network timing errors. Positional errors result from complex interactions of the above factors. We have measured spatial displacement threshold for a typical flight task (formation flight) as a function of angular size of an aircraft and presence or absence of smoothing. Angular displacement thresholds were independent of target-size and, although smoothing approximately doubled displacement thresholds, the result was not statistically significant. We will discuss the implications of these results as well as provide suggestions as to when algorithms such as smoothing should be enabled.

To provide a realistic team-training environment, simulators are networked enabling exercises emulating real-world conditions. To broadcast information over a network, information from simulators is formatted into messages called protocol data units (PDUs), which are used to describe specific aspects of the simulation environment. When change occurs in the simulated environment, i.e. a weapon is fired or an entity leaves a predictable motion path, the entity causing the reaction or experiencing it first transmits a message informing others of the event and describing what has occurred.

The level of realism of an entity in a networked simulator depends upon factors such as thresholds used to determine when an error should be corrected, algorithms used to correct the errors, and network timing errors. Positional errors, that is, discrepancies between the position provided by a local model of a remote entity and the actual position of the entity in the virtual environment result from a complex interaction of the factors outlined above. Recently, Slater and Covas (2007) characterized the position errors that result from increasing network load. The purpose of this research was to identify the conditions that produce positional errors and to suggest engineering solutions to reduce them. However, this research did not address the perceptual consequences of positional error. In this companion paper, we examine the perceptual saliency of positional errors by measuring spatial displacement thresholds.

The smallest spatial displacement that a human can reliably detect is dependent on stimulus size (Mestre et al. 2001), velocity (Nakayama and Tyler 1981) and retinal eccentricity (McKee and Nakayama 1984). In addition, displacement thresholds measured in the presence of a reference, often called relative motion thresholds, are smaller than those measured without a reference, referred to as absolute motion thresholds (Tyler and Torres 1972; Levi et al. 1984). In this study, we measure absolute motion thresholds of targets presented to the fovea. Previous research shows that these conditions produce displacement thresholds of approximately 0.5 arcmin (McKee et al. 1990).

The purpose of this experiment is to determine the effect of target velocity and target size on displacement thresholds using a representative flight task (i.e. formation flight). Target velocity was varied by moving the target to a new position over 1 s or 1/60th of a second timeframe. These two velocity conditions model the 1-s smoothing or no smoothing procedures often used in networked simulators. The angular size of the aircraft was varied by altering the simulated distance between the aircraft and observer. The simulated distances used in this study were 100, 200, and 400 m, which is typical of the distances encountered in formation flight.

7.1 Methods

Observers. Five observers served in the experiment. All observers had normal or corrected-to-normal visual acuity (as measured with an Optec 2000P, Stereo Optical, Inc.). All observers gave documented informed consent. Some observers had served in previous unrelated research studies.

Apparatus. The IG was a Windows PC equipped with an nVIDIA GeForce4 graphics card set to 32-bit color, 4x antialiasing, a pixel format of 1,280×1,024, and a 60-Hz update rate. Visual scenes were rendered with MetaVR's real time visualization application, Virtual Reality Scene Generator (VRSG)TM. Custom software, written in C++, was used to control the characteristics, timing, and sequencing of the target motion, to communicate with VRSG, and to record the observer's responses. A Barco 808 CRT projector set to a refresh rate of 60-Hz was used to rear project the test stimuli (Proscreen Inc., Medford OR, gain = 1.2). The mean luminance of the display was 4.2 fL. The projected-image size was 51.5×41.1 in., which subtended a visual angle of 71.15×59.44° at the viewing distance of 36 in. The projector was set to 1280 H by 1,024 V which corresponded to a pixel size of 3.34 arcmin horizontal and 3.48 arcmin vertical.

The measured number of resolved lines was 986×971 as determined by a standardized measurement procedure (Winterbottom et al. 2004). A Microsoft Sidewinder Joystick interfaced with the host was used by observers to respond to the stimuli.

Stimuli. We used a blue sky and a fully textured, realistic three-dimensional model of an F-16. The model had a specified size of 15.85×10.00×4.85 m. The aircraft was shown at simulated distances of 100, 200 and 400 m, corresponding to visual angles of 8.99°, 4.52°, and 2.26° respectively. Figure 7.1 shows the aircraft used for this experiment at a position of 90° yaw angle and the 100 m simulated distance.

A trial consisted of a 2-s presentation of the F-16 on the blue-sky background that was present continuously throughout each trial. The aircraft was stationary for 0.25 s at the beginning and end of each trial, leaving 1 s for the aircraft to translate forward and return to its starting point, the extent to which was determined by the simulated distance and the target displacement amplitude. Simulated distance was varied within each block of trials. Thirteen target-displacement amplitudes were tested for each simulated distance. Displacement amplitude varied from 0.001 to



Fig. 7.1 F-16 aircraft model used in this experiment at 90° yaw orientation at the 100 m distance

2.048 m, 0.002 to 4.096 m, and 0.004 to 8.192 m, for the 100 m, 200 m, and 400 m distances respectively. All amplitudes were a factor of two from the lowest amplitude for each distance, with each distance including a no-displacement condition.

We used two different levels of smoothing, no-smoothing and one-second smoothing. In the no-smoothing condition, the aircraft took one frame (16.7 ms) to move up and one frame to move back. In the 1-s smoothing condition, which is typical of that used in a simulation environment, the aircraft took 1 s to move up and 1 s to move back. Therefore, the speed of the aircraft varied as a function of the displacement amplitude, the simulated distance, and the level of smoothing.

7.2 Procedure

Each observer received written as well as verbal instructions for the task they were to perform. Prior to each session, the observer was positioned 36 in from the screen. This viewing distance approximates the viewing distance in many Air Force flight simulators. Large target displacements were used to demonstrate clearly visible spatial displacements and smoothed and non-smoothed motion. Prior to the collection of data, one or more practice sessions were used to familiarize the observers with the task. Following each trial, text was presented on the screen asking the observers if they detected jitter. The observer indicated either “Yes” or “No” by moving a cursor to their selection and then pressing the trigger on the joystick. The next trial was initiated immediately after an observer’s response. Observers were encouraged to take breaks between blocks of trials, as needed, to minimize fatigue.

Each block of trials consisted of the 13 displacement amplitudes and three distance conditions with both smoothing conditions tested randomly within each block of trials. Each block of trials required 3–5 min to complete. Each experimental session consisted of 15 blocks of trials and lasted about 60 min. We had a total of: 13 (displacement amplitudes) × 3 (distances) × 2 (smoothing) × 5 (observers) = 5,850 data points.

7.3 Results

Figure 7.2 shows the proportion of reported jitter, as a function of displacement amplitude for one observer and one distance-smoothing condition. The data were fit to a modified Weibull psychometric function of the form:

$$P_{\text{Reported}} = 1 - e^{-\left(\frac{D}{T}\right)^Q},$$

where D is the displacement amplitude and Q , T are fitted constants. The Q parameter determines the steepness of the psychometric function; the T parameter provides an estimate of displacement threshold corresponding to a proportion of reported jitter of 0.63.

Figure 7.3 shows the displacement thresholds, averaged across observers, for the two smoothing conditions and the three distance conditions. The figure shows that

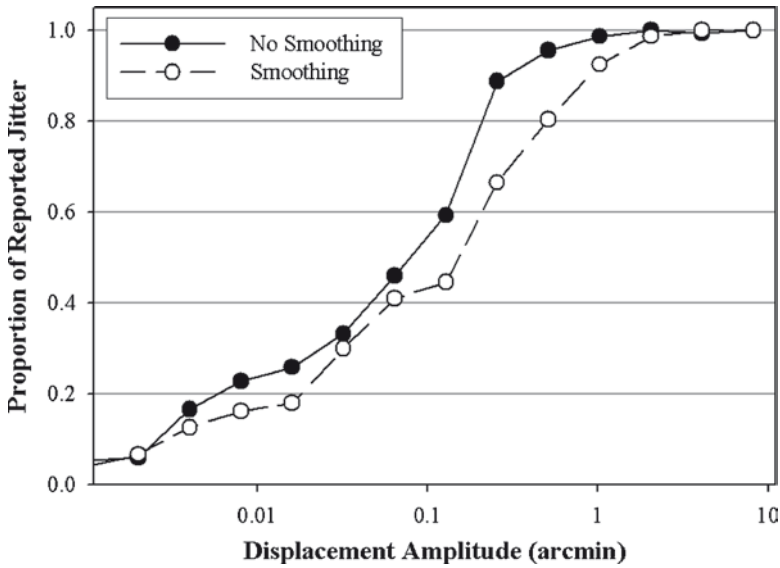


Fig. 7.2 Mean response data as a function of target displacement with Weibull fit (represented by long dashed line) for one observer for the 100 m, with smoothing condition. The darker line represents the performance level and the short dashed line represents threshold for this combination of conditions

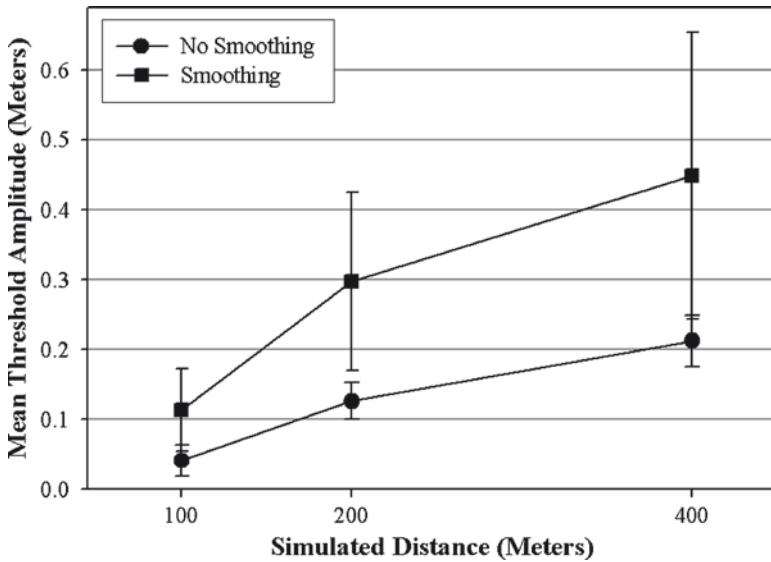


Fig. 7.3 Mean target-displacement thresholds for three simulated target distances and two smoothing conditions. Circles represent the no-smoothing condition, and squares represent the 1-s smoothing condition. Error bars represent ± 1 SEM intervals about each data point

displacement thresholds generally increased as a function of increased distance and smoothing. An ANOVA revealed that the effect of simulated distance was statistically significant $F(2, 8)=6.97, p<0.05$. A post-hoc pairwise comparison showed that the following simulated distances were significantly different ($p<0.05$): 100 m vs. 200 m and 100 m vs. 400 m. We did not find a significant overall effect of smoothing, or an interaction with distance.

In Fig. 7.4, the experimental data is transformed from simulation space units (meters) into observer units (visual angle). Repeated-measures ANOVA did not reveal a significant effect of either angular target size or smoothing.

7.4 Discussion

The results show that increasing simulated distance increases the amount of displacement, in simulation space, required to detect target aircraft motion. Because a constant displacement in simulation space produces an angular displacement that is inversely proportional to simulated distance, this result was expected. The angular displacement, not displacement in simulation space is the appropriate metric to evaluate the encoding of displacement by an observer.

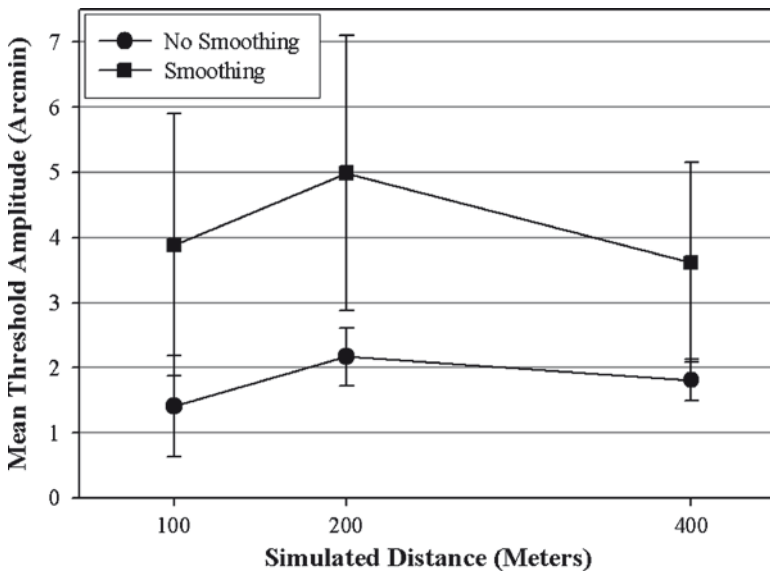


Fig. 7.4 Mean target-displacement thresholds for another three simulated target distances and two smoothing conditions. The circles represent the no-smoothing condition, and the squares represent the 1-s smoothing condition. Error bars represent ± 1 SEM intervals about each data point

We found that the angular displacement thresholds were independent of the angular target size. Previous research (Mestre et al. 2001) shows that motion thresholds decrease as the angular width of the stimulus is increased up until approximately 0.75° and saturates thereafter. Because the minimum width of our target was about 2.25° , our results are consistent with this finding and suggest that for the distances employed in typical formation flight task, angular target size will not significantly influence displacement thresholds.

The average of the unsmoothed displacement thresholds measured in this study was approximately 2 arcmin which corresponded to approximately a $\frac{1}{2}$ pixel displacement on the display. The sub-pixel displacement value suggests that thresholds measured in this study were display-limited, that is, restricted by the ability of the display to accurately present the displaced target. This analysis also illustrates why the displacement thresholds measured in this study are much larger than the 0.5 arcmin found in other studies which measured eye-limited thresholds (McKee et al. 1990).

Smoothing approximately doubled angular displacement thresholds although the result was not statistically significant. The weak effect of smoothing may seem counter intuitive given that smoothing can greatly reduce the visual saliency of displacements and is routinely used in a networked environment to provide a more realistic looking flight path to an observer. However, as shown above, the unsmoothed thresholds were limited by the spatial resolution of the display and it is likely that the ratio of smoothed to unsmoothed thresholds was reduced in this study because of a ‘floor effect’. Displacements that are much higher than the display spatial resolution, typical of those caused by network errors, would probably reveal a greater effect of smoothing on the perceived motion magnitude.

7.5 Impact

In a Distributed Interactive Simulation (DIS) environment, networked entities communicate by broadcasting formatted messages called protocol data units (PDUs). The maximum time between PDU updates is called the system heartbeat and is typically set to 6 s. In order to reduce intra-heartbeat network traffic, a “dead-reckoning” procedure is used in which position information is broadcast only when an entity leaves a predictable motion path. For example, an aircraft flying at a constant velocity and direction will not broadcast an intra-heartbeat positional update. To implement this procedure, each simulator maintains a detailed record of its position as well as simple dead reckoning model which is used to extrapolate its future position. When its position changes, the simulator compares its new position to the value of the dead-reckoning prediction and, a PDU is broadcast only if the difference exceeds a threshold value. Other simulators in the network use this information to update their dead-reckoning model of the entity that sent the PDU and alter its rendered flight path.

Currently, most dead-reckoning algorithms use a fixed threshold that is independent of the spatial arrangement of the networked entities. However, because an observer encodes angular displacement and not displacement in simulation space, a dead reckoning threshold, based on observed angular displacement may provide a more efficient use of network resources. Using this procedure, simulators would use small dead reckoning thresholds when they are near other networked entities and large thresholds when they are distant. This technique, called adaptive dead reckoning (Lee et al. 2001), has been shown to substantially reduce the number of update packets required to maintain a criterion dead reckoning accuracy.

When an update packet is received, the flight path generated from this new information is often smoothed to eliminate unrealistic looking jumps in position. Although smoothing provides a more realistic looking flight path, the smoothed flight path will often provide an inaccurate representation of the true flight path (Lin et al. 1995; Slater and Covas 2007). If the positional error produced by smoothing is less than an observer's displacement threshold, then the errors cannot be encoded by the visual system and would, therefore, have no influence on performance. However, the current results show that displacement thresholds measure using currently fielded simulators are display limited. Therefore, positional errors that are rendered to any currently fielded simulator display can be encoded by the visual system. Furthermore, because the smoothed flight path looks realistic, the pilot may respond to the error whereas the unrealistic jump that occurs when smoothing is disabled is less likely to influence performance.

In summary, when networked entities are in close proximity, as in close formation flight, smoothing should be turned off and dead reckoning thresholds should be small. For the foreseeable future, the size of the dead reckoning threshold required to completely eliminate perceived positional errors will be limited by the spatial resolution of the display. Small dead reckoning thresholds will, of course, increase network traffic. In order to reduce overall network traffic, we recommend that dead reckoning thresholds should be increased when the distance to other networked entities is large. In addition, smoothing can be utilized to produce a more realistic simulation of the flight path of distant and objects.

Acknowledgments This work was supported in part by US Air Force Contract FA8650-05-D6502 to Link Simulation and Training (a division of L-3 Communications Corp.). In addition, the following agencies provided financial and/or technical support for this research: Air Combat Command (ACC/A8AT), Secretary of the Air Force International Affairs (SAF/IA), and Defence Research and Development Canada – Toronto. We would like to thank Bill Morgan (The Boeing Co.) for programming support; Ryan Amann (Link Simulation) for image capture assistance; and Craig Eidman (AFRL/HEAE) for subject matter expertise. The views expressed here are those of the authors and do not reflect the official policy or position of the Department of Defense or the US Government.



Chapter 8

Evaluation of the Spatial and Temporal Resolution of Digital Projectors for use in Full-Field Flight Simulation

James P. Gaska, George A. Geri, Marc D. Winterbottom, and Byron J. Pierce

Abstract Digital projectors based on liquid crystal or micro-mirror technologies provide higher spatial resolution than the CRTs typically used in flight-simulator applications. However, their temporal response is limited by long hold-times. Hold-time is the interval that pixels are activated during each frame, and it is related to both the design of the driver circuitry and the display itself. Long hold-times result in the blurring of moving imagery, which can affect the detection and identification of simulated targets as well as the perceived realism of simulated scenes. Described here are the results of an evaluation of several digital projectors that implement various techniques for reducing pixel hold-time. The evaluation included estimates of both spatial and temporal resolution, as well as perceptual tests used to assess the validity of those estimates. The techniques described here can be used to assess the potential utility of candidate projectors for a given simulator application. The following projectors were tested: ferro-electric liquid crystal on silicon (FLCoS), liquid crystal on silicon (LCoS), liquid crystal (LCD), digital light projector (DLP, a digital micro-mirror device, or DMD), and cathode ray tube (CRT). The hold-time of the FLCoS and LCD projectors was varied using mechanical shutters, and the hold-time of the DLP projector was varied using an electronic shutter. The CRT and LCD projectors were used mainly for comparison and so their hold-times were not varied.

The number of resolved lines for the projectors tested ranged from about 1,200 lines for the CRT projector to about 4,000 lines for the DLP and LCoS projectors. Projector hold-times ranged from less than 1 ms for the CRT to about 16.7 ms (i.e., the full video frame) for the LCD. For the CRT projector, perceived blur, as measured using the moving-line test, increased from about 5 to 12 arcmin as target

J.P. Gaska (✉) and G.A. Geri
Link Simulation and Training, 6030 South Kent Street, Mesa, AZ 85212-6061, USA
e-mail: james.gaska@mesa.afmc.af.mil

M.D. Winterbottom and B.J. Pierce
Air Force Research Laboratory, Mesa, AZ USA

speed increased from about 5 to 68 deg/s. For the LCD, LCoS, and DLP projectors tested with hold-times of over 10 ms, perceived blur increased from 5 to 10 arcmin for target speeds below 10 deg/s to 40–60 arcmin for target speeds greater than about 50 deg/s. As hold-time was decreased for the LCoS and DLP projectors, perceived blur also decreased, although the decrease was generally proportional to target speed. For the LCoS and DLP projectors, perceived blur was not significantly different from that of a CRT when hold-time was reduced to about 4–6 ms. The threshold roll-detection magnitudes were about 10, 14, and 16 deg for aircraft-target speeds of 8, 17.5, and 35 deg/s, respectively. These data indicate a clear correlation between the roll-detection data and both the temporal-response measurements and the moving-line test data.

The U.S. Air Force and Navy are considering, or are in the process of, replacing CRT displays with digital displays for use in flight simulators. However, low temporal resolution and concomitant moving-image blur is a significant problem for the use of these digital displays. The display evaluation techniques developed at AFRL, Mesa provide a relatively simple way to objectively measure both display spatial and temporal resolution, and additionally, to predict their perceptual consequences. The results of these evaluations show that, for all projector types tested, reducing hold-time to about 4–6 ms effectively reduced moving-image blur to that of a CRT projector, which is currently the accepted standard for projector temporal response. Finally, the results of the temporal-resolution measurements were found to correlate with performance on an air-to-air task that is representative of those performed in full-field flight simulators. In order to assess projectors in applied settings, analogous tasks will need to be devised and implemented for the flight simulators of interest.

Sommaire Le flou des images mobiles est un problème important, surtout pour les écrans numériques (c.-à-d. à cristaux liquides ou à micromiroirs) des applications d'entraînement au pilotage et de simulation de vol. La réponse temporelle limitée des écrans numériques, qui cause le flou dans les images mobiles, est une conséquence de deux caractéristiques principales : 1) lenteur de l'activation et de la désactivation, et 2) longue durée de maintien (c.-à-d. le temps pendant lequel les pixels persistent sur chaque image), et elle est liée tant à la conception du circuit pilote qu'à celle de l'écran lui-même. Nous présentons ici les résultats d'une évaluation de plusieurs écrans numériques mettant en œuvre diverses techniques permettant de réduire la durée de maintien des pixels. Les techniques utilisées ici font partie d'une procédure d'évaluation d'écrans conçue pour quantifier les caractéristiques spatiales et temporelles des écrans utilisés dans les simulateurs de vol. Plusieurs types d'écrans d'affichage ont été testés : des écrans à cristaux liquides ferroélectriques sur silicium (FLCoS), des écrans à cristaux liquides sur silicium (LCoS), des écrans à cristaux liquides (LCD), un projecteur numérique (DLP) (dispositif numérique à micromiroirs) et un projecteur à tube cathodique. Dans certains cas (projecteur à tube cathodique et un projecteur à cristaux liquides), les affichages ont surtout servi pour comparaison et ils n'ont pas été modifiés. Par contre,

plusieurs écrans d'affichage ont été modifiés au moyen d'obturateurs mécaniques ou électroniques afin qu'on puisse régler la durée de maintien. Nous avons testé plusieurs de ces projecteurs : FLCoS (obturateur électronique), LCD (obturateur mécanique), LCoS (obturateur mécanique) et DLP (obturateur électronique).

La définition spatiale, fondée sur une technique basée sur la norme VESA, allait d'environ 1200 lignes pour un affichage à tube cathodique à près de 4000 lignes pour les affichages DLP et LCoS. La réponse temporelle, ou durée de maintien, allait de près de 16,7 ms (image vidéo complète) à environ 1 ms. Ainsi, les diverses méthodes utilisées pour réduire la durée de maintien se sont révélées efficaces. Par contre, la répartition de la luminosité, en réponse à une mire de 30 Hz, a varié considérablement en fonction du type d'écran d'affichage. Les écrans FLCoS et DLP fonctionnaient selon la modulation d'impulsions en durée et, par conséquent, présentaient des courbes de réponse en ondes carrées. Les écrans LCD et LCoS présentaient des temps d'activation et de désactivation plus lents, tandis que les écrans à tube cathodique, comparativement aux projecteurs LCD et LCoS, présentaient essentiellement une réponse en impulsions. À mesure que la vitesse d'essai a augmenté à près de 70 degrés par seconde, le flou perçu variait considérablement en fonction du type d'affichage. Le flou perçu était pire avec les écrans LCD non modifiés et LCD, LCoS et FLCoS lorsque la durée de maintien n'était pas réduite. Le flou perçu dans ces cas atteignait jusqu'à 60 minutes d'arc. Dans le cas des affichages à tube cathodique, et pour les affichages à durée de maintien réduite, le flou perçu était considérablement réduit. Dans le cas des écrans LCoS à obturateur mécanique, le flou perçu n'était pas vraiment différent de celui d'un écran à tube cathodique lorsque la durée de maintien était réduite à 4 ms ou moins.

L'Air Force et la Navy des États-Unis envisagent, ou sont en train, de remplacer les écrans à tube cathodique par des écrans numériques dans les simulateurs de vol. Toutefois, une faible définition temporelle et le flou d'images mobiles connexe constituent un problème important dans l'utilisation de ces écrans numériques. Les techniques d'évaluation des affichages mises au point à l'AFRL, à Mesa, offrent un moyen relativement simple de mesurer objectivement les définitions spatiale et temporelle et, de plus, de prédire leurs conséquences sur la perception. En outre, les résultats de ces évaluations montrent que la réduction de la durée de maintien, quel que soit le type d'affichage, est un moyen efficace de réduire le flou des images mobiles. Bien que l'effet du flou sur des tâches pertinentes en mission air-air et air-sol reste à déterminer, la réduction de la durée de maintien de 4 à 8 millisecondes était suffisante pour les images testées dans le cadre de la présente étude.

Digital displays can generate imagery with high spatial resolution and luminance, but their slow temporal response has limited their use in some applications, such as flight simulation, where rapidly moving imagery must be displayed. In this paper, we describe the evaluation of several digital displays whose temporal response is controlled either electronically or by the use of mechanical shutters. The evaluation relates the perceived blur of a moving test target to projector

temporal response. We show that perceived blur is equal to the product of the time that the image is present within a video frame and the average speed of a visually tracked target. Thus, the temporal response of a display can be selected to reduce moving-image blur to the level required for a particular simulator application.

Digital displays are rapidly replacing CRTs in many commercial and technical applications (Menozzi et al. 1999; Poor 2003). Among the digital displays that have become particularly popular are liquid-crystal displays (LCDs), liquid crystal on silicon (LCoS) displays, and digital light processor (DLP) displays. The major limitation of CRT technology is spatial resolution, which is related to the number of discriminable pixels that can be displayed. Desktop CRTs have shadow masks that increase spatial resolution, but shadow masks are not practical in full-color, high-intensity, and wide-field CRT projectors. In addition, increasing a CRT's luminance generally decreases its spatial resolution. Thus, at optimal settings, a wide-field CRT display may be limited to a relatively low luminance level. Wide-field digital displays, on the other hand, typically have a spatial resolution that significantly exceeds that of projection CRTs, and in addition they have much higher luminance levels. The major disadvantage of digital displays, however, is their slow temporal response, which is a consequence of primarily two characteristics. The first is long onset and offset times, which are inherent characteristics of the display technology (Yamamoto et al. 2000). For instance, early LCDs had onset and offset times that exceeded the video frame time, and as a result, temporal changes in the displayed image could be lost, or artifacts such as long tails on moving objects could be produced. The second characteristic of digital displays that may reduce temporal response is long hold-time (i.e., the time that imagery is presented during each frame) (Shima et al. 1995). More recent LCDs and LCoS displays have onset and offset times that are much less than the frame time but the image is displayed over the entire frame, usually for the purpose of increasing image intensity (Yamamoto et al. 2000).

One technical area where digital displays are of particular interest is flight-simulation. In this application, it is often necessary to simulate both wide-field imagery and rapidly moving objects. Displaying wide-field imagery requires a high pixel count, and, in addition, an accurate depiction of rapidly moving objects requires both fast temporal response and high pixel density (i.e., high spatiotemporal resolution). These properties have been combined, in principle, in the latest 20-MP laser projectors (see Bloom and Tanner 2007) but the capabilities of these devices have not been fully evaluated to date. However, various 8–10 MP LCoS projectors do have both high pixel counts and high spatial resolution. Those projectors could support wide-field flight simulation if their temporal responses were sufficiently fast.

We have assessed display temporal response through measurement of perceived tracking blur. Tracking blur can be understood by considering that a digital display represents motion as a temporal sequence of still images, whereas an eye that is tracking the motion moves continuously. Therefore, within a frame, a stationary display element is imaged by a continuously moving eye. This eye movement blurs the retinal image, and the magnitude of the blur increases with both the speed of the tracking motion and the display hold-time.

We describe here the results of the evaluation of several digital displays. These displays are based on a variety of technologies, and they employ various techniques for reducing pixel hold-time and hence reducing perceived tracking blur. The evaluation techniques used here are part of a display assessment and calibration procedure designed to quantify the spatial and temporal properties of displays used in flight-simulator applications (Geri et al. 2004; Winterbottom et al. 2004).

8.1 Methods

8.1.1 *General Evaluation Methods*

8.1.1.1 Spatial Resolution

Display spatial resolution was characterized using procedures adopted from accepted measurement standards (VESA: Video Electronics Standards Association 2001; Winterbottom et al. 2004). The number of resolved lines was obtained by dividing the number of addressed lines by the line width needed to obtain a minimum 25% contrast. Thus when the contrast of a single-pixel width line falls below 25%, the number of resolved lines is estimated as less than the number of addressed lines. Brightness, contrast, and display gamma functions were also obtained as part of the display evaluation.

8.1.1.2 Temporal Response

Display temporal response was measured using a photodiode-based circuit and an oscilloscope. The photodiode was directed at a 30-Hz flashing square generated by our test program. A Fluke ScopeMeter was used to record the photodiode response. For temporal responses with extended onset and offset times, as were obtained for the two projectors with mechanical shutters, hold-time was defined as width of the temporal response function at half-maximum (i.e., 50%) amplitude.

8.1.1.3 Tracking Blur

Perceptual assessment of display tracking blur was based on techniques previously developed for use with flight-simulator displays (Winterbottom et al. 2004). This assessment was performed using a test-image consisting of a pair of vertical white bars, each five pixels wide (see Fig. 8.1). The bars moved horizontally across a dark field at various speeds. The bars extended across approximately one-half of the vertical extent of the display, and moved, either left-to-right or right-to-left, across the full horizontal extent of the display. Observers adjusted the gap width between

Fig. 8.1 The moving line-pair test stimulus used to assess perceived display temporal resolution. The *arrow* shows one direction of stimulus motion

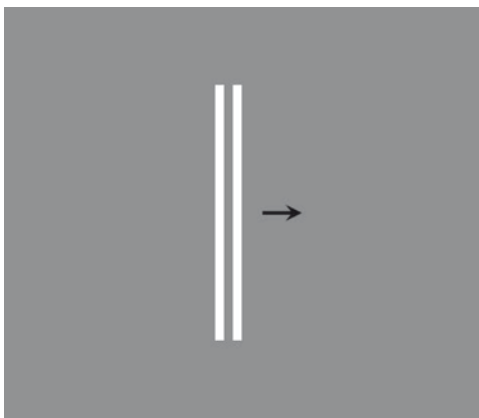


Table 8.1 Image characteristics of the projectors

Projector (shutter)	Pixel format (H×V)	Image size (deg)	Pixel size (arcmin)	Resolved pixels (H×V)
LCoS (electronic)	1,280×1,024	63×53	3.0	Not measured
LCoS (mechanical)	1,280×1,024	70×58	3.3	1,150 (H)
DLP (electronic)	1,400×1,050	73.7×59	3.2	3,976×2,814
LCD (mechanical)	1,600×1,200	71.7×61.8	2.7	1,662×2,544
LCD (none)	1,920×1,280	71.7×40.4	2.2	3,840 (H)
CRT	1,280×1,024	70×58	3.3	1,200 (H)

the bars, using the keyboard arrow keys, until there was a just-noticeable separation between the two lines. This threshold gap width was recorded when the observer pressed the space bar.

8.1.1.4 Projector Characteristics

The relevant pixel formats, projected image sizes, derived pixel sizes, and resolved pixels for each projector tested are shown in Table 8.1. If the contrast of the one-on/one-off grill pattern was 0.25 or greater, the number of resolved pixels is shown as equal to the number of addressed pixels. If the contrast was less than 0.25, the number of resolved pixels was calculated by dividing the addressed pixel count by the threshold VESA grille-line width for the horizontal and vertical dimensions.

8.1.2 Projector-Specific Methods

8.1.2.1 LCoS-Electronic Projector

Perceptual data were obtained from four observers between the ages of 24 and 55. All observers had normal, or corrected to normal, vision as determined by the

acuity, binocular vision, color vision, and phoria tests of the Optec Vision Tester (Stereo Optical Co., Inc., Chicago, IL).

A moving line-pair test image was presented through the left-eye channel of a Kaiser Proview SR80 Head Mounted Display (HMD). The right eye was occluded during testing. The monocular field-of-view was approximately 63° (H) \times 53° (V). The display imagery was produced by a FLCoS microdisplay (Model SXGA-R2-H1, CRL-Opto Ltd, Middlesex, U.K.) that provided a $1,280 \times 1,024$ (SXGA) pixel image. Thus, each pixel subtended about 3 arcmin. The microdisplay used pulse width modulation (and, color multiplexing, when required) to generate the image. Display illumination was provided by an LED array. The luminance of the displayed test images was about 2 fL and was measured by imaging the exit pupil of the HMD on the aperture of a Minolta Model 101 luminance meter.

The hold-time (i.e., the width of the luminance profile within each frame) for this projector was determined by five electronic timing sequences that controlled the LCoS panel. The form of the temporal response for this projector was similar to that of the responses shown in Fig. 8.2 (center). The timing sequences tested included a standard full-color sequence with a hold-time of 13.4 ms, a similar full-color sequence with the hold-time reduced to 8 ms, and three special-purpose monochrome sequences, representing hold-times of 3, 4, and 6 ms. The hold-times were verified directly by measuring the light output of the HMD using a photodiode placed at the exit pupil. The monochrome sequences were generated by supplying the same signal to the R, G, and B inputs of the chip. The color sequences are more complex because separate sets of signals were sent to each color channel. Due to the color-sequential design of the LCoS chip, there was some color separation at the leading edge of the trailing test line, which might have caused some confusion in defining the gap between the lines. Observers were therefore instructed to define gap width based on the un-illuminated region between the two lines rather than on changes in color between the test lines.

All perceptual tracking blur data were obtained in a single experimental session for each observer, and each of the five timing sequences was tested in a random order within each session. For each timing sequence, the 12 combinations of the six test-image speeds and two test-image directions were tested twice in random order, for a total of 24 trials for each sequence. No effects of test-target direction were found, and so the data for the two directions were combined. Thus, each data point for each observer was obtained by combining the gap-width settings from four trials (i.e., two replications for each of the two directions).

8.1.2.2 LCoS-Mechanical Projector

The hold-time for this projector was determined by a rotating metal disk containing a cutout whose width could be varied (i.e. a mechanical shutter). For the tracking blur measurement, the observers first adapted for 6–8 min to the ambient illumination of the experimental room. The moving test stimulus was then

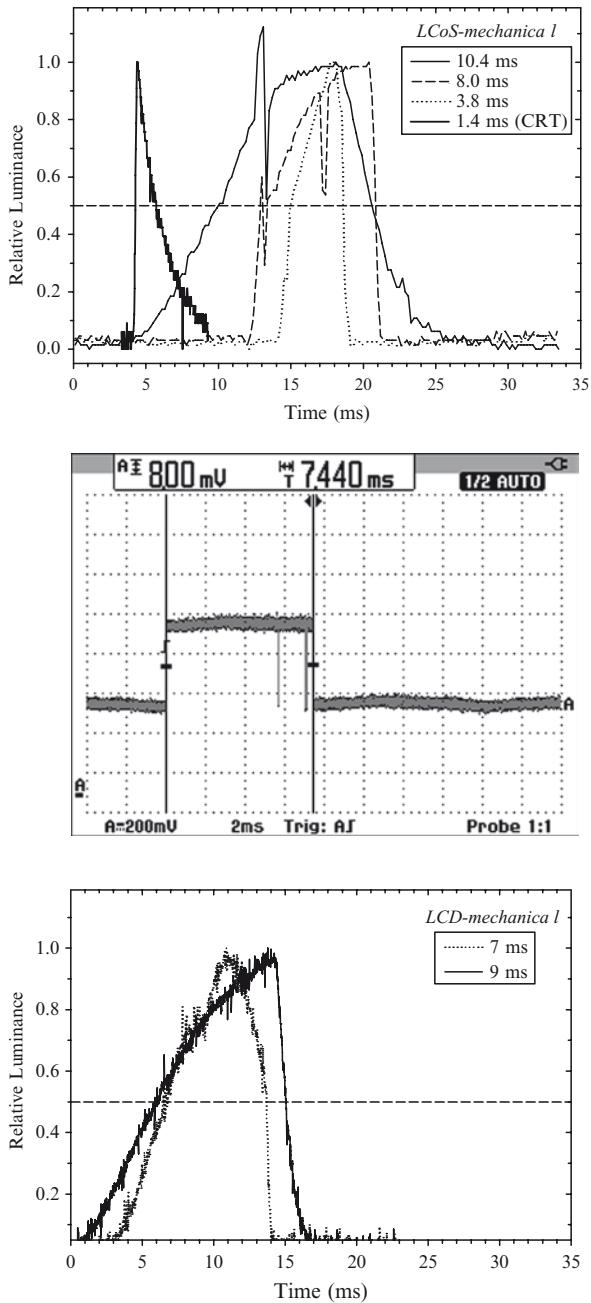


Fig. 8.2 (Top) Temporal response (i.e., light output over time) for selected hold-times for the LCoS-mechanical projector. Also shown is a typical temporal response from the CRT projector. (Center) Temporal responses for an intermediate hold-time (7.4 ms) of the DLP-electronic projector. The temporal responses for other hold-times as well those obtained from the LCoS-electronic projector had a similar form. (Bottom) Temporal response of the LCD-mechanical projector for the two short hold-times tested

presented, and the observers were instructed to track the stimulus and adjust the separation of the vertical lines such that the gap between them was minimized while keeping the width of the two lines equal. Each of the seven test-stimulus speeds was repeated ten times, and speeds were presented in a random order. The hold-times were tested in separate experimental sessions, and the order in which they were tested was also randomized.

8.1.2.3 DLP-Electronic Projector

The hold-time of the *DLP-Electronic* projector is shown in Fig. 8.2 (center). The procedures for obtaining the tracking blur measures were similar to those described above for the *LCoS-mechanical* projector.

8.1.2.4 LCD-Mechanical Projector

The hold-time for this projector was varied using a mechanical shutter similar to that of the *LCoS-Mechanical* projector and also involved metal disks with various cutouts. The shutter was synchronized with the activation of the blue and green LCD panels. Because the activation of the red LCD panel was reversed relative to the blue and green, and therefore not synchronized with the shutter, only the blue and green imagery were used for this evaluation. Again, the procedures for obtaining the tracking blur measures were similar to those described above for the *LCoS-mechanical* projector.

8.1.2.5 LCD and CRT Projectors

These projectors were used primarily for comparison purposes. The *LCD* projector was a standard off-the shelf projector and thus had a relatively long hold-time (14.6 ms at half-amplitude), and so would be expected to produce the greatest image blurring. The *CRT* projector was expected to display moving imagery with very little blurring since its hold-time was about 1.4 ms (see Fig. 8.2, top). The *CRT* is currently used in many applications, and has become the de facto standard for adequate display temporal resolution.

8.2 Results

8.2.1 Spatial Resolution

The measured resolution and the pixel format for each of the projectors tested are shown in Table 8.1. A viewing distance of 36 in was used to estimate the size of each displayed pixel in arcminutes.

8.2.2 Temporal Response

Temporal response functions from the *LCoS-mechanical* projector for various hold-times are shown in Fig. 8.2 (top), along with the *CRT* projector response. Figure 8.2 (center) shows the temporal response of the *DLP-electronic* projector for an intermediate hold-time (7.44 ms). The temporal response of the *LCoS-electronic* projector was similar in form to that of the *DLP-electronic* projector. Figure 8.2 (bottom) shows the temporal responses of the *LCD-mechanical* projector for two hold-times. The responses of the *LCoS-mechanical* and *LCD-Mechanical* projectors are qualitatively similar.

8.2.3 Tracking Blur

Tracking blur was characterized using the moving line-pair stimulus. The size of the gap to see the lines as separate was taken as a measure of perceived blur. The gap width (in arcmin) as a function of line-pair speed is shown in Fig. 8.3 for various hold-times associated with the six projectors.

LCoS-electronic projector: For the 13.4-ms hold-time, the threshold gap-width was about 7 arcmin for a test-image speed of 9.8 deg/s. The threshold gap width increased roughly linearly to about 60 arcmin as test-image speed was increased to 59 deg/s. For the 8-, 6-, and 4-ms timing sequences, the threshold gap-width was also about 7 arcmin for a test-image speed of 9.8 deg/s. This value remained between about 7 and 12 arcmin as test-image speed increased from 9.8 to 59 deg/s. The threshold gap size decreased slightly to about 5 arcmin for the 2-ms timing sequence, and remained between about 7 and 10 arcmin as test-image speed increased.

A two-way, repeated-measures factorial ANOVA was performed on the *LCoS-electronic* data. Main effects of Hold-Time and Target-Speed were both significant [$F(4,16)=98.2$, $p<0.001$ and $F(5,20)=81.2$, $p<0.001$, respectively] as was their interaction [$F(20,80)=50.5$, $p<0.001$]. Within-observer contrasts indicated that the 13.4 ms hold-time data were significantly different from the data for the other hold-times [$F(1,4)=521$, $p<0.001$]. None of the other hold-times were significantly different from one other.

LCoS-mechanical projector. For the data obtained from the *LCoS-mechanical* projector, a repeated-measures ANOVA indicated that the main effects of Speed [$F(6, 30)=65$, $p<0.001$], Hold-time [$F(5, 25)=26$, $p<0.01$], as well as the Speed \times Hold-time interaction [$F(30, 150)=22.5$, $p<0.001$] were all statistically significant. A set of a priori contrasts indicated that the 10.4 ms and 8.0 ms hold-times were significantly different from those obtained for the *CRT* projector, but that the 3.8, 1.8, and 1.0 ms hold-times were not.

DLP-electronic projector. An ANOVA for the *DLP-electronic* projector showed significant main effects of Speed [$F(4, 40)=451$, $p<0.001$], and Hold-time [$F(3, 40)=395$, $p<0.001$], and a significant Speed \times Hold-time interaction [$F(12, 40)=73$,

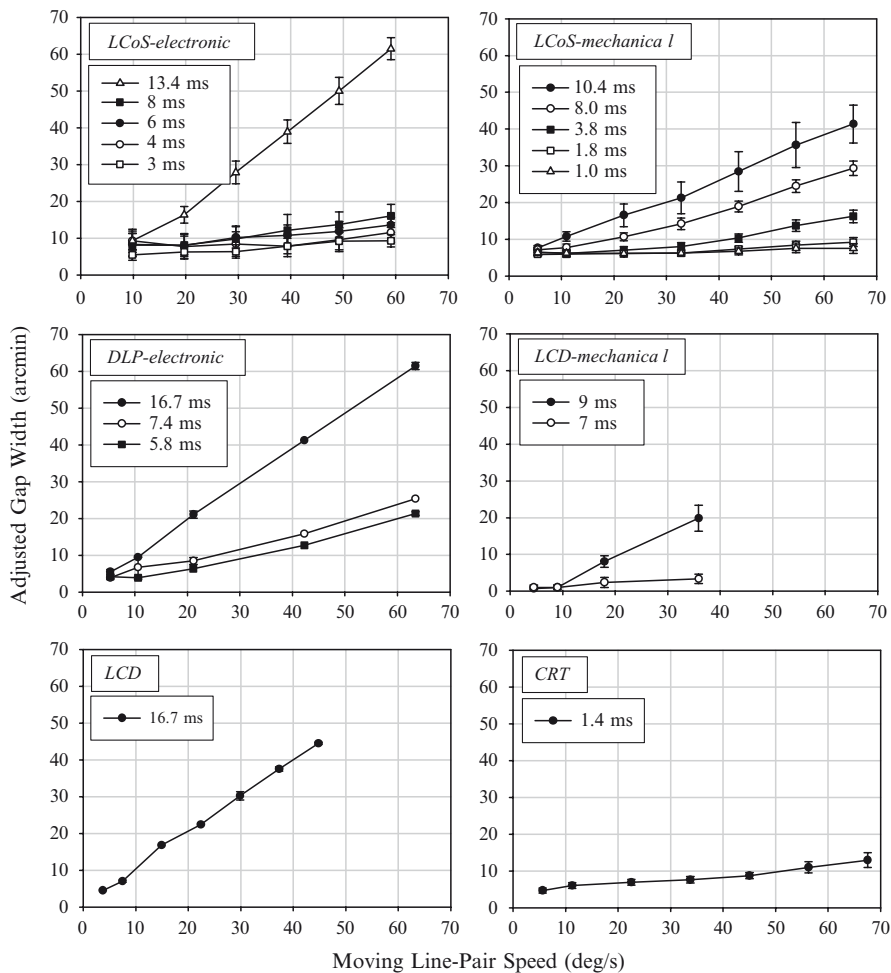


Fig. 8.3 The adjusted gap-width between the two lines making up the moving line-pair test stimulus as a function of the speed of the line-pair and the projector hold-time for each of the six projectors tested. Adjusted gap-width is taken as a measure of perceived blur. All visual angle calculations are based on the pixel formats shown in Table 8.1

$p < 0.001$]. Post-hoc tests indicated that perceived blur for the 16.6 and 7.4 ms hold-times was significantly greater than for the CRT condition. Averaged across speed, perceived blur for the 5.8 ms hold-time was not statistically different from the CRT condition. Tracking blur data (Fig. 8.4) for this projector show that perceived blur for the longest hold-time (16.7 ms) is similar to that for comparable hold-times for the LCoS-electronic projector and LCD projectors. Data for the two lower hold-times (7.4 and 5.8 ms) are similar to those obtained for comparable hold-times with the LCoS-mechanical projector.

LCD-mechanical projector. Increasing test-stimulus speed significantly increased perceived blur [$F(3, 9) = 21.3, p < 0.001$] for both the 9 ms and 7 ms hold-times with

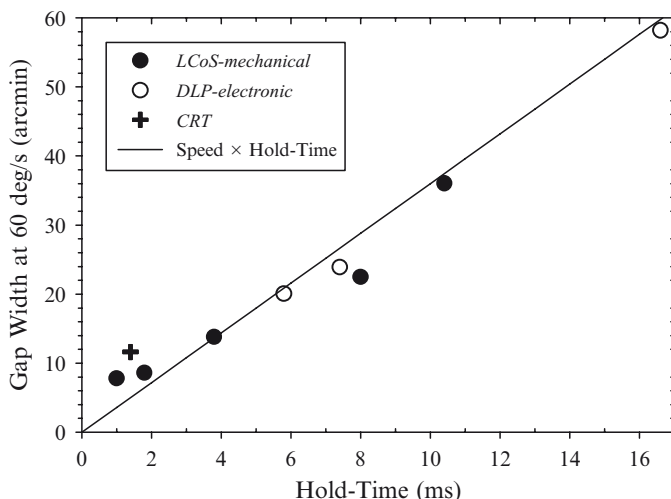


Fig. 8.4 The adjusted gap-width for a 60 deg/s (i.e., 3.6 arcmin/ms) moving line-pair for various projector hold times. The gap-widths were interpolated from the LCoS-mechanical, DLP-Electronic, and CRT data of Fig. 8.3. The *solid line* shows the moving-line blur that would be expected for each hold time for a moving-line pair speed of 60 deg/s

this projector. Reducing hold-time from 9 to 7 ms significantly decreased perceived blur [$F(1, 3)=88.9$; $p<0.01$].

LCD and CRT projectors. Tracking blur measures were obtained for the LCD and CRT projectors primarily for comparison with the other projectors tested. As shown in Fig. 8.4, the perceived blur for these two projectors was similar to that for the other displays when the hold-times were of similar duration.

8.3 Discussion

For the LCoS-electronic projector (see Fig. 8.3), a reduction in hold-time from 13.4 to 8.0 ms, resulted in a significant reduction in the blurring of a moving image. Lee et al. (2001) concluded that when the LCD light-level reached its maximum in the first half of the frame, image quality became similar to that for much faster displays such as CRTs. The present data provide some support for that, but it is not clear why continuing to illuminate the display during the second half of the frame did not result in image blurring in the Lee et al study. The discrepancy may lie in the subjective nature of the evaluation procedure used by those authors who provided no criteria for either defining the perceived blur or for concluding that additional decreases in LCD onset/offset times or hold-times were not necessary.

The data of Fig. 8.3 for the *LCoS-electronic* projector are quantitatively different from those of the other projectors tested in that the adjusted gap-width (i.e., the perceived blur) decreases abruptly when hold-time is reduced from 13.4 to 8.0 ms;

it changes relatively little for further decreases in hold-time. One explanation is that the LCoS-electronic projector is a single-chip device that can produce color breakup in the moving line-pair test target (see Sect. 8.1). This breakup appears as a blue trailing edge on the leading line of the target. Because of the lower luminance of blue light, and the lower spatial resolution of the eye to blue light, the blue trailing edge may have been difficult to distinguish from the “black” gap. This would have been true particularly at higher target speeds and longer hold-times, and it may have resulted in smaller differences in adjusted gap width under those conditions.

For the *LCoS-mechanical* projector, perceived blur was significantly greater for a hold-time of 8 ms (Fig. 8.3) than was the case for the *LCoS-electronic* projector. Data for other hold-times are consistent for both projectors. Data for the *LCoS-mechanical* projector are consistent with available data for the *DLP-electronic* projector, despite differences in the forms of their temporal response (Fig. 8.2, top vs. center). This is evidence for the validity of the output level (i.e., 50% of maximum) chosen to define the hold-time. The analogous comparison to the *LCD-mechanical* data is less conclusive. Stated hold-times for this projector were also defined by a 50% output-level. The data obtained for the 9 ms hold-time are consistent with those for the *LCoS-mechanical* and *DLP-electronic* projectors, but the 7 ms data are not. There is no clear difference in temporal response of the *LCoS-mechanical* and *LCD-mechanical* projectors, but relatively little data is available on the latter.

In Fig. 8.4, the adjusted gap-width (i.e., the perceived blur), for a moving line-pair speed of 60 deg/s, is plotted as a function of display hold-time for the *LCoS-mechanical* and *DLP-electronic* projectors, as well as for the CRT projector. The *LCoS-electronic* data were excluded from this analysis for the reasons discussed earlier. The solid line in the Figure is the expected blur-width calculated by multiplying the target speed (i.e., 60 deg/s, or 3.6 arcmin/ms) by the hold-time. The close correspondence of this line with the data suggests that a very simple model is sufficient for characterizing data of this kind. Perceived blur can be accurately predicted from measurements of projector hold time and tracking speed.

Taken together, the perceived-blur data of Fig. 8.3 suggest that for LCD, LCoS, and DLP projectors, a hold-time of 3–4 ms results in perceived blur approaching that of a CRT. However, this evaluation represents only a first step in the perceptual assessment of flight-simulator imagery displayed with various hold-times. Evaluation of the effects of hold-times on tasks that are more like those performed in flight simulators is needed before conclusive recommendations can be made. Examples of such tasks would be the detection and identification of moving targets, as well as changes in their size or motion.

Acknowledgements This work was supported by U.S. Air Force contract FA8650-05-D-6502. In addition, the following agencies provided financial and/or technical support for this research: Air Combat Command (ACC/A8AT), Secretary of the Air Force International Affairs (SAF/IA), and Defence Research and Development Canada. We thank Bill Morgan (The Boeing Co.) for developing the simulation and data collection software. The digital displays evaluated here were provided by Barco Simulation (LCD, LCoS), Christie Digital (DLP), and Kaiser Electro-Optics (LCoS). The views expressed here are those of the authors and do not reflect the official policy or position of the Department of Defense or the US Government.



Chapter 9

A Spatial Cognition Paradigm to Assess the Impact of Night Vision Goggles on Way-Finding Performance

Michelle Gauthier and Avi Parush

Abstract Although Night Vision Goggles (NVGs) support and enhance visual perception in low-light or dark conditions the optical and electro-optical characteristics of NVGs impact visual perception differently than normal light conditions. Consequently, most published studies investigating the impact of NVG use on human performance have been based on psychophysical and perception research approaches [e.g., Macuda et al. (Proceedings of the Society of Photo-optical Instrumentation Engineers (SPIE) 5442:36–44, 2004); Niall et al. (Human Factors 41(3):495–506, 1999)]. However, anecdotal reports have suggested that NVGs affect spatial orientation and way-finding, implying that NVG use also influences human cognitive functions. Few studies have systematically characterized the cognitive nature of using NVGs, particularly on spatial behavior.

This paper aims to present an empirical methodology to study NVG use on way-finding and orientation performance, by introducing a spatial cognition research paradigm. The paradigm is a between subjects design composed of two main phases (1) a learning and practice phase; and (2) tests of acquired spatial knowledge. In the learning and practice phase, participants learn the environment through active navigation and way-finding, searching for targets within a life-sized maze with or without NVGs. In the second phase, knowledge of the environment is tested with two spatial memory tests (a judgment of relative direction and map drawing task). It is proposed that such an approach can be used to study both the perceptual and cognitive aspects of using head mounted vision enhancing devices, particularly for search and way-finding tasks. Furthermore, this approach can be utilized in the comparative and acceptance testing of new vision enhancing technologies. The methodology used in this study can also be utilized in developing and assessing training guidelines and strategies that are more compatible with humans' spatial cognition processes. Some practical implications regarding NVG training and

M. Gauthier and A. Parush (✉)

Human Oriented Technology Laboratory, Department of Psychology, Carleton University,
1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6
e-mail: avi_parush@carleton.ca

possible field research to explore improvements in the design and deployment of vision-enhancing devices are also discussed.

In the current study, findings show that participants using NVGs while navigating and way-finding had longer navigation times and more excess turns compared to those not using NVGs. Moreover, a significant decrease in navigation times and navigational steps compared to controls. In the judgement of relative direction task, relative direction pointing to searched objects across rooms and to distractors in the same room was more accurate for those who performed the search without NVGs. In the map drawing task, participants using NVGs were more likely to position more objects incorrectly and receive worst map goodness scores. The results demonstrate that NVGs affect spatial navigation and way-finding performance and the acquisition of spatial knowledge. By objectively characterizing the impact of NVGs on spatial way-finding and orientation, the current results provide empirical evidence beyond that of anecdotal reports. The findings empirically demonstrate the effectiveness of the research paradigm.

Sommaire Le but du présent article est d'exposer une méthode empirique pour l'étude de la portée de l'influence du port de lunettes de vision nocturnes (LVN) sur le repérage et la performance en matière d'orientation. Les LVN améliorent la perception visuelle dans des conditions de faible lumière ou d'obscurité; toutefois, l'influence exercée par leurs caractéristiques optiques et électro-optiques sur la perception visuelle est différente par rapport à celle observée dans des conditions de lumière normale. Par conséquent, la plupart des études publiées sur l'influence du port de LVN sur la performance humaine ont été fondées sur des recherches dans les domaines de la psychophysique et de la perception. Toutefois, des signalements anecdotiques indiquent que le port de LVN influe sur l'orientation spatiale et le repérage en exerçant un effet sur les fonctions cognitives. Quelques études ont systématiquement caractérisé la nature cognitive du port de LVN en général et, plus particulièrement sur le plan du comportement spatial.

Un paradigme de la recherche en cognition spatiale est présenté pour étudier l'influence exercée par le port de LVN. Ce paradigme est fondé sur deux étapes principales : 1) l'apprentissage et l'exécution d'exercices et 2) l'évaluation des connaissances spatiales acquises. À la première étape, les participants ont découvert un environnement par la navigation et le repérage actifs, en recherchant des cibles à l'intérieur d'un labyrinthe grandeur nature avec et sans LVN. À la seconde étape, la connaissance de l'environnement a été évaluée au moyen de deux tests de mémoire spatiale. Les résultats ont indiqué que la navigation et le repérage semblaient être plus difficiles avec des LVN (groupe expérimental), les temps de navigation étant plus longs et les virages inutiles étant plus nombreux comparativement à l'exécution de ces mêmes tâches sans LVN (groupe témoin). De plus, au cours d'essais de repérage menés avec le groupe expérimental, on a observé une diminution considérable et rapide des temps de navigation et une réduction des étapes de navigation par comparaison avec le groupe témoin. Dans une tâche d'appréciation de l'orientation relative, le positionnement relatif à des objets recherchés et à des

distracteurs se trouvant dans une même pièce était plus juste avec le groupe qui n'utilisait pas de LVG. Dans une tâche de dessin de carte, les membres du groupe portant des LVN avaient plus tendance à positionner incorrectement des objets et ont obtenu les pires résultats.

Ces résultats démontrent clairement que les LVN influent sur la performance en navigation spatiale et en repérage ainsi que sur l'acquisition de connaissances spatiales. En caractérisant objectivement l'influence des LVN sur le repérage et l'orientation spatiale, les résultats actuels fournissent une preuve empirique qui vient s'ajouter aux signalements anecdotiques. De plus, ces résultats peuvent constituer une démonstration empirique de l'efficacité de ce paradigme. Il est proposé d'utiliser ce type de méthode pour étudier les aspects perceptuels et cognitifs inhérents à l'utilisation de dispositifs d'amélioration de la vision portés sur la tête, en particulier dans des tâches de recherche et de repérage. En outre, cette méthode peut être utilisée pour l'essai comparatif et l'essai de réception des nouvelles technologies d'amélioration de la vision. Enfin, la méthode employée dans cette étude peut servir à l'élaboration et à l'évaluation de lignes directrices et de stratégies en matière de formation, en accord avec les processus de cognition spatiale humaine. Certaines répercussions d'ordre pratique, intéressant la formation sur l'utilisation des LVN et la recherche éventuelle sur le terrain, sont examinées, et ce, afin d'explorer les améliorations possibles sur le plan de la conception et de l'utilisation de ces dispositifs.

Most published studies investigating the impact of NVG on human performance have been based on psychophysical and perception research approaches (e.g., Macuda et al. 2004; Niall et al. 1999). These approaches have greatly enhanced our understanding of the impact of NVGs on performance of visual tasks. While NVGs support and enhance visual perception in low-light or dark conditions, these studies have shown decrements in visual performance when using NVGs.

NVGs are electro-optical devices that enhance visibility in low light. The electro-optical components of NVGs that amplify available light create scintillating noise (i.e., a “grainy” appearance similar to static noise on a television) within the visual display and may influence depth perception, distance estimation to light sources and targets, and colour perception (Macuda et al. 2004; Niall et al. 1999). In addition, optical characteristics of NVGs such as the limited (40°) Field-Of-View (FOV) contribute to a number of problems which include an increased need for head movements (Geri et al. 2002). One might consider such a visual environment as impaired or possibly “degraded” relative to normal day vision, while significantly enhanced in comparison to unaided night vision.

Anecdotal reports have suggested that NVGs affect spatial orientation and way-finding, implying that NVG also influences human cognitive functions. Despite the paucity of empirical investigations on NVGs and spatial cognition, some evidence suggest that NVG impact on visual perception also influences spatial perception and navigation performance (e.g., McLean et al. 1997; Geri et al. 2002; Macuda et al. 2004). Indeed, studies have shown that the visual limitations due to NVGs can

result in increased spatial disorientation, poorer way-finding performance, and can increase workload (Braithwaite et al. 1998; McLean et al 1997; Salazar et al. 2003). However, few studies have systematically characterized the cognitive nature of using NVGs, particularly with spatial performance.

This paper presents a spatial cognition research paradigm to study the impacts of NVG, in general, and the impact on spatial performance in particular. The research approach is based on two main phases (1) learning and practice; and (2) tests of acquired spatial knowledge. The approach applied in the current study is based on well established methodologies used in spatial cognition paradigms (Goldin and Thorndyke 1982; McNamara 1986; Bigelow 1996; Klatzky 1998). In the learning and practice phase, participants learn the environment, with or without NVGs, while searching for targets within a life-sized maze. In the second phase, acquired spatial knowledge is tested with two spatial memory tests. This paper demonstrates the efficacy of the paradigm by presenting an experiment done with it, along with some critique and lessons learned.

9.1 Method

Study design. The experimental design was a between-participant design with participants randomly assigned to one of two conditions:

1. NVG experimental group: Participants used NVGs in a target search and way-finding task.
2. Non-NVGs control group: Participants performed a target search and way-finding task without NVGs.

Materials and apparatus. The test environment was a 36×28 ft maze constructed from wood dividers covered with a black cloth. Prominent, highly visual, and salient landmarks (i.e., garbage can, broom, plant) were placed at various locations in the maze. Some of these landmark objects were used as the search targets. Placement of the objects created “fuzzy themes” in each room including a baby room, utilities room, sports room, kitchen and hallway area (Fig. 9.1).

Participants in the NVG group were fitted with a pair of ANVIS9 (Aviator Night Vision Imaging System), F4949 set of NVGs for the target search and way-finding task. For safety reasons, the experiment was conducted in lighting conditions. The amount of light entering the NVGs was reduced with a pinhole cover that was placed in front of the NVG objective lenses which also modified the focal length of the NVG system. This also allowed items close to the observer to be perceived (e.g. 6 in. to 20 ft.). The resulting light level that the participant saw in the NVG imagery was consistent with half-moon illumination conditions (RCA 1974). The resulting field of view (FOV) was nominally set to 30°. Since the experimental environment was lit, peripheral vision was blocked with a mask that fit over the face and goggles. NVGs were mounted on a Canadian issued military helmet and secured with a NVG head-mount. Participants from the control group were also required to wear

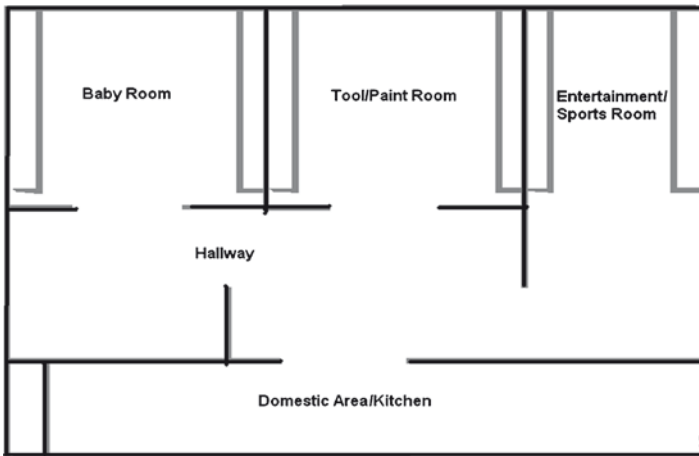


Fig. 9.1 Layout of experimental environment, including the fuzzy theme of each room

the helmet during the target search and way-finding task to control for the possible influence of wearing a helmet.

9.2 Experimental Tasks

Learning phase: Target search and way-finding task. Participants searched for various objects within a life-size maze, with or without NVGs. There were 12 search trials designed to start at different locations in the maze and with headings of every 30° such as 0, 30, 60, 90, and so on. The mean number of minimal turns (3.5) and distance traveled (33 m) were counter-balanced across the 12 search trials. The trials were then randomized within and across participants.

Spatial knowledge phase: spatial tasks. Following the learning phase, three spatial memory tests were administered to assess the acquisition and accuracy of spatial knowledge.

1. *Judgment of relative direction (JRD) task.* Participants were asked to draw on a circle the direction of a given test object relative to an imagined position within the maze (e.g., “Imagine you are facing the bicycle, standing one meter away. Point in the direction of the skis.”). Participants were instructed to draw a line from the dot in the center of the circle to the edge of the circle to indicate the direction of the objects they were asked to point to. Participants were told that the top of the circle in which they recorded their direction decision, indicated with an X, was always considered oriented in the direction they are facing (see Fig. 9.2).
2. *Distance judgment task.* Adopted from Bigelow (1996), this task asked participants to determine, from memory, which one of three objects within the experimental environment was closest, in a straight-line distance, to a specified object (Fig. 9.3).

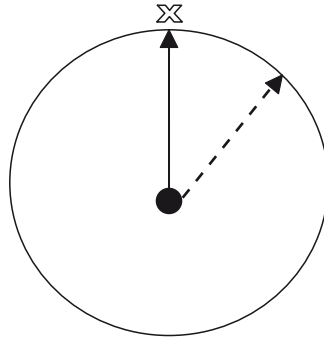


Fig. 9.2 Example of the circle used in the JRD task in which participants recorded their direction decision. The *dotted arrow* symbolizes the correct direction of the object

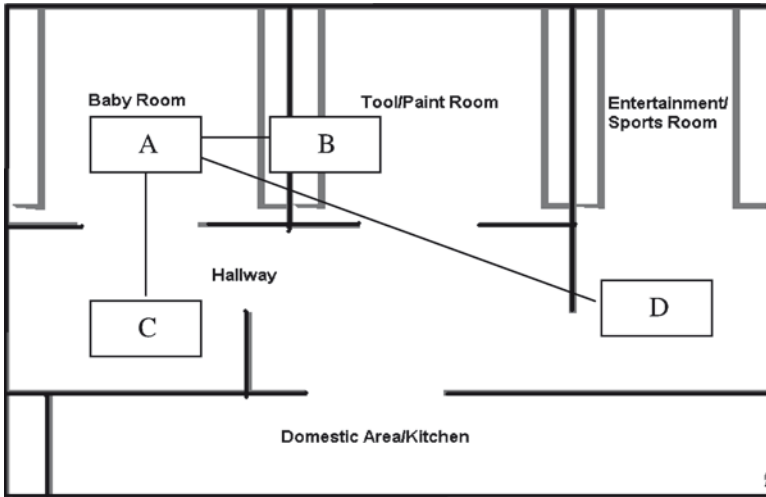


Fig. 9.3 A visual example of a distance judgment. One location, (B), was closest in straight-line distance to the test object (A), one location was clearly closest by route (C), and one location that was neither closest by route nor in a straight-line distance (D)

- 3. *Map drawing.* Participants were asked to draw, from memory, a map of the environment they learned including everything they could remember and label all objects and features.

9.3 Procedure

Participants received a general description of the experiment and its objective and were then randomly assigned to one of the two between-participant experimental conditions. Participants were tested individually. First, they performed the Peters

et al. (1995) Mental rotations test (MRT) – A to assess their spatial abilities. Participants' visual acuity was then measured using the Snellen acuity test. Participants in the NVG group were then fully briefed on adjustment and focusing procedures in an interactive session with the experimenter and spent approximately 5–10 min adjusting and focusing their goggles. Acuity measurements taken following the adjustment and focusing procedures were nominally in the range of 20–30 to 20–40, which is the typical range of acuity values with NVGs. Participants in the NVG group were then given 5 min to walk around wearing the NVGs to become familiar with the visual distortions that NVGs can produce.

Participants were then led into the maze to a pre-defined starting location and heading to begin the first trial. At the beginning of each trial, participants in the control group were blindfolded while participants in the NVG group had the NVGs turned off. While still blinded, participants were instructed to find a given target within the maze. The participant began looking for the target when the experimenter said “Begin” following the removal of the blindfold or the NVGs turned on. Once the participant had located the target, they were required to face the target perpendicularly, point to it, and announce, “Object found”. At the end of the trial, the blindfold was again placed over the eyes for the control group (or the NVGs turned off for the NVG group) and the participant was then taken to a new starting location, facing at another 30° heading and given the next task and so on. Once the 12 search/way-finding trials were completed, participants moved out of the maze and did three spatial memory tasks to assess their level of survey knowledge.

9.4 Measurements

9.4.1 *Navigation Performance*

Navigation performance measurements included the following:

1. *Time to target.* From the time that the blindfold was taken off of the participant in the control group or NVGs turned on until the participant announced “object found”.
2. *Number of excess turns (relative to optimal route).* The difference between the route taken by the participant and the shortest possible route (the minimum number of turns the participant needed to take to locate the object from starting location).

9.4.2 *Spatial Knowledge Assessment*

Orientation deviation. The absolute deviation, in degrees between 0° and 180°, between the correct directions of the test object and the direction drawn by the participant.

Distance judgment. The correct response for judging target-to-target distance.

Map drawing. Due to the subjective nature of sketched maps, the data acquired in this study were used as a qualitative assessment of the participants' cognitive maps. The following two measurements were used to assess the quality of participants' drawn maps:

1. *Proportion of correct object placement.* The ratio of number of correctly placed objects drawn on the map and the total number of objects drawn on the map.
2. *Map goodness scores.* Adapted from Billinghamurst and Weghorst (1995), goodness was rated on a scale of 1–5 by three raters who were familiar with the experimental environment but unaware of the participant's identity and group affiliation. The raters were told to rate how useful the maps would be as a navigational tool in the maze. They were told to ignore the participants' drawing ability and to focus on how well the map represented the experimental environment and the locations of the objects.

9.5 Results

9.5.1 *The Learning Phase: Impact of NVGs on Navigation and Way-Finding Performance*

To analyze the performance of search and way-finding with and without NVGs, the 12 search trials were divided into four blocks, each consisting of three trials. Analyses of the means of actual time to target and the number of excess navigational turns were performed for each of the blocks. Results showed that all participants, with and without NVGs, became quicker and more accurate in carrying out the search and way-finding task over the course of the experiment. These results simply reflected the practice effect. However, participants using NVGs while navigating and way-finding had longer navigation times and more excess turns compared to those not using NVGs. Also, over the course of the experiment, there was an earlier decrease in time to target with NVGs. Finally, there was a significant decrease in excess turns when using the NVGs compared to the control group. Taken together, the differences in performance during the search and way-finding task may reflect the level of acquired spatial knowledge, particularly survey knowledge, which was tested later.

9.5.2 *Tests of Acquired Spatial Knowledge: Level and Accuracy of Survey Knowledge*

JRD task. The judgment of relative direction task required participants to indicate, from memory, where they thought an object is located within the maze, relative to an imagined position and heading. Since targets were out of sight, the assumption

is that participants had to determine the direction of the target using their survey knowledge. Participants in both groups tended, in the judgment of relative direction tasks, to point closer to the correct direction of objects that were search targets in the navigation task compared to distracters or a mix of both. This finding may imply that better spatial knowledge is associated with explicit learning of objects that were search targets, as opposed to implicit learning of objects that were distracters. This general finding varied as a function of using NVGs and room location. The control group did better than the NVG group in judging the relative direction between targets compared to distracters, in relative pointing to searched objects across rooms, and to distracters in the same room as compared with the NVG group.

Distance judgement task. The distance judgement task required participants to indicate, from memory, which one of three objects within the experimental environment was closest, in a straight-line distance, to a specified object. The results indicated that there were no differences between the NVG group and the control group. Accuracy scores were on average around 50%.

Map drawing task. The map task required the participant to draw a map, from memory, of the environment they learned in the learning phase. While the overall map drawings were surprisingly accurate, participants who wore NVGs during the search and way-finding phase were more likely to position objects incorrectly and to receive worse scores for map goodness compared to participants who did not wear NVGs.

9.6 Discussion

The findings demonstrate the efficacy of the spatial cognition research paradigm in assessing the impact of NVGs on performance, particularly spatial cognition and wayfinding. By objectively characterizing the impact of NVGs on spatial wayfinding and orientation, the results provide well-controlled empirical evidence beyond that of anecdotal reports. The following section describes how the methodology used in this study can be used to study both the perceptual and cognitive aspects of using head-mounted vision enhancing devices, particularly for search and way-finding tasks. This is followed by a discussion on how this methodology can be applied to possible field research to explore improvements in the design and deployment of vision-enhancing devices and for practical implications regarding NVG training. Finally, a critique of the applied methodology is discussed. A theoretical discussion of the findings is presented in Gauthier et al. (2008).

9.6.1 Implications for NVG Design and Procurement

One benefit of the spatial cognition paradigm is its efficacy and high ecological validity. Most tasks performed with NVG and other vision-enhancing HMDs are spatial, and this paradigm assesses them directly and realistically. Consequently, aspects of the performance data collection and the proposed performance metrics

can be standardized and applied in making design, acceptance, and procurement decisions. This paradigm can be used for comparative studies, where comparisons can be relative to established and substantiated benchmarks or relative to other equivalent devices. NVGs with different fields of view, including panoramic NVGs, provide an example.

Another approach is to use the paradigm in virtual worlds where the direction of simulating vision-enhancing devices would be investigated by adjusting various parameters (e.g., FOV, depth of field, acuity levels) to examine their impact on spatial cognition. Design improvements can then be recommended based on acceptance.

9.6.2 Implications for Training

Simply practicing with NVGs results in improved way-finding performance, and contemporary training programs are based on this principle. However, the implications of using NVGs on spatial cognition suggest that training guidelines and strategies should be modified so that they are more compatible with human spatial cognition. Training programs can be modified to increase and maintain spatial navigation and orientation. A way to implement this strategy would be to develop a simulated navigation and way-finding training task with “pop quizzes” to frequently test NVG users’ spatial knowledge (see Parush et al. 2007; Blades et al. 2002). For instance, certain spatial tasks, such as relative pointing, verbalization or modeling (map drawing) could be performed during a search and locate task or free active exploration. NVG users should be exposed to as many different viewpoints as possible. The impact of NVGs on the acquisition of spatial cognition could then be assessed at different times throughout the training session. And finally, formative feedback could then be given to correct any likely inconsistencies in their spatial performance and memory. Feedback techniques have been successful in training NVG users to estimate distances. For instance, there was a significant improvement in distance estimates when observers had previous knowledge of the distances involved and feedback of their performance during a training session. Niall et al. (1999) confirmed the value of direct verbal feedback for distance estimation with NVGs. They showed that observers typically underestimated the true physical distance only if they had limited experience using NVGs and had no feedback on their performance.

The present methodology can also be used in developing, comparing, and assessing different training strategies with the use of vision-enhancing devices. For instance, this approach can assess and develop different training strategies on way-finding and orientation performance to compensate for a narrow FOV. The impact of different training strategies on positional awareness, head movements and scanning can be investigated. Another promising avenue is comparing simulation-based training (e.g., simulating wayfinding with NVGs on a computer screen) and the transfer of learning to NVGs in the field. Several studies have found that simulated-based training has positive transfer training effects to real world tasks (Witmer et al. 1996;

Loftin and Kenney 1995). In the simulated approach, a simulated image from a visual database is generated and displayed on a high-quality graphics workstation displaying a virtual environment on a standard computer monitor, on a head-mounted display or on a large front-mounted projection screen. If the image is to follow the observer's line of sight, a head-tracking system that updates the display in response to natural head movements is included. Visual effects, including NVG-specific effects such as FOV, monochrome vision, scintillating noise, etc could be manipulated. In the stimulation approach, imagery would be provided that is visually and spectrally correct to stimulate the actual NVGs worn by an observer. The experience should be similar to or nearly identical to viewing the real world scene with NVGs under realistic night conditions.

9.6.3 Critique and Lessons Learned

While the advantages of using the spatial cognition paradigm are straightforward, there are several limitations to this approach that should be developed and addressed.

Navigation task. The navigation performance of participants in the NVG group never reached the optimal level achieved by the control group. It was not possible to indicate if navigation performance with NVGs would ever reach the same level of proficiency as navigation without NVGs. What is the optimal performance level of NVGs without training and can we reach asymptotic levels with additional training? Many studies have assessed performance on spatial tasks longitudinally (Bigelow 1996; Gillner and Mallot 1998; Blades et al. 2002). A method often used is to assess how long it takes participants to learn an environment, determined by the number of trials needed to master some spatial task (i.e., no error rates in navigation and/or orientation performance). It would be interesting to see if it was possible to master navigation and wayfinding with NVGs and, if so, how long it would take compared with not having NVGs. Future studies could address this question by assessing how long it takes participants to learn a simulated/stimulated or real operational environment, determined by the number of trials needed to master spatial tasks (i.e., no error rates in navigation and/or orientation performance). The impact of NVGs on the acquisition of spatial knowledge could then be assessed at different times throughout the study.

Distance estimation task. There were no significant differences between the NVG and control group in the distance estimation task. Indeed, the results indicated that participants' responses were barely above chance. It is unlikely that the poor performance in the distance estimation task is due to lack of survey knowledge. It is well documented that the knowledge of the inter-object Euclidean (i.e., straight-line target to target) distances (Thorndyke and Hayes-Roth 1982; Bigelow 1996) is indicative of survey knowledge. Increased experience with an environment leads to improved performance on distance estimation tasks, indicating greater survey knowledge (Thorndyke and Hayes-Roth 1982; Bigelow 1996). The distance

estimation task used here was adopted from the study by Bigelow (1996). In Bigelow's study, performance in the distance estimation task improved with increased experience within the environment, indicating increased survey knowledge. Performance on the distance estimation task in the present investigation was barely above guessing level, indicating that this task may not be a good measure. The poor performance in this study is probably not due to lack of survey knowledge but how this task was assigned. In Bigelow (1996), participants had no time limit to make distance estimation judgments. In the current study, participants had a time limit of 15 s to make their distance judgments. This task required participants to recall four objects, their location, then make accurate estimations about their relative distance to each other. Fifteen seconds may not have given users enough time to do this task. Anecdotal reports from participants following this task were that many of their answers were guesses. Pilot tests of this task were done as a paper and pencil test with unlimited time, and judgments were quite accurate. In summary, while distance estimation tasks appear to be a reliable measure of survey knowledge, the methods used in this study need to be explored further.

Expert vs. novice users. The results of this study should be interpreted with caution when generalizing to experienced NVG users. Walking in normal visual conditions usually requires little attention: it is "automatic" (Montello 2005). Walking with NVGs, particularly when first learned, appeared to require attentional effort, thereby demanding explicit strategies. Explicit strategies have been defined as procedures that are conscious and intentional (Montello 2005). The application of these strategies, when they are first learned and applied, requires attentional resources (Montello 2005). With increased practice, tasks become more automatic, requiring fewer attentional resources. It has been reliably shown that navigation performance improves with increased practice with the environment and the task (Blades et al. 2002; Gillner and Mallot 1998; Parush and Berman 2004; Parush et al. 2007). In the present investigation NVG users quickly learned to navigate with NVGs, and became more efficient. Increased practice with NVGs resulted in a reduced demand for explicit strategies. Experienced NVG users would already have had plenty of practice with NVG, potentially leaving many attentional resources towards the search and wayfinding task. NVG experts are usually trained navigators – which could affect navigation and wayfinding.

Map drawing task: missing objects. In the map drawing task, only the location of correct object placement was measured. The emphasis of the map drawing task was to assess the effect of NVGs on participants' accuracy of the spatial location of objects and not object memory. Therefore the number of missing objects in participants' maps was not assessed. The number of missing objects could have shown the level of spatial knowledge acquired during the learning phase. And finally, attentional resources and operator confidence may have been two other factors involved in group differences in performance. Future studies should investigate the possible influence of attentional resources and operator confidence on spatial cognition.

The spatial cognition paradigm used in this study can serve as a basic experimental paradigm in testing the development or acceptance of these new technologies. Our interest is to further training methodologies and to develop technology.



Chapter 10

Psychophysics of Night Vision Device Halos

Robert S. Allison, Tracey Brandwood, Margarita Vinnikov, James E. Zacher, Sion Jennings, Todd Macuda, Paul Thomas, and Stephen A. Palmisano

Abstract In modern Night Vision Devices (NVDs) ‘halo’ around bright light sources remains a salient imaging artifact. Although a common feature of image intensified imagery, little is known of the perceptual and operational effects of this device limitation. This paper describes two related sets of experiments. In the first set of experiments, we provide quantitative measurements of Night Vision Device (NVD) halos formed by light sources as a function of intensity and distance. This characterization allows for analysis of the possible effects of halo on human perception through NVDs. In the second set of experiments, the effects of halation on the perception of depth and environmental layout are investigated psychophysically. The custom simulation environment used and results from psychophysical experiments designed to analyze halo-induced errors in slope estimation are presented. Accurate simulation of image intensifier physics and NVD scene modeling is challenging and computationally demanding, yet needs to be performed in real-time at high frame rates and at high-resolution in advanced military simulators. Given the constraints of the real-time simulation, it is important to understand how NVD artifacts impact task performance in order to make rational engineering decisions about the required level of fidelity of the NVD simulation. A salient artifact of NVD viewing is halo, the phenomenon where the image of a bright light source appears surrounded by disc-like halo. High-fidelity physical modeling of these halo phenomena would be computationally expensive.

R.S. Allison (✉), T. Brandwood, M. Vinnikov, and J.E. Zacher
Centre for Vision Research, York University, Toronto, ON, Canada M3J 1P3
e-mail: allison@cse.yorku.ca

S. Jennings and T. Macuda
Institute for Aerospace Research, National Research Council of Canada, Ottawa, Ontario, Canada

P. Thomas
Topaz Technologies, Toronto, ON, Canada M4C 2R2

S.A. Palmisano
School of Psychology, University of Wollongong, Wollongong, NSW, Australia

To evaluate the level of approximation that would be sufficient for training purposes human factors data is required.

NVD halos generated by light sources in a scene have a size that is approximately invariant with intensity and distance. Objective and subjective measures of halo geometry indicate that halo size, when halo is present, is relatively invariant of target distance or intensity. This property results in perceptual distortions and strong illusions with isolated stimuli. In complex scenes, systematic distortions of slant are predicted due to an imposed texture gradient created by the halo. We investigated this hypothesis in psychophysical experiments. The results suggest that perception of slant and glideslope in complex scenes is remarkably tolerant of texture gradients imposed by NVG halo. These results are discussed in terms of NVG simulation and of the ability of human operators to compensate for perceptual distortions.

Sommaire Effectuer des mesures quantitatives des halos autour des sources lumineuses perçus par les dispositifs de vision de nuit (DVN) selon l'intensité et la distance, décrire une méthode visant à simuler leurs effets en laboratoire et présenter les résultats des expériences psychophysiques ayant pour but d'analyser les erreurs induites par les halos dans l'estimation de la pente. Simuler avec précision la physique des intensificateurs d'images et modéliser les scènes observées à l'aide des DVN est difficile et exige beaucoup de calculs, or il faut réaliser ces activités en temps réel avec une fréquence d'images et une résolution élevées dans des simulateurs militaires de pointe. Étant donné les limites inhérentes à la simulation en temps réel, il est important de comprendre les incidences des artefacts des DVN sur l'exécution des tâches afin de prendre des décisions rationnelles techniques sur le niveau de fidélité requis. La présence d'un halo ayant la forme d'un disque autour des sources lumineuses est un artefact propre aux DVN.

Lorsque les repères indiquaient de façon évidente que la scène observée était inclinée, les participants ont perçu une pente proche de celle que l'on trouvait en l'absence de halo, tel que prévu. L'agencement régulier des lumières a permis d'obtenir différentes perspectives de profondeur, y compris la perspective linéaire, les gradients de texture, la compression (et l'effet de rapprochement) et la possibilité d'inférer un horizon implicite. Lorsque des halos sont présents dans une scène et qu'ils sont associés à une surface inclinée, leur grandeur varie, dans une certaine mesure, avec la distance apparente (constance de la grandeur). Il n'y a que peu de conflit dans ce cas particulier, étant donné que les repères de pente dominant et que l'invariance du halo est perçue comme un gradient de taille. Nous examinerons les résultats de la simulation NVG et la capacité de l'utilisateur à compenser les distorsions liées à la perception.

Night vision devices (NVD or Night Vision Goggles, NVGs) are critical to night operations for military aviators and ground forces. The devices allow forces to 'own the night' by intensifying ambient illumination, providing visibility under

reduced light conditions. However image intensifiers do not provide daytime equivalent vision and the devices suffer from a number of limitations or artefacts. For example, the image is monochromatic, contaminated by image noise at low light levels, the unusual spectral sensitivity can result in contrast inversions and field of view is limited in most devices. These limitations and artefacts presumably underlie the reported deficits in perception of space, depth and motion (for example Berkley 1992; Bradley and Kaiser 1994; Braithwaite et al., 1998; DeLucia and Task 1995; DeVilbiss et al. 1994; Geri et al. 2002; Hughes et al. 2000; Jennings and Craig 2000; Knight et al., 1998; Macuda et al. 2004; Niall et al. 1999; Rabin and Wiley 1994; Sheehy and Wilkinson 1989; Task 2001).

In high-fidelity simulation these limitations require special attention for a number of reasons. First, limits on operator perception and performance need to be simulated; any deficits in simulation fidelity need to be understood and quantified to allow an appropriate test of the perceptual capabilities required in a given situation. Second, accurate simulation of NVD characteristics can allow trainers to highlight and illustrate artefacts or limitations of the devices. Third, device limitations can influence operational procedures, which in turn may need to be rehearsed and simulated. While accurate simulation of image intensifier physics and NVD scene modeling is extremely challenging and computationally demanding, it needs to be performed in real-time at high frame rates, and at high-resolution in advanced military simulators. Given the constraints of the real-time simulation it is important to understand the nature of NVD artefacts and how they impact task performance in order to make rational engineering decisions about the level of fidelity required and level of implementation effort to commit to modeling the device.

One salient artefact of NVD viewing is halo. Halo in the context of NVDs refers to the phenomenon that a bright light source viewed through NVDs appears to be surrounded by a corona or halo that is much larger than predicted by the point spread function of the device. If a bright light, such as a NVD incompatible vehicle light, is viewed then the user typically reports seeing the image of the light source surrounded by disc-like halo. The brightness of the disc depends on the intensity of the light source and can appear transparent for relatively weak lights allowing for visibility of scenery beyond the halo. Examination of these halos is important for NVD simulation and to understand limitations on their use in operational settings (as well as to develop and train compensatory strategies).

While there have been many anecdotal reports and descriptions of the phenomenology and effects of NVD halo, published data in the open literature is sparse. Metrics such as halo intensity, transparency, symmetry, and profile shape and their dependence on source intensity, distance, shape and spectral characteristics are important, but relatively unexplored in the open literature (Craig et al. 2005; Thomas et al. 2005). With current technology, halo is a ubiquitous feature of both the built environment and natural scenes (e.g. the stars in the night sky). The phenomenon is superficially similar to the physiological halo reported in normal and diseased eyes and the coronas seen when viewing light sources through the atmosphere. However the presence of halos around numerous light sources is both an unusual and unnatural stimulus.

The effective brightness of the image of a light source depends on intensity, direction and spectral content. Informally, we have noted that halo angular size in the image is largely independent of source intensity and distance. Once a light is bright enough then a halo will appear. Increasing the intensity of this light increases the brightness of the ‘disc’ and diminishes its transparency. However, the halos of both very bright and moderately bright light sources will have the same angular diameter. With very bright sources the primary halo appears to be surrounded by a weaker secondary halo. NVD images may have other artefacts created by bright light sources (e.g. lens flare) and care must be taken not to confuse these with halo. Halos of nearby or extended sources can merge and form extended halos surrounding the extended configuration. NVD halos are generated in the image intensifier tubes. Being device artefacts, they have characteristics that are significantly different from the associated environmental features in the image. These distinctions are important and predict specific distortions of perceived environmental layout and movement. We provide quantitative psychophysical and objective descriptions of the halos formed by light sources as a function of intensity and distance and report psychophysical experiments designed to analyze halo-induced errors in estimates of slope and aimpoint.

10.1 Variation in Halo Size with Source Distance and Intensity

Evaluation of the perceptual effects of halo depends on an understanding of their image characteristics. We designed a study to compare perceptual and objective measures of NVD halo size directly using identical laboratory conditions.

10.2 Methods

A custom built light source and optical bench were designed and built to present variable intensity stimuli at a range of distances. The observer’s head was supported in a head and chin rest and placed in front of the NVD eye piece. The NVD was a standard ANVIS-9 with GenIII image intensifier tubes. The target light source was an LED mounted in a custom housing and driven by a custom driver board under computer control. A small 0.5 mm aperture was mounted at the output of the LED source to ensure the target could be regarded as a point source. Driving the LED with a pulse-width modulated digital signal permitted a wide range of light source intensity. The PWM frequency (1,000 Hz) was sufficient such that no flicker was observed.

In separate blocks of trials the target light was presented at one of three distances from the nodal point of the NVD objective (2, 4 and 8 m) at approximately the centre of the NVD field of view. At each distance the target was presented at one of five intensity levels scaled for viewing distance. The target was either presented in darkness (the room was blackened, extraneous light baffled and suppressed with

matte black cloth, paint or paper) or in the presence of an illuminated surround. The illuminated surround filled the periphery of the NVD field of view but did not illuminate the target. A gap of 10–15° was placed between the edge of the surround and the target. The purpose of the surround was to study the effect of the NVD automatic gain control (AGC) on halo size. Thus there were 30 different conditions (3 distances × 5 intensity levels × 2 background illumination conditions). These conditions were repeated five times per observer resulting in 150 measurements per observer within a counterbalanced design to control for any order effects.

For each condition we made three measurements in separate trials. Two measurements were subjective and intended to measure apparent halo angular size (image size as opposed to linear size in the world) and the third was an objective measure using a digital camera. All measures were cross calibrated to each other and to standard targets at known distances to get commensurable data in terms of visual angle at the NVD.

The first subjective measurement was direct and used a fine-grained reticule mounted in the NVD eyepiece as a gauge. The 18 mm reticule had horizontal and vertical scales extending 5 mm from the centre that were marked in steps of 5 μm per minor division with major divisions marked and enumerated every ten steps. The cross hair formed by the intersection of the axis was centred on the target and the observer was required to estimate the radius of the halo in terms of the number of divisions covered.

The second subjective measurement used an approach similar to a linear stage micrometer. A long-travel, motorized linear translation stage was mounted with its direction of travel perpendicular to the viewing direction just in front of the light source (fixed together on a rigid plate that could be moved between the viewing distances). Mounted to the stage was a long, vertically oriented illuminated line (formed by LEDs) that could be translated horizontally in front of the target. Care was taken that the line did not produce a halo. At the beginning of each trial the line was aligned visually with the halo-producing target. Then the observer moved the stage outward until the inside edge of the illuminated line of the stage micrometer was aligned with the perceived edge of the halo (a precise Vernier alignment task). The stage was moved with a stepper motor and an encoder was used to measure the stage position with a resolution of 2,048 counts per cm of translation. The measurements were repeated starting well outside the halo region to account for hysteresis effects. The halo diameter was estimated by comparing the distance between the indicated left and right edges of the halo.

For the objective measurements, a Nikon Coolpix 5400 digital camera was placed in the position of the observer's eye and used to image the NVD output through the ocular. The camera is based on a 2,592 × 1,944 pixel colour CCD sensor. The camera was set to manual focus within the shutter priority control mode and a short focal length was chosen from the camera range of 5.8–24 mm (35 mm equivalents 28–116 mm). The camera was shrouded to prevent light contamination. Five LED intensities were chosen for each distance to provide overlap in equivalent intensity ranges between distances and to produce a wide variety of halos (none, single, and double).

10.3 Objective Measures

Examples of images captured by the digital imager are shown in Fig. 10.1. Typically as light intensity increased the image changed from showing a small spot to having an obvious halo to showing signs of double halos. To estimate the halo widths the centre of mass of the image of the light source was calculated. Then cross-sections at 1° intervals were made through the centre of the spot and averaged to reduce noise (see Fig. 10.1). Estimates of the half-width of the principal halo were based on the distance between maxima in change in slope in these cross-sections and plotted in Fig. 10.2 as a function of intensity for both the background and no background conditions.

It can be seen that halo size is approximately 1.7° when halo is present and that there is no consistent variation in intensity despite a fourfold variation in distance and a 1,000-fold change in intensity. Small estimates at the lowest intensity correspond to situations where a primary halo could not be detected in the image. Moving a real object from 8 m to 2 m would have resulted in a fourfold increase in image size whereas halo size does not vary with distance.

10.4 Subjective Measures

Both the reticule and halo measures were consistent with the objective data and indicated a perceived halo of roughly 1.7° of visual angle (Fig. 10.2). Responses at 2 m were slightly smaller than at 4 and 8 m but this reduction was less than 5%.

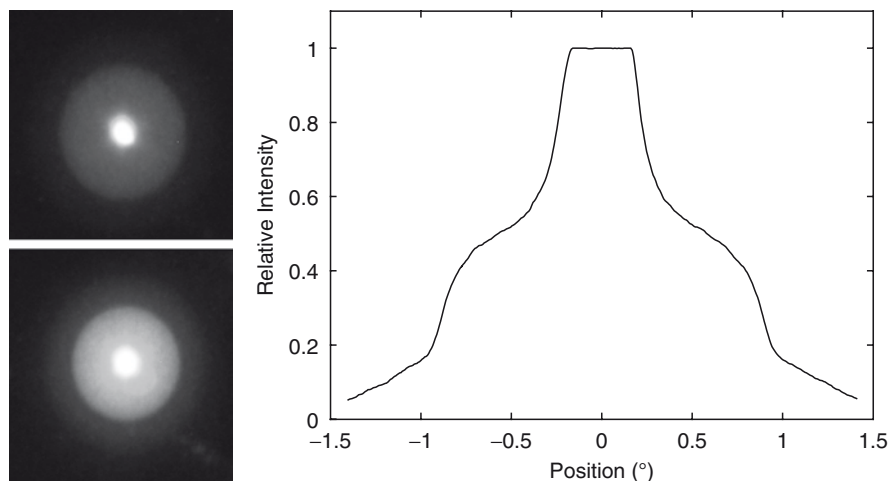


Fig. 10.1 Typical halo imagery for two intensities. On the *upper left*, the bright central spot is centered on the point source target and a 'disc' halo surrounds the spot. A secondary halo is more pronounced in the lower image. Dynamic range limitations of the camera are apparent in the saturation of the bright central spot. The apparent whitening and widening of the central spot in very bright halos is not apparent when viewing by eye. The *right-hand plot* shows a cross-section

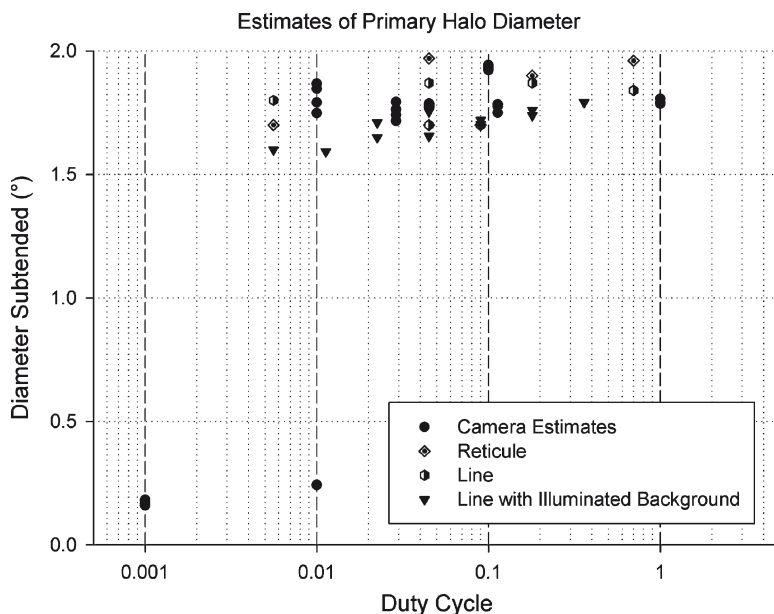


Fig. 10.2 Estimated halo width as a function of intensity

More data would be required to determine if this effect was reliable and if it was perceptual in origin. Background lighting drove the AGC of the NVD effectively. Observers were less likely to perceive single or double halos with the background light than without at any given intensity. However, the background light did not appear to have a significant effect on the halo size. Halo estimates were variable with the AGC engaged and the observers reported that the halo edge was less distinct and the judgements more difficult.

10.5 Discussion

If a point source is bright enough to generate a halo then the size (but not intensity or transparency) of that halo is effectively constant over changes in distance and intensity – at least until secondary halos are seen. Any change in apparent size is small compared to the more salient effects of halo disappearance or double halo appearance as the source intensity is decreased or increased respectively (we assume that the principal effect of distance is on effective intensity). Halo intensity profile falls with eccentricity from the centre of the spot but is remarkably flat over the ‘disc’ portion of the halo. This effect needs to be modeled precisely but provides justification for our simple disc model of halo in the psychophysical experiments.

How do scene characteristics affect the physical (as opposed to perceptual) halos generated by NVDs? Presumably, the principal parameter controlling halo generation

is effective source intensity. This allows us to generalize to larger distances and to natural scenes. Typically the device is focused at a far distance (i.e. optical infinity) and depth of field is not an issue beyond a few meters. If the distance is large enough that the target is effectively a point source within the device's depth of field, then distance cannot be a determinant of halo size per se beyond its effect on effective source intensity. Similarly, the effective intensity of a source depends on its spectral characteristics and the wavelength selectivity of the NVD photocathode. Finally, the effective intensity is also a function of the 'gain' of the NVD, which is determined by scene illumination. As discussed above, the source intensity affects the likelihood that a primary and secondary halo will be perceived and the vividness of the halo but has little effect on halo image size when halo is present.

10.6 Halos as Visual Stimuli

The fact that primary halos do not change their angular size as a function of the distance of their generating sources lends them interesting perceptual properties. The image size of a real object is determined by its egocentric distance according to the laws of perspective projection. However, halos are generated in the sensor and are therefore similar to the afterimages seen when closing one's eyes after viewing a bright light, which have a fixed retinal size. If one then gazes around an environment, an afterimage will appear to change size depending on the distance of the surface on to which it is projected. 'Emmert's law' describes how the apparent linear size of an afterimage depends on its perceived distance (Emmert 1881). Since halos also have a fixed retinal image size, Emmert's law predicts that their apparent linear size will: (a) grow as their perceived distance from the observer increases; and (b) shrink as their perceived distance decreases. Brightness is also a cue to distance and an isolated bright halo should appear nearer than a dim one even if further away. We have verified these predictions in the lab. It is important to note that complete size constancy is not to be expected and size constancy is reportedly poorer in NVD imagery than natural viewing (Zalevski et al. 2000).

When approaching a landing zone in a helicopter, a pilot must make judgements about the suitability of the terrain and their current approach. Similarly in terrain following or nap-of-the-earth flight pilots must make continuous judgements of the layout and respond accordingly. Besides affecting judgements of their depth and size, halos could have effects on the perception of the layout of the environment and surfaces within it as well as one's movement through it. Judgements of surface slant (and tilt) can be used to estimate the orientation and layout of surfaces in the environment and provide critical information during helicopter low-level flight and landing. When making judgements of slant, humans rely on a number of visual cues including perspective, binocular disparity and motion parallax.

One perspective-based cue that humans could use is known as texture gradient. If a homogeneously textured flat surface (such as a ground plane) is slanted in depth, then its retinal image will contain a gradient of texture element image size

from near to far (Cutting and Millard 1984). This gradient will be manifest in the size of texture elements, their spacing, foreshortening and their density and will be present in NVD imagery. In the case of NVD halos, patterns of lights on the ground have an added texture corresponding to the halos generated by the NVD tubes. However, as halo size (and shape) is not related to source distance, this will be in conflict with perspective-based information in the scene.

When aircraft position or orientation changes with respect to the environment, additional cues are available within the dynamic retinal image that can indicate surface slant and the motion of the observer with respect to the environment. Judgement and control of the glideslope is a critical flight task. During approach, the stream of retinal images contains changing perspective and optic flow that could be used to determine the glideslope. During an NVD approach, the perspective change in the location of objects in the optic array is consistent with the observer's self-motion (global optic flow), but the lack of local optical expansion/contraction of their halos is not. Thus, we might expect that glideslope estimation to be impaired when other cues to self-motion and environmental layout are weak.

10.7 Simulation Environment

In order to assess the effects of halo and other artefacts we have implemented a simulation environment for NVD human factors experiments. Three-dimensional computer graphic simulation of flight over modeled terrain was rendered by a cluster of Linux-based PC workstations. Scenes were modeled in 3D Studio Max based on digital terrain maps. We used an in-house developed virtual environment API (VE 2.2) to control and configure the simulation, display and input devices. The simulation was primarily visual and aircraft dynamics were not modeled. However, the simulation gave considerable flexibility for inclusion of various artefacts and for script based experimental sequencing. Extensive use of state-of-the art shader language techniques allowed real-time generation of the modeled NVD halo. The program was designed to allow for implementation of a flexible halo model. While various physical models can be implemented, halos were initially modeled as disks subtending a constant visual angle. The experiments were conducted in a large format stereoscopic virtual immersive environment. Mirrors mounted at $\pm 45^\circ$ were located in front of the left/right eyes so that each eye viewed a large projection screen located to the side. Images were projected onto the screens via BARCO 808 projectors (Barco N.V., Belgium) with a resolution of $1,280 \times 1,024 \times 100$ Hz. Each screen was driven by a separate graphics workstation in a Linux based graphics cluster. The video cards (NVIDIA Quadro FX 3000G, NVIDIA Corp. Santa Clara CA) for the displays were genlocked and the simulations synchronized.

Simulated helicopter approaches to a runway were rendered with imposed NVD effects from a physics-based model (only monochromatic display and halos were modeled). The modeled world contained a large flat plateau with a landing strip in the centre. The plateau was surrounded by simulated mountains that were unpredictable

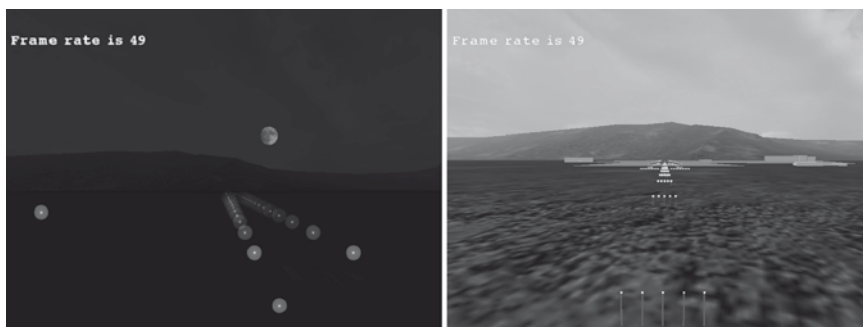


Fig. 10.3 Screen shots from simulated approaches. The light on the *left* is isolated for illustrative purposes. The frame rate indication was turned off for the experiment and was half the video refresh rate

in location, height and distance on the plateau to prevent their being used as reliable visual cues.

Stereoscopic images of simplified night scenes were rendered with halo or non-halo inducing light sources distributed on the ground plane. The intrinsic texture and perspective cues to depth in the scenes were varied by changing the regularity and configuration of the light sources. The regular lighting was composed of a runway lighting pattern based on a Precision Approach Category I lighting system (Transport Canada 1993, pp.5–25) with 282 lights arranged in a runway pattern. There was a rectilinear set of lights on the approach path, a horizontally extended set of lights marking the threshold and two rows of lights outlining the runway beyond the threshold (Fig. 10.3). The runway was 53 m wide by 1,850 m long and surrounded by gray tarmac that was visible in the daylight condition. For the irregular pattern of lights, the runway lights were randomly redistributed on the ground throughout a bounding box that bounded the regular landing light pattern.

10.8 Halo Effects and Slope Judgements

The effects of halo on judgments of slope were studied during simulated level-flight helicopter flight. The halo size estimation data allowed us to predict expected effects on texture gradients in the scene. We modeled halo image size as invariant with distance, which is true to first-order in both our objective and perceptual halo measures. Thus, the similar/identical relative sizes of these halos suggested that the observer was viewing a frontal surface. However, this interpretation was inconsistent with the information provided by the other depth cues in the display, including binocular disparity, motion parallax, texture gradients of light position and light density gradients. These latter cues will of course scale image size with depth

according to the laws of perspective. We expected that the effects of such cue conflicts on slant judgements would be most pronounced when veridical cues were weak and minimal when strong cues to slant from texture, motion and binocular disparity exist. Conversely, we expected that when surface slant was correctly perceived, the halo would be interpreted as a feature in the environment. In the following experiment we investigated the effects of superimposed halo on slant percepts when the surface was defined by regular or irregular patterns of lights and under static or dynamic conditions.

10.9 Methods

In these experiments, observers made judgments about their attitude with respect to the ground. During the simulation the observers were set at a slant with respect to the ground (via virtual camera pitch) and were required to make judgments of surface orientation in depth. The lighting pattern and halo were controlled as described above (Simulation Environment). The test scenes were either static, or depicted simulated lateral motion, or a simulated level flight approach. Observers were instructed to estimate the slope, in pitch, of the aircraft (the virtual camera) with respect to the ground (or equivalently the slope of the ground with respect to the virtual camera). Following the stimulus a fully-lit, full-cue daylight scene was displayed with random pitch angle. The observers were given control of the pitch of the virtual camera and were asked to match the attitude of the virtual camera to their estimate. The match setting was recorded and the next trial began.

To ensure that observers could perform the matching task reliably they were pre-trained. Observers trained on this task by estimating a large range of surface slants presented in full-cue daylight conditions. We reasoned that this condition would give the most reliable slant percepts and the best estimates of measurement error in the matching method. Following each presentation of the training stimulus subjects made two match settings. Observers received feedback indicating the sign of their error after their initial setting to maximize their performance. They were requested to make second setting and were not provided feedback on this setting. Pre-training was continued until the response variance reached acceptable levels and was stable. All observers required two training sessions to reach this level of performance. There was a minimum period of 24 h between training sessions.

The following manipulations were made in a factorial, repeated measures experiment: approach type (forward, lateral, static), aircraft pitch angle (-10° , -5° , 0° , 5° , 10°), light pattern (regular vs. irregular), and lighting condition (day, night no halo, night with halo). The sequence began with the aircraft positioned 53 m above ground level (172' AGL) at a distance of 431.5 m from the end of the runway. Velocity during dynamic conditions was 10 m/s. The stimulus duration for all conditions was 5 s.

10.10 Results

All observers required two training sessions. Training was rapid and all observers could reliably indicate a full-cue slant within the criterion set ($R^2 > 0.85$ with a slope coefficient greater than 0.80). A multivariate repeated-measures analysis of variance (MANOVA) demonstrated that there was no significant effect of type of approach ($F(1,5) = 3.14, p = 0.127$). Mean slant estimates are shown in Fig. 10.4. Subjects perceived increasing slant with increasing simulated slant in the daylight conditions and in night-time conditions with regular patterns of lights in the absence of halo ($R^2 = 0.77, F(1, 59) = 191.84, p < 0.001$ and $R^2 = 0.69, F(1, 59) = 257.14, p < 0.001$, respectively). There was no significant difference between the slopes estimated in daylight and night-time conditions. In all conditions the slant was underestimated and tended toward zero. In the presence of halo this underestimation was much more pronounced and slant estimates were small and not significantly related to the portrayed slant. Irregular patterns of lights resulted in a poor correlation between perceived and simulated slant. The pattern of slant estimates as a function of simulated slant under halo for regular lights was similar to that seen in all night-time irregular lighting conditions.

When slant was seen in the regular light arrangement and halos were present, observers reported a strong impression of an increase in the perceived size of the halos with simulated distance even though the halo image size was constant across the image. This was size constancy consistent with Emmert’s law. Anecdotally observers reported that in these conditions they could ‘see through’ the halo to the slanted surface, which suggested that they could segregate the slant of the surface from the frontal slant specified by the halo.

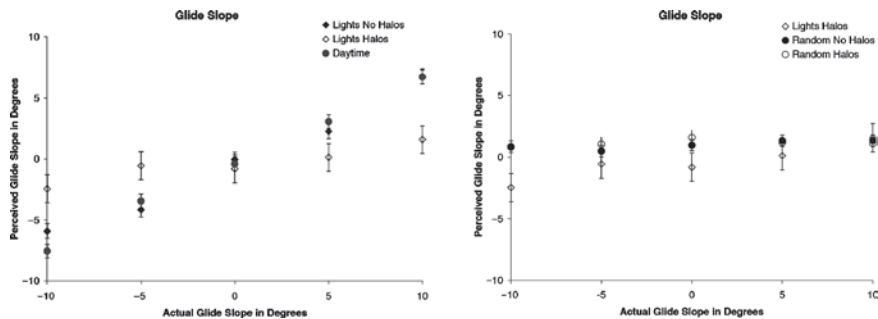


Fig. 10.4 Slant estimates as a function of simulated slant for the halo and lighting combinations averaged across six observers (mean \pm s.e.m. plotted). *Left hand panel* shows the effects of halo on slope estimation for the structured lighting (runway lights). The *right hand panel* shows estimates for the random lighting condition

10.11 Discussion

The regular pattern of lights provided a variety of perspective cues to depth including linear perspective, texture gradients, compression (and foreshortening) and the possibility of inferring an implicit horizon. When halos are present in a scene and associated with a slanted surface, their size scales with apparent distance at least to an extent (size constancy). There is little conflict here as the strong slant cues dominate and the halo invariance is seen as a size gradient. In the daylight or at night when slant cues in the scene were strong (due to a regular light pattern and absence of halo), observers perceived slant that was near the simulated slant as predicted. However, even under these conditions matched slant generally fell short of simulated slant. Gibson (1950) reported that observers consistently underestimate the slant of surfaces defined by a texture gradient in the absence of other cues. He noted that this regression to the frontal plane was much stronger for irregular textures than for regular textures. Here we have a similar finding where slant was underestimated in all cases, except that the regression was to the level ground plane rather than the frontal plane. It is likely then that either the level ground plane or the frontal plane can act as a 'norm' for slant judgements. In Bayesian terms the 'norm' would reflect a 'prior' assumption of the visual system that favours level or frontal surfaces (e.g. Knill 1998; Knill and Saunders 2003). Whether the frontal plane or level ground is preferred in Gibson's regression to the norm likely depends on the viewing situation.

This observed tendency to underestimate slant was exaggerated when we used an irregular (as opposed to regular) pattern of lights at night. With this irregular pattern, slant percepts were markedly reduced in both halo and non-halo conditions. One effect of texture irregularity is to add noise to estimates of texture gradient. Young et al. (1993) have provided evidence that under cue conflict, percepts shift to the more reliable cue (or toward a norm) when noise degrades information from the other. The current results suggest that the irregular pattern of lights provided significantly less reliable slant information than the regular pattern. We had hypothesized that changing perspective due to motion would provide particularly compelling slant information because stronger assumptions could be made than in the static case even with irregular lights (Allison and Howard 2000). However, perceived slant was as weak under dynamic conditions as under static conditions, indicating that motion was not able to compensate for the lack of regularity in the lighting pattern.

We had hypothesized a strong effect of adding halo under irregular lighting conditions. We reasoned that gradient of halo size, which was consistent with a frontal surface, should have dominated when the cues indicating the simulated slant of the ground were weakened by the use of irregular textures. However, the irregular texture manipulation may have been too strong, removing any reliable percept of surface slant change thus preventing any possibility of a halo effect (a floor effect).

In contrast, slant estimates with the regular landing light pattern at night were similar to daylight estimates and changed appropriately with changes in portrayed slant in the absence of halo. With regular light patterns, addition of halo had a

marked effect and resulted in weak slant percepts. Interestingly, the slant did not tend toward the frontal plane (consistent with the halo size gradient) but rather toward the level ground norm. Thus, the effect of adding halo to the scene was similar to the effect of using irregular rather than regular lighting patterns. This equivalence suggests that addition of halo has an effect of degrading the percept of slant from the texture gradient. The visual system thus treats the estimate as less reliable and slant matches reflect the prior bias for level ground rather than the degraded slant from texture.

Interestingly, observers viewing natural scenes sometimes report being able to see both a frontally oriented pattern of halos and to see through it to a slanted scene. Such dual percepts are sometimes seen in cue conflict situations. Van Ee et al. (2002) have claimed that for slant perception these dual percepts are alternating and bistable (like the famous Necker cube). However, for slanted surfaces the subjective impression is usually simultaneous rather than alternating. Study of the resolution of the cue conflict created by halo in scenes that are nearer frontal (i.e. a steep hill or cliff face), so that the slants specified by the halo texture gradient (a frontal surface) and the true surface are more similar and thus more likely to be combined rather than bistable, may be informative.

10.12 Halo Effects and Aimpoint Estimation

In the second set of experiments, observers watched a simulated approach to the runway and estimated the aimpoint or touchdown point given their current heading. Estimation and control of glideslope and aimpoint is traditionally thought to rely on processing of optic flow and perspective based cues in the visual image (Palmisano and Gillam 2005) that could be disrupted by halo.

10.13 Methods

The simulation environment was similar to the previous experiment except that approach to a runway along a fixed glideslope was simulated. The environment, lighting patterns and lighting conditions were the same as in the previous experiments. For each trial, the stereoscopic simulation began with the aircraft set at an altitude of either 76 m above ground level (248' AGL) or 152 m AGL (495' AGL). The aircraft began 431 m from the end of the runway and descended toward the runway along a fixed glide slope that was varied between trials. The attitude of the aircraft was aligned so that the virtual camera pointed along the simulated glideslope. The aircraft was then translated along the glideslope at a constant forward velocity of 10 m/s for 5 s. The descent rate was set by the glideslope and varied from 0.53 to 1.76 m/s.

At this point the animation stopped and a horizontally extended red line appeared across the screen drawn across the terrain at a random distances. Using buttons on a gamepad the observer adjusted the vertical screen position (i.e. perceived distance) of the line so that it appeared to be aligned with/intersect their perceived future touchdown point. The scene then disappeared and the next trial began. To ensure that observers could perform the aimpoint task reliably they were pre-trained. Observers trained on this task while estimating aimpoint in full-cue day-light conditions for a large range of glideslopes (-5° to -50°). Observers made two settings for each trial. They received feedback after their initial setting. Following this, they were requested to make another aimpoint setting without feedback. This training was continued until the response variance reached acceptable levels and was stable.

The following dependent variables were varied in a factorial, repeated measures experiment: glideslope (-3° , -4° , -5° , -6° , -7° , or -8°), altitude (high or low), light pattern (regular vs. irregular), and lighting condition (day, night no halo, night with halo).

10.14 Results and Discussion

The results are shown in Fig. 10.5. There were no significant effects of lighting condition or lighting pattern. Further, there was no significant interaction between these variables and the simulated glideslope or altitude. A prominent trend for both the high and low approach conditions was that observers appeared to overestimate distances for steeper glideslopes (mean glideslope error for -8° glideslope of 0.72 and 1.31 for the low and high conditions respectively) and underestimate aimpoint distances for shallow glideslopes (mean error for -3° glideslope of -0.69 and -0.98 for the low and high conditions respectively). This was consistent with the findings of Palmisano and Gillam (2005) who found that the bias was accentuated when the simulated ground plane was covered with randomly positioned dots (compared to a grid pattern). These authors concluded that optic flow information alone was insufficient for unbiased estimation of glideslope or aimpoint and argued that the insufficient information was a source of ‘black hole illusion’ landing errors (see also Gibb 2007).

Gibson (1950) argued that we use properties of the optic flow field, such as the focus of expansion, to estimate our aimpoint in a scene. When we move relative to the environment the image projected on the optic array (a theoretical projection surface fixed to the observer) flows out and away from a focus of expansion that lies in the direction of the observer motion. Extracting the optic flow from the retinal flow is complicated with real, mobile eyes but we could estimate our direction of travel from optic flow. In the current experiment, the global pattern of optic flow was consistent with the simulated motion but the lack of expansion of the halos was not consistent with self motion through a rigid environment. Observers could have perceived this non-rigidity as object motion of the halo light sources although this

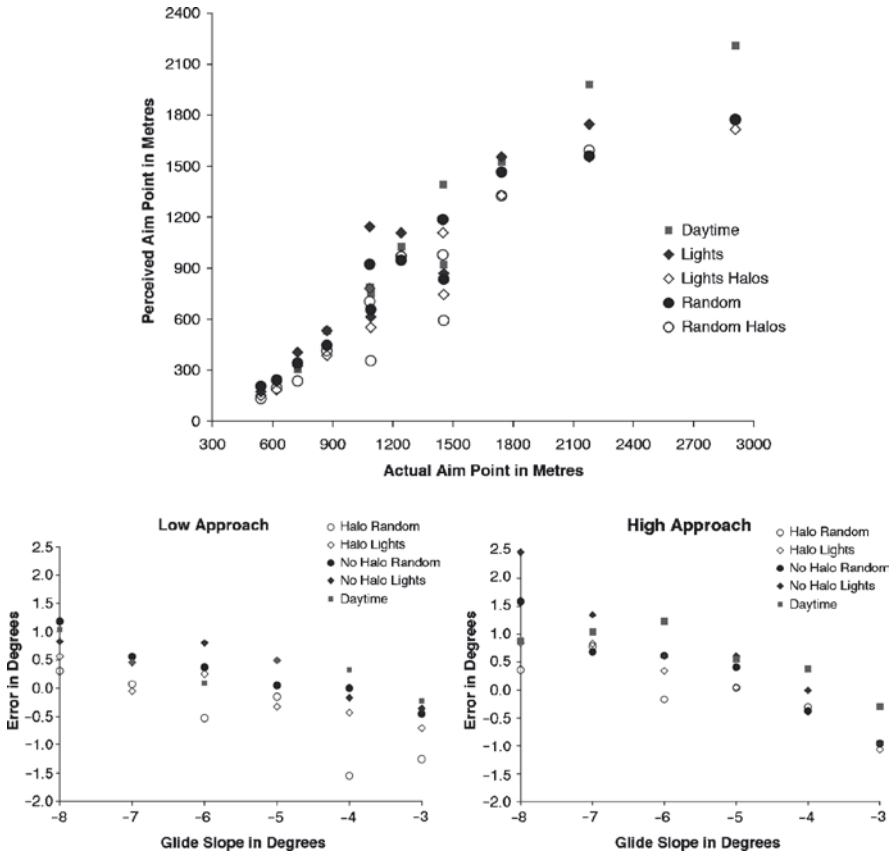


Fig. 10.5 The *top panel* shows aimpoint estimates as a function of true aimpoint. The *bottom panels* show glideslope error as a function of glideslope for the low altitude (*bottom left*) and high altitude (*bottom right*) conditions

was not reported. Instead they appeared relatively immune to halo - aimpoint bias and precision was similar to that under non-halo conditions. Thus we appear to largely ignore the halos when estimating aimpoint and instead rely on the overall pattern of motion in the landing light configuration.

This finding is similar to previous reports for time-to-contact judgements. When an object moves in depth the image of the object on the retina dilates; thus image expansion or looming is a cue to motion in depth. Gray and Regan showed observers simulated texture patches composed of arrays of circular texture elements that underwent simulated motion in depth due to image expansion (Gray and Regan 1999). When the entire pattern underwent simulated motion (simultaneous scaling of spacing, density and element size) time to contact was slightly underestimated. When the size of the circular elements was held fixed the time to contact was significantly overestimated if the elements were larger than a few minutes of arc.

However, the subjects still perceived the overall pattern as approaching which indicated dominance of the global expansion of the pattern. Harris and Giachritsis have shown that estimates of time to contact in optic flow displays consisting of clusters of dot patterns is based primarily on global rather than local image expansion (Giachritsis and Harris 2005; Harris and Giachritsis 2000). Interestingly, cells in the medial superior temporal area of the monkey brain, which is believed to be specialized for processing of optic flow, are reportedly more sensitive to the overall pattern of image motion than to the size changes in texture elements (Tanaka et al. 1989).

Besides optic flow a number of other potential cues that could indicate aimpoint position on the runway. Many of these are perspective based including splay angle of the runway (related to linear perspective), depression, runway aspect ratio cues, and so on. These configural cues depend upon spatial relations between features in the scene. If these features can be picked up in the image then halo should have relatively little effect since it will affect the image of individual lights but not their spatial configuration. Thus the splay angle of the runway (Flach et al. 1997) or its aspect ratio (Galanis et al. 2001) should not be affected as long as the halos do not obscure the light positions or merge into competing features. The fact that we did not find a detrimental effect of halo is consistent these configural strategies. However, many configural strategies depend on or are enhanced by regular patterns of lights, which was not evident from our data. Interestingly, Palmisano and Gillam (2005) did find a significant effect of light arrangement but their (non-stereoscopic binocularly-viewed) patterns had visible horizons and were sparser making them more likely to show a pattern effect. Thus, further simulation testing with both regular and irregular light patterns might provide useful insights into configural strategies that are being used to perceive aimpoint during the final approach for landing.

10.15 Conclusions

Objective and subjective measures of halo geometry indicate that halo size, when halo is present, is relatively invariant of target distance or intensity. Any change in apparent size is small compared to the more salient effects of halo disappearance or double halo appearance as the source intensity is decreased or increased respectively. Halo intensity profile falls with eccentricity from the centre of the spot but is remarkably flat over the 'disc' portion of the halo. These halo characteristics predict systematic distortions of slant and glideslope due to an imposed texture gradient and interference with optic flow processing. We investigated these hypotheses in a series of psychophysical experiments. Halos appear to act to make slant estimates less reliable but do not cause a bias toward the frontal plane when viewing ground plane surfaces. When slant is seen and halos are present, subjects report a strong impression of an increase in the perceived size of the halos with simulated distance although halos are constant size over the image. This is appropriate size constancy as found in Emmert's law. Anecdotally subjects report that they can 'see

through' the halo to the slanted surface suggesting they can segregate the slant of the surface from the frontal slant specified by the halo. Consistent with previous work in time to contact perception there appears to be little effect of halo on perceived aimpoint during simulated landing. Future work could incorporate increasingly more realistic physical halo models and address active perception and control of glideslope during simulated landing in the presence of halo.

Acknowledgements This work was performed for the NRC Flight Research Laboratory under PWGSC Contract #561982 in support of the Advanced Deployable Day/Night Simulation Technology Demonstration Project led by DRDC Toronto. Alex Tumanov and Jason Telner assisted in data collection for preliminary experiments related to this research. Portions of this work were reported in the proceedings of the SPIE Defense and Security conference held in Orlando, FL, April 9–13, 2007.



Chapter 11

Effects of Screen Resolution and Training Variation on a Simulated Flight Control Task

Don C. Donderi, Keith K. Niall, Karyn Fish,
and Benjamin Goldstein

Abstract One hundred and four inexperienced observers were trained on a PC-based flight control task in order to measure the effect of two flight simulation variables; screen resolution and training speed, on post-training performance. A throttle and joystick controlled the speed, radius and inclination angle of an orbiting F-18 chase aircraft. The control task was to overtake and intercept a target F-18 flying on another orbit with a different radius and inclination. The simulation included a realistic F-18 cockpit for the chase plane, a realistic F-18 target aircraft, textured terrain, a horizon, clouds and sun. The variables of interest were screen resolution (768×1,024, 960×1,280, 1,200×1,600 [criterion]) and simulated training speeds of 450, 900 (criterion) or 1,350 knots. Experiment 1 was carried out in a single session at criterion resolution with 50 observers. It included a pre-test at criterion speed and a training session that presented either 18 training trials at 900 knots, 18 at 1,350 knots, 9 trials at 450 and 9 at 900 knots, or 6 trials at each of 450, 900 and 1,350 knots to different groups of observers. Consistent fast (1,350 knots) training generated the largest training errors but the best post-training (criterion speed) performance. Experiment 2, with 52 observers, used the same task. The three screen resolutions were varied between observers while training speeds were varied within observers on three sessions over 3 days. Familiarization and the criterion pre-test were completed on day 1. A single training session and criterion post-test were completed on each of days 2 and 3. All permutations of pairs of 450, 900 and 1,350 knot speeds were used during the training sessions over days 2 and 3. Following the first training session, 1,350-knot training at high (criterion) resolution and medium resolution produced the best post-training performance results. Following both training sessions, the best post-training results occurred following

D.C. Donderi (✉), K. Fish, and B. Goldstein
Human Factors North Inc., 174 Spadina Avenue, Suite 202, Toronto, ON, Canada, M5T 2C2
e-mail: hfn@hfn.ca

K.K. Niall
Defence Research and Development Canada, Embassy of Canada, 501 Pennsylvania Avenue
NW, Washington, DC, 20001, USA

high resolution training at the fast (1,350 knot) training speed in both training sessions. Performance under low-resolution conditions was significantly worse than under high (criterion) or medium-resolution training. These results re-emphasize the importance of screen resolution during simulator training, and they support earlier findings that greater-than-realistic speed training improves performance on a simulated flight control task carried out at a realistic (criterion) post-test speed.

Sommaire Deux expériences ont été suivies dans le cadre d'une tâche de pilotage simulée au moyen d'un PC. Une manette des gaz et un manche à balai servaient à commander le mouvement d'un F 18 de poursuite décrivant des cercles, et la tâche de pilotage consistait à rattraper et à intercepter un F 18 cible décrivant des cercles de rayon et d'inclinaison différents. Les variables d'intérêt étaient la résolution d'écran (faible, moyenne et élevée) ainsi que la vitesse d'entraînement simulée (supérieure à la vitesse de référence, égale à la vitesse de référence (900 nœuds) et inférieure à la vitesse de référence). L'expérience un, effectuée à haute résolution, présentait dix-huit essais d'entraînement qui faisaient varier la présentation de vitesses supérieures à la vitesse de référence, égales à la vitesse de référence et inférieures à la vitesse de référence au cours d'une seule journée et mesurait les variations des performances dans le cadre de prétests et de post tests. L'entraînement à une vitesse supérieure à la vitesse de référence a donné lieu aux erreurs d'entraînement les plus importantes et aux meilleures performances de référence dans le cadre de post tests. Dans le cadre de l'expérience deux, on a utilisé la même tâche. On a fait varier la résolution d'écran d'un observateur à l'autre, ainsi que les régimes d'entraînement au sein des observateurs sur trois jours. On a procédé à la familiarisation et aux prétests le premier jour, puis on a procédé à l'entraînement à une vitesse inférieure à la vitesse de référence, égale à la vitesse de référence et supérieure à la vitesse de référence selon différentes combinaisons au cours des deuxième et troisième jours, après quoi on a effectué un post test de référence chaque jour. L'analyse des données est incomplète, mais les données partielles laissent croire que l'entraînement à résolution élevée et l'entraînement rapide permettent des meilleures performances.

The goal of the ADDNS project is to develop a high-resolution military simulator that can be quickly programmed with real-time data, making it useful as mission scenarios evolve. Two relevant questions are (1) how much fidelity is required in the real-time data? and (2) what is the best way to train using the real-time data? Both questions have been asked before and the results of previous studies are summarized in this report.

We carried out two experiments on the effects of varying screen resolution during training on a performance task carried out at a constant criterion resolution. We also varied the mix and speed of training from slower-than-realistic speed, thorough realistic speed, to a faster-than realistic speed in simulated flight. One approach to visual motor skill training emphasizes gradual, error-free introduction to the task – a

technique introduced by Pavlov while teaching dogs to discriminate ellipses (Pavlov 1928, p. 342) and further developed by students of B.F. Skinner who “faded” rats, pigeons and children from easy to hard discrimination tasks (Hilgard and Bower 1966, p. 566; Lieberman 1990, p. 204). This suggests that starting slowly might lead to smoother or faster acquisition of the criterion task. Another approach practiced by the drill instructors of the Roman legions, was to practice military exercises under harder conditions than would be met on the battlefield (Gibbon 1776). This suggests that making the training task harder than it will be will improve criterion performance. It is important to find a method, perhaps combining the Pavlovian and the Roman legionnaire techniques, that optimizes some combination of cost and success in training pilots to criterion on a simulator task. We place our experimental results in the context of the theory of transfer of training as well as in the context of what has already been learned about aviation simulation resolution and training. We first review relevant previous results about transfer of training. Then we review evidence about the use of personal computers in flight simulation and then the effects of varying training difficulty on success in transfer of training.

11.1 Transfer of Training

We focus on training variety and training difficulty as influences on transfer. Difficulty will be discussed later in the context of Above-Real-Time Training (ARTT). An early and influential study of training variety was carried out by Duncan (1958). He had people practice a manual control task which was to learn to move a lever into one of thirteen slots cut into a metal sheet placed in front of the observer. Once the lever was moved into a slot, a light lit up over the correct slot. The correct slot for each trial was determined by a set of 13 visual symbols that were displayed in random order in the aperture of a memory drum. Each symbol was associated with one of the slots, and the observers’ task was to learn which slot went with which symbol, and to move the lever into the appropriate slot. Duncan varied the total amount of practice (2, 5 or 10 days at 20 trials per day) and the number of separate symbol lists to learn (1, 2, 5, or 10 separate lists). Each symbol list was constructed by varying a single visual theme (e.g. a circle, or some letter) and was unlike every other list. For example, someone trained for 1 day on one list would experience 20 trials on that list. Someone trained for 10 days on ten lists would experience 20 trials on 1 day on one list, 20 trials the next day on another list, and so on. Someone trained on two lists for 10 days would experience 5 days (5×20 or 100 trials) of training on one list, and then switch to another list for 5 days and 100 more trials. The two transfer tasks were the same for everyone. They consisted of learning lever responses to new symbol lists. One list was non-sense syllables (not like any of the symbols used before) and the other list was symbols constructed as variations on the letter H, which had not been used before. Duncan found that the longer the training sessions, the greater the transfer, and the greater the variety of training sessions, the greater the transfer. The effect of the

amount of training on transfer performance was independent of the effect of the variety in training. People who practiced ten different symbol sets on 10 days at 20 trials a day transferred to new tasks more successfully than did people who practiced the same symbol set for 10 days, 20 trials per day. This experiment is valuable because it illustrates that, other things being equal, variety in training facilitates transfer to a new task, and that, other things being equal, the more the practice the better the transfer to a task that has something in common with the training task.

11.2 Personal Computer (PC) Based Flight Simulation

A full-task simulator is a realistic cockpit with visual displays that cover much of the visual field (e.g. Menendez and Bernard 2000). Full-task simulators are an accepted substitute for flight experience when a pilot is advancing in rank or qualifying on a new type of aircraft. By contrast, a PC-based simulator usually has a single display screen subtending perhaps $30^{\circ} \times 40^{\circ}$ of the visual field and that provides only a small fraction of the spatial and temporal resolution available to a pilot. PC-based simulators have, however, been approved as part of FAA-approved flight training. Approved PC simulators must include a specific and required set of attitude and aircraft systems controls (Koonce and Bramble 1998).

Lintern (1991) suggested that if what was being trained in flight simulation was detecting invariants in visual-motor response, then this transfer goal might be achieved on a less than full-fidelity simulator. But in using a PC-based flight simulator, Lintern and Walker (1991) found that reducing the scene detail led to worse simulated landing performance in a quasi-transfer experiment. "Quasi-transfer" means that the criterion task is also a simulator task, not an actual flight task. However, by adding flight path indicators and an augmented (e.g. predictive) aircraft symbol on the display, Lintern and Koonce (1992) demonstrated improved landing performance in a quasi-transfer task. Presumably the predictive flight augmentation and the flight path indicators, used when a flight path deviation needed correction, served to highlight the appropriate invariants for the learner. The deleterious effect of reducing scene detail anticipates a finding here.

Taylor et al. (1999) showed that PC-based training is effective in reducing expensive flight time by the substitution of less expensive simulation time during flight training. This conclusion is valid for students at the beginning stages of flight training, and is relevant to less-practiced skills, both procedural and visual-motor. However, at the expert end of flight training, Roessingh (2005) found no objectively – measured quality transfer benefit in practicing acrobatic skills on a PC-based simulator in advance of actual flight practice. He did find that the simulator-trained group completed more maneuvers per hour during flight than was the control group. It seems that PC-based training improves procedural abilities as well as or in addition to visual-motor abilities, but as suggested by the Roessingh result, the highest-level visual-motor abilities, like those required to perform accurate aerobatics, do not improve through PC-based simulation.

11.3 Above-Real-Time Training

Above-real-time training (ARTT) is a procedure in which visual-motor tasks are trained at speeds faster than those at which they are normally performed. This means that either more repetitions of the task can be practiced in the same elapsed time, or less time is required to complete the same number of practice tasks. The success of ARTT is measured by performance on the criterion task following either ARTT or normal training. Other possibilities involve mixing standard training trials with ARTT, or using various speeds during ARTT; again, success is measured by comparing the results of the various training procedures on a common transfer task following the training.

Guckenberger was the first to experiment with ARTT in aviation, although previous non-experimental reports had suggested that ARTT would work (Ali et al. 2000). He and his co-workers report two experiments involving different observer populations and different tasks. In the first and simpler experiment, 25 male college undergraduates were trained and then tested on a tank gunnery training simulator. The simulator physically duplicated the controls on a tank gun, and it presented a visual simulation of either night (infrared) or daytime gunnery targets. Five observers were assigned to each of five different time acceleration training conditions. Each condition started with five familiarization trials which were not analyzed. Then there were 15 trials under each training condition. Three groups had all trials at either 1.0, 1.6, or 2.0 times realistic target speed, a fourth group had a random mixture of the three training speeds, and the fifth group experienced sequential increases of training speeds from 1.0 to 1.6 to 2.0 across the training session. Training session tasks were divided evenly between shooting at a daytime moving helicopter, a nighttime (infrared) helicopter, or a nighttime (infrared) tank. Following the 15 training trials, each observer completed six test trials at a realistic target speed (the 1.0 speed condition during training). Two of the six test trials were with each of the three different training targets. Performance measures included the time it took to fire at the target, the percentage of hits and misses, and a composite indicator that combined azimuth and elevation errors and time to fire. Test performance measured by percentage of hits increased steadily in order of the 1.0, 1.6, 2.0, sequential and random training conditions. The easiest task, shooting at the daytime helicopter target, did not show improvement from training to test trials, but the other two (night infrared target) improved from training to transfer. Training was completed fastest by the 2.0 times realistic speed group, followed by the 1.6 speed group, then the random and mixed training groups, and slowest by the realistic speed group.

The second task was carried out using 24 USAF F-16 pilots who were trained and tested in an F-16A avionics simulator, which simulated a fixed-base F-16A cockpit and produced radar and instrument inputs and control outputs that provided realistic audio, kinesthetic and visual scenarios for electronically controlled air-to-air combat. First, each pilot was given 45 min of familiarization on the simulator, which was for an early model F-16 that few of them had flown. Then each pilot

completed three consecutive sets of ten training trials followed by four test trials. Each set was for a different simulated task. The three tasks were: maneuvering to attack (with air-to-air missiles) two simultaneously approaching enemy aircraft, following and attacking (with missile or gun) a single enemy aircraft, and detecting and avoiding an enemy electronic pulse weapon by completing a complex emergency procedure, followed by defeating a hostile aircraft. The ten training trials were followed by four transfer trials in which the same task was performed at realistic (1.0) speed. Each pilot was trained and tested on all the tasks under a single simulation speed condition: either 1.0 (realistic speed), 1.5 times realistic speed, 2.0 times realistic speed, or a random mixture of the three simulated speeds. The first training trial in each training session was considered familiarization and not analyzed. Performance on the remaining nine training trials was reported in three successive blocks of three trials each, and performance on the four transfer trials was reported in two blocks of two trials each.

Three performance measures for each task differed slightly depending on the task. A common measure on all tasks was average number of hostile aircraft kills, with a maximum of two in the combat maneuvering task and one in the stern attack and the emergency procedure task. Another common measure for the two combat tasks was time to complete the kills. The stern attack task also measured distance from target at first radar lock, which is the electronic signal that a missile or gun kill was achievable. The greater the distance, the more efficient the kill. The combat maneuvering task measured a related variable, the time to first radar lock. Again, the sooner the lock or the further from the target, the better. Missile hit percentage was a performance measure used in the combat maneuvering task. The time to complete the emergency procedure, as well as the percentage of the task that was completed correctly, were the two other performance measures in the emergency procedures task.

Although the results differed slightly across the different performance measures, a comprehensive description of the results is that for every measure and on two of the three tasks, test trial performance was worst following real speed (1.0) training, and better following any one of the 1.5, 2.0 or mixed-speed training conditions. Number of hostile aircraft kills was highest for the 1.5 and 2.0 speed groups in the emergency procedures task, and the emergency procedure was followed most accurately by the mixed and the 2.0 speed training groups. The stern attack was completed most efficiently (radar lock achieved faster) by the mixed speed group, followed by the 2.0 speed group. However a composite performance stern attack score did not differ across groups, leading the authors to surmise that this relatively easy task (from an experienced pilots' point of view) might not benefit as much from above real-time training as less practiced tasks. In the combat maneuvering task, the 1.5 speed group had the highest hit percentage and the mixed group achieved their first radar lock sooner. The authors are cautious about their results because of the variations among measures and across tasks, but given that they tested only five pilots under each speed condition, the consistency across groups outweighs the variation among measures. For all measures in most tasks, and most measures in all tasks, the faster training times led to better test trial performance (Guckenberger et al. 1993).

Another above-real-time training study was conducted using a Predator UAV flight simulator. The training included an initial explanatory video, introductory sessions of straight-and-level flight, and then another explanatory video followed by training sessions on climbing, turning, and straight-and-level flight. Three groups of seven students with little or no flight training were tested. One group was trained at real simulated speeds with performance feedback which included audio tones when altitude or heading instructions were violated as well as a strip chart printout of actual versus instructed altitude, heading and speed. Another group was trained in the same way without the feedback. A third group was trained at above-real-time simulated Predator speeds, with feedback. They completed 20 separate training flight segments in the same time as the previous two groups completed 16 segments. Each segment of the above-real-time training was completed in two-thirds the time at 1.5 times the velocity of the standard training sequence. After the training had been completed, each observer carried out two additional test flights at real-time speeds. There were no significant or important differences in performance scores between any of the groups, and therefore none of the experimental comparisons: feedback versus no feedback, or above-real-time training versus real-time training, showed any significant differences at all. (Ali et al. 2000).

In a land-vehicle ARTT study, 60 undergraduate men between 19 and 29 years old participated in a tank gunnery simulation experiment. Thirty were unpaid undergraduates who participated for course credit, and 30 were ROTC (Reserve Officer Training Corps) students who were paid \$5.00 per hour to participate. The task required each student to learn to use a tank gunnery simulator which had a realistic set of tank gun controls and computer-generated reduced-fidelity imagery of targets that differed in difficulty because they moved at different speeds (from as slow as 5 mph to as fast as 35 mph) and at different ranges (from a near range of 1,200 m to a far range of 2,400 m). Each trial in the experiment allowed the observer to take one shot at the target. Time to fire, hit percentage and aiming error were collected as performance measures.

There were 36 different target scenarios, each defined by a combination of target speed (slow or fast) and target range (near or far). Near, slow targets were called "easy", near and fast, or far and slow, targets were called "intermediate" and far and fast targets were called "difficult." The difficult training condition included three successive blocks of 36 trials, all of which were drawn from the "difficult" target set. The mixed training condition included three blocks of 36 trials, 9 of which were drawn from the "easy" target set, 18 from the "intermediate" target set and 9 from the "difficult" target set. The progressive training condition included three successive blocks of 36 trials, the first block drawn entirely from the "easy" target set, the second block from the "intermediate" target set and the third block from the "difficult" target set. Following the three training blocks, each participant completed an identical mixed test block, which like the mixed training block, had 9 "easy" targets, 18 "intermediate" targets and 9 "difficult" targets. Knowledge of results was provided during training. The entire experiment, including introduction and debriefing, took about 1½h.

The number of target hits made over the entire test block did not differ significantly across training conditions. Nor did the other measures (time to fire, azimuth and elevation errors). ROTC students performed better on the test block than the other undergraduates. The “difficult” targets in the mixed test block were hardest to hit when analyzed separately by target type. The progressive training condition produced more hits on the difficult targets than did the mixed or difficult training condition (the difference was statistically significant for the progressive versus mixed comparison), and the difficult training condition produced fewer hits on the easy targets than did either of the other two conditions (significantly fewer for both comparisons). The reason for these interactions between the training and test conditions is that in progressive training, the last block of 36 trials is all difficult, and that provided direct transfer to the difficult targets of the mixed test block. And the difficult training condition provided no examples of easy targets at any point during training, so the range and speed conditions of the easy targets in the mixed test block were completely unfamiliar to observers in the difficult training condition (Bliss et al. 1992).

Transfer of training experiments and studies of PC-based flight simulation suggest that flight simulator training should be carried out at the highest possible resolution. The data also suggest that above-real-time flight simulator training (ARTT) should provide better transfer than real-time flight simulator training. The effectiveness of training which progresses from easy to intermediate to difficult is not clear, based on the existing data. The simulation experiments carried out using progressive training or using randomly mixed easy, difficult and hard tasks are few, with few observers tested. The experiments reviewed here did not include data that would allow analysts to measure and control the effect of individual differences, and we show in our studies that individual differences are a major source of variance at least when testing untrained observers.

11.4 Method

The two questions that determined the design of these experiments were: does screen resolution during training influence criterion performance following training, and does simulated aircraft speed during training influence criterion performance following training? Screen resolution and training speed were represented by three levels of variation within each factor. Screen resolution varied from low (738 × 1,024 pixels) to medium (1,024 × 1,280 pixels) to high (1,200 × 1,600 pixels, the criterion value). Simulated training speeds varied from low (450 knots) to medium (900 knots, the criterion value) to high (1,350 knots). The criterion values (high resolution, medium speed) were used to present the target intercept task on the pre-test and the post-test trials.

The experiments required inexperienced observers, university undergraduates, to play a flight simulation game in which they controlled a chase aircraft. Their task was to intercept a target aircraft. The simulation program presented the

image of an F-18 fighter jet as a target aircraft, and a pilot's-seat view of the cockpit and out the windscreen from another F-18 fighter jet as the chase aircraft. The out-the-window world included a sky containing clouds and a sun hidden behind clouds, a brown-toned, textured earth (a flat earth, with four sides and corners) and the target aircraft. The simulated scene was presented on an IBM ThinkVision L220 20-in flat panel LCD powered by an IBM IntelliStation Z Pro type 6223 desktop computer.

The visual display area was 31 cm high by 41 cm wide. It was placed at eye height about 60 cm from the observer and at that distance subtended a 28 deg high by 36 deg wide field of view. Low-resolution ($738 \times 1,024$) pixels were 0.40×0.40 mm, medium resolution ($1,024 \times 1,200$) pixels were 0.34×0.29 mm and high-resolution ($1,200 \times 1,600$) pixels were 0.25×0.25 mm. At 60 cm, low-resolution pixels subtended about $2'14'' \times 2'76''$ of arc, medium-resolution pixels subtended $1'30'' \times 1'43''$, and high-resolution pixels subtended $1'26'' \times 1'26''$. Normal human visual acuity at medium levels of visual contrast is about $1'$ of arc (Michaels 1985). Since observers under all three resolution conditions in our experiments could discriminate one pixel from another at medium contrast levels, there was a system-imposed visual resolution upper limit which was less sensitive than average human resolution at all three resolution levels.

The flight trajectories of both target and chase aircraft were orbits around an imaginary fixed vertical. The radius and inclination of the orbit of the target aircraft (relative to an orbit parallel to the earth) could be set by the experimenter and the speed of the target aircraft could also be set by the experimenter. The initial orbit radius, inclination and speed of the chase aircraft were also under experimenter control but they were modified in flight by the observer who controlled them by manipulating a throttle (left-hand, a CH Products Pro Throttle) to control the chase aircraft acceleration, and a two-degrees-of-freedom joystick (right hand, a CH Products 568 CombatStick), to control the chase aircraft orbit radius and inclination. The throttle was programmed as a second-order manual controller. The neutral position provided no acceleration and maintained current velocity. Pushing the throttle forward produced progressively more acceleration; pulling it back produced progressively more deceleration.

Two experiments were carried out using the same software and hardware. The first experiment was a large-scale trial using an observer task that took about an hour and varied only training speed, one of the two factors of interest. The second experiment varied both factors of interest and extracted more information from the observers, who participated over three successive days.

11.4.1 Experiment 1: Training Speed

This 1-h experiment was divided into four separate parts: familiarization, pre-test, training, and post-test. All of the experimental trials were presented at high screen resolution ($1,200 \times 1,600$ pixels).

11.4.1.1 Familiarization

The first 10 min were dedicated to explanations and to obtaining informed consent. The researcher first explained the experimental objective, which was to intercept the target aircraft with the chase aircraft, and explained how to control the chase aircraft trajectory using throttle and joystick. Then the observers familiarized themselves with the interception task. Each familiarization trial either ended on a successful intercept or was timed out after a minute. On the first three familiarization trials, observers used only the throttle to approach the target aircraft and pace it without catching it. On trial one, a dynamic, schematic three-dimensional visualization of the observer and target aircraft flight paths was presented just below, but simultaneously with, the visual “out-the-window” simulation. On familiarization trials four to six, observers controlled the chase aircraft orbit radius by moving the joystick to the left or to the right. The speeds of the two aircraft were set so that if the participant properly adjusted the chase aircraft orbit radius to match that of the target aircraft there would be an interception. Should the observers fail to intercept and fly past the target aircraft on the first run, the experimenter took control of the throttle and allowed the observers a second try. The three-dimensional schematic visualization was presented below the “out-the-window” display on trial four. On trials seven through nine, observers used only the orbit inclination control (move the joystick forward or back) in order to catch the target aircraft. The speed of the simulator was set so that if the observers properly adjusted their inclination to match that of the target aircraft they would intercept it. Again, if an observer failed to intercept and flew past the target aircraft on the first try, the experimenter took control of the throttle to allow a second try. The three-dimensional visualization was presented on trial seven. To avoid forming a bias for a speed condition, Latin squares were used to set one in every three trials to the slow speed, one to the criterion speed, and one to the fast speed. The Latin squares assured that no two observers experienced the same sequence of speed variations, but that every participant experienced the three speeds equally often.

11.4.1.2 Pre-test

Observers completed four pre-test trials. Each trial was performed at criterion (900 knot) speed. Observers saw the “out-the-window” display (Fig. 11.1). Each trial ended either with a successful intercept or timed out after 2½ min.

11.4.1.3 Training

There were 18 training trials which took a total of about 30 min to complete. One of the four training trial conditions was assigned to each participant in rotation based on their order of appearance in the laboratory. Men and women were rotated independently in order to maintain a constant male to female ratio in each group.

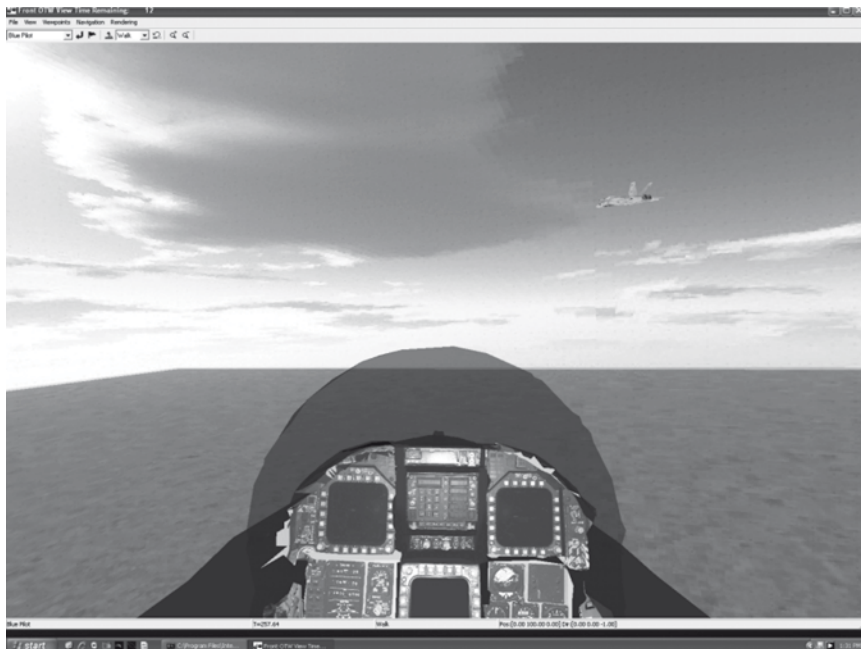


Fig. 11.1 Screen shot of the “out-the-window” display as seen during pre-test, training and post-test trials

The single independent variable, simulated speed, was varied across conditions: criterion ($n=13$), fast ($n=13$), slow-criterion ($n=13$) and slow-criterion-fast ($n=11$). In the criterion condition, all trials were performed at a simulated speed of 900 knots. In the fast condition, all trials were performed at 1,350 knots. In the slow-criterion condition, the first nine trials were performed at 450 knots and the last nine trials at 900 knots. In the slow-criterion-fast condition, there were six slow, then six criterion, then six fast trials. Observers saw the out-the-window display. Trials ended with an interception or after 2.5 min.

11.4.1.4 Post-test

The post-test was the same as the pre-test. There were four trials, each trial was performed at criterion (900 knot) speed, and the out-the-window display was used.

11.4.1.5 Observers

Ten men and 53 women were recruited from the McGill University psychology participant-pool. The results of two women were unusable because of an equipment

calibration error, and nine women and two men could not complete the experiment during the allotted hour. These 13 people were replaced by others in order to complete the experimental design. The mean age of the 50 people who completed the experiment was 21 (range 17–27, SD 5) Gender was balanced among the conditions. Informed consent was obtained from all observers prior to the experiment and observers received a full debriefing at its conclusion.

11.4.2 Experiment 2: Resolution and Training Speed

Observers completed three sessions on three successive days. On the first day they provided relevant personal information, were familiarized with the task on 36 training trials, and completed a pre-test set of 40 trials. On the second day they completed 50 training trials followed by 20 post-test trials. On the third day they completed 50 more training trials and another 20 post-test trials, following which they were debriefed.

The pre-test and post-test sessions were always completed at the average flight speed of 900 knots and the criterion resolution of $1,200 \times 1,600$ pixels. The 36 familiarization trials were completed under four Latin square combinations of equal numbers of high, medium and low resolution trials and equal numbers of low, medium and high speed trials. Each observer was always trained at the same resolution: low, medium, or high (criterion). Simulated training speeds across the first and second training days were varied within observers. All possible pairs of simulated speeds were used. Labeling the speeds as slow (S), criterion (C), and fast (F), the nine different speed conditions presented during training sessions on days 2 and 3 were S-S, S-C, S-F, C-S, C-C, C-F, F-S, F-C, F-F. Multiplying nine speed combinations times three resolution conditions gives 27 distinct experimental regimes: Low, medium and high (criterion) resolution with nine different training speed conditions tested under each resolution. Two observers (one male, one female) were assigned to each of the 27 conditions for a total of 54 observers.

Observers were recruited from the undergraduate population of McGill University in Montreal, either through an undergraduate subject pool or through response to advertising posters. Each person participated in three sessions that were completed on three successive days. Each session lasted between 45 min and 1½h, depending the day, the task and the skill of the participant.

11.4.2.1 Instructions

On the first of three testing days, each observer signed a consent form and filled out a questionnaire regarding his or her driving and video-game-playing experience. After creating a user profile in the program software for the observer, the researcher then explained the task. Observers were told that they would be playing a game of

aerial tag in which they were flying a chase aircraft. They would be chasing a target aircraft, and the goal was to catch up to and intercept the target aircraft. Observers were told that successful interception of the target aircraft would cause the image to freeze and end the trial.

11.4.2.2 Familiarization

Next, the researcher explained the structure of the circular flight orbits. Observers were told to imagine that both their aircraft and the target aircraft were tethered to an invisible vertical pole. With the aid of hand gestures, the researcher explained that the target aircraft's orbit was fixed on each trial. Following the tethered metaphor, this was because the "string length" and the "string inclination" were fixed, as was the target aircraft's speed. Then the researcher explained that orbit radius ("string length"), orbit inclination ("string inclination") and aircraft speed were all variable for the observer's aircraft, and that they were controlled by the joystick and throttle. Training began once the observer understood these concepts.

On the first day, training was carried out in four sets of nine trials each. During the first nine training trials, observers were instructed how to use the throttle to control speed. Observers' instructions were to approach and pace the target aircraft, neither intercepting nor lagging behind. The researcher used a debriefing program to show observers a dynamic, schematic three-dimensional recreation of the observer and target aircraft flight paths over two of the trials. The debriefing trials were chosen from among the nine completed trials either because they were the trials that gave the participant the most trouble, because they illustrated a recurring problem, or if the observer had no particular trouble, they were chosen at convenience.

On the next nine trials observers learned how to control the size of the observer aircraft's orbit ("the length of the string attaching the aircraft to the pole") using right and left joystick motions. Observers were instructed to catch the target aircraft while being careful not to move the joystick forward or back, and not to touch the throttle unless they accidentally passed the target aircraft. Following the nine orbit familiarization trials, observers were once again debriefed using dynamic three-dimensional re-creations of two of the preceding trials.

The third set of nine familiarization trials trained observers to control an inclined orbit. On each trial, the observer and chase aircraft were started in orbits that had the same radius and differed only in inclination from the vertical. The instructions were to catch the target aircraft by pivoting the joystick forward (thus tipping the observer aircraft and the orbit inclination down) or back (tipping the observer aircraft and the orbit inclination up). At the end of the nine trials, three-dimensional dynamic re-creations were provided for two trials; again either the two most difficult trials or, if the task was easy, two chosen for convenience.

The fourth and final set of nine familiarization trials required observers to combine all three skills for the first time, providing them with experience in using both controllers simultaneously and in controlling all three degrees of freedom

(acceleration, orbit radius and orbit inclination) simultaneously. Three-dimensional dynamic re-creations were provided for two of the nine trials.

11.4.2.3 Pre-test

Following the 36 familiarization trials, observers completed 40 pre-test trials on day 1. During these trials there was no researcher feedback and no dynamic debriefing. Pre-test trials were carried out at the criterion display conditions: high resolution (1,200×1,600 pixels) and average 900 knots aircraft speed. Each trial was started with the target aircraft in sight of the observer aircraft, but the observer aircraft's speed, orbit radius and inclination differed from the target's speed, orbit and inclination. Each trial ended with a successful intercept or after a maximum of 2 min without an intercept. The 40 pre-test trials were run in succession, with a few seconds intervening as the parameters for the next trial were automatically set up by the simulation program. Each successive pre-test trial began when the observer use a mouse to click a "Start next trial?" button on the screen. After completing the pre-test, observers were thanked and day 2's appointment was arranged.

11.4.2.4 Day 2: First Training and First Post-test

Day 2 included the first training and the first post-test session. The training condition (the resolution/aircraft speed combination for training sessions on days 2 and 3) was chosen based on the next available spot in the rotation for a male or female participant. All observers completed the same sequence of post-test trials. Like the pre-test trials, post-test trials were at high resolution and criterion aircraft speed. Each trial began with the target aircraft in sight of the observer aircraft, but with the observer aircraft having a different speed, orbit radius and inclination than the target. Each trial ended with a successful intercept or after a maximum of 2 min without an intercept. The 20 post-test trials were run under the same conditions as the previously described pre-test trials.

Observers were told that that the training and post-test trials were just like the last set of nine familiarization trials they had completed the previous day and that they were to intercept the target aircraft using the skills that they had learned the day before. When they had finished the 50 training trials, they completed 20 post-test trials. If they asked about differences between the training and post-test trials, they were told that all would be explained at the end of Day 3. Upon completion of the post-test trials, the next day's appointment was confirmed.

11.4.2.5 Day 3

Day 3 was like day 2, the only differences being that the resolution/speed combination for the day 2 training trials depended on the condition being completed by the observer, and there was a complete debriefing after all the trials were done.

11.5 Results

The simulation software recorded the x, y, z coordinate positions of the target and the chase aircraft to a text file on each trial every 1/10 s. These data were converted offline into three different measures of difference between the positions of the chase aircraft and the target aircraft. The first measure was the distance in kilometers between the chase and the target aircraft (Distance, D). The second measure was the difference in radius (in km) between the orbits of the two aircraft (Radius, R). The third measure was the difference between the angles of inclination (in radians) between the orbits of the two aircraft (Angle, A). In the process of intercepting the target aircraft, each of these measures should approach zero. Each of the three measures was summed separately over each experimental trial, and the performance measures were the summed values of A, R and D recorded for each observer on every trial. The greater the number, the worse the performance.

11.5.1 Experiment 1: Training Speed

The distributions of A, R and D, were positively skewed. In order to correct for the unequal variances accompanying the positive skew, all statistical analyses were carried out on the common logarithms of the data scores, and the summary results are presented as the geometric mean error scores (the antilog of the mean log error scores) accompanied by the standard error of the mean for each condition. Analysis of covariance was used to calculate the significance of the differences among the training and the post-test error scores on the three measures across the four conditions. The covariate for each observer on each measure was the pre-test log error score summed over four trials, and the dependent measures for each observer were the log error scores for the 18 training trials summed over three blocks of six trials each, as well as post-test log error score summed over four trials.

For all measures, the fast training condition produced the largest error scores during training and the smallest error scores during the post-test trials. Statistical analysis confirms the significance of this result. The effects of both training conditions (Fast, Criterion, Criterion-Slow and Criterion-Slow-Fast) and error measures (Angle, Distance and Radius) were significant determiners of the corrected mean log post-test error scores (in every case, $p < .0001$). But the interaction between training conditions and measures was nonsignificant. In other words, the differences among conditions were the same, and were all significant, for all three error measures. Orthogonal contrast analysis shows that each of the log post-test error means for conditions was significantly different (minimum $p < .02$) from every other.

Experiment 1 showed that 18 training trials of above-real-time training in a flight simulation task produced much better performance on a post-training performance test than did 18 training trials carried out at the real simulated speed, or than a progression

of nine slow speed to nine criterion speed trials, or than a progression from six slow to six criterion to six above-real-time training trials. All trials were carried out at the highest available resolution (1,200×1,600 pixels). The implications of the results of this experiment will be considered in the discussion to follow.

11.5.2 Experiment 2: Resolution and Training Speed

The distribution of error measures was skewed, averaged over measures, for each observer in Experiment 2. In order to limit the biasing effect of variance differences across conditions and observers on the statistical evaluation of the results, all of the raw error measures were transformed to common logs and the statistical analyses were then carried out on the log transformed data. Results are graphed and tabulated as the geometric mean values, which are the antilog of the mean log transformed data for each condition. As explained in the Methods, the experiment was carried out over 3 days. The first day included familiarization and pre-test trials. Data from familiarization trials were not analyzed. The 40 pre-test trials were carried out at criterion (1,200×1,600 pixels) resolution and criterion (900 knots) simulated airspeed. The log mean error pre-test trial data were used as covariates for the training and post-test trial data collected on days 2 and 3. Post-test trials, like the pre-test trials, were carried out at criterion resolution and speed. On day 1, three resolution conditions (low, medium and high) were combined with three speeds (slow, 450 knots [S], criterion, 900 knots [C] and fast, 1,350 knots [F]) to produce nine training conditions. Six observers were tested in each of the nine conditions.

On day 3, each of nine speed conditions tested on day 2 was subdivided by adding three new speed training conditions, so that on day 3 there were a total of nine different speed training conditions. They were: C-C, C-F, C-S, F-C, F-F, F-S, S-C, S-F, and S-S, where the first letter refers to the speed training condition on day 2 and the second letter refers to the speed training condition on day 3. On day 3 each observer continued to be trained on the resolution to which he or she had originally been assigned. Thus there were a total of 27 conditions on day 3: the nine speed conditions described above, each presented at one of three resolutions: high, medium or low. One male and one female observer experienced each combination resolution and speed training on day 3.

High resolution and fast training speed independently produced the smallest cumulative error score on the first post-test trials, and high resolution independently produced the smallest cumulative error score on the second post-test trials. The results for all of the combinations of training resolution and training speed are presented in Table 11.1.

Analysis of covariance of the least squares log means confirmed that there were significant differences among the three measures (Angle, Radius and Distance, $p < .0001$), across the three resolutions (Low, Medium and High [Criterion], $p < .02$) and across the three training speeds (Slow, Medium [Criterion] and Fast, $p < .04$), a

Table 11.1 Experiment 2: geometric mean average error per trial, for first training and first post-test sessions, by resolution and training speed

Resolution	Training speed	Training 1			Post-test 1		
		GM	(se)	Resolution GM	GM	(se)	Resolution GM
Low	Slow	198.1527	1.0575	197.6059	145.8143	1.0735	145.8143
	Criterion	179.2669	1.0578		135.4877	1.0737	
	Fast	216.7205	1.0575		156.7473	1.0732	
Medium	Slow	229.4563	1.0578	230.4623	161.1388	1.0735	153.7800
	Criterion	233.2384	1.0578		176.6038	1.0735	
	Fast	222.2798	1.0590		125.2853	1.0752	
High	Slow	178.2379	1.0578	210.9110	138.6117	1.0735	140.1120
	Criterion	207.6825	1.0583		147.1635	1.0740	
	Fast	261.3966	1.0575		137.8796	1.0732	

significant interaction between training speed and resolution ($p < .0001$), but there were no interactions between measure and any other condition. This justifies presenting data averaged over the three error measures.

11.5.2.1 Interaction Between Training Speed and Resolution

The significant interaction between the factorial variables of resolution and training speed can be resolved quantitatively. The three training speeds (450, 900 and 1,350 knots) bear a simple 1:2:3 quantitative relationship to each other. Based on the total number of pixels on each screen, the three training resolutions of $738 \times 1,024$, $1,024 \times 1,200$ and $1,200 \times 1,600$ are in the ratio 1:1.67:2.41.

If we simplify the quantitative relationship among the factors by expressing the resolution and speed values on each factor as the appropriate ratio value, and if we then express the interaction between the two factors by multiplying the ratios of the relevant main effects, we have a quantitative expression for the interaction of the resolution and speed factors which can replace its factorial expression. Table 11.2 illustrates the conversion between factorial and quantitative terminology for each of the nine interaction conditions.

The significant resolution \times speed interaction on post-test 1 can now be expressed graphically, as in Fig. 11.2.

From this figure it is clear that, across all combinations of training resolution and training speed, the higher the first day training resolution and the higher the first day training speed, the better the performance on the first day post-test. The medium-resolution condition produced the largest cumulative error score and the low-resolution condition the smallest. All three of the resolution condition covariance-corrected log mean errors were significantly different from each other (minimum $p < .02$).

Table 11.2 Experiment 2, day 1: transformation of training resolution and training speed factors into a quantitative value for each interaction combination

Training factors				Quantitative expression
Resolution		Speed		
Factorial	Quantitative	Factorial	Quantitative	Product
Low ($738 \times 1,024$)	1	Low (450)	1	1
	1	Medium (900)	2	2
	1	High (1,350)	3	3
Medium ($1,024 \times 1,200$)	1.67	Low (450)	1	1.67
	1.67	Medium (900)	2	3.34
	1.67	High (1,350)	3	5.01
High ($1,200 \times 1,600$)	2.41	Low (450)	1	2.41
	2.41	Medium (900)	2	4.82
	2.41	High (1,350)	3	7.23

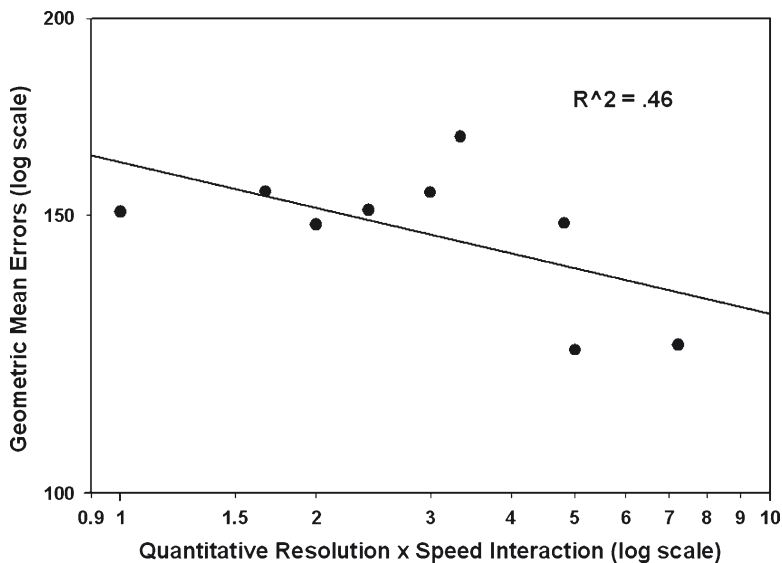


Fig. 11.2 Experiment 1, first day: decrease in geometric mean error as a function of the interaction between training speed and training resolution

11.5.3 *Speed and Resolution Interactions on the Second Post-test*

A subset of the nine observers who experienced each of the three first session training speeds (S, C, and F) at each training resolution (L, M, H) experienced each of the possible second training session speeds at the same training resolution. There were nine training speed conditions under three different resolutions for a total of 27 different combinations of training speed and resolution. Two observers – one man and one woman – provided data for each of those combinations. Data analyses were carried out on the log error scores, and the data are reported as geometric means (the antilog of the mean log scores) for each condition. People trained at high resolution generated fewer errors during the second post-test, a result consistent with the outcome of the first post-test session.

Analysis of covariance showed a significant effect of resolution ($p < .001$), and orthogonal comparisons showed that high and medium resolutions were significantly different than low resolution ($p < .0001$) but not significantly different from each other. Training speed on both the first and second training sessions influenced post-test two error scores significantly ($p < .0001$) as did the interaction between first and second training speeds and resolution ($p < .0001$). Error measure was a significant source of variance ($p < .0001$), but did not interact significantly with training speed on the first or second post-tests, nor with resolution. But there was a significant four-way-interaction among first and second session training speed, resolution, and error measure ($p < .02$).

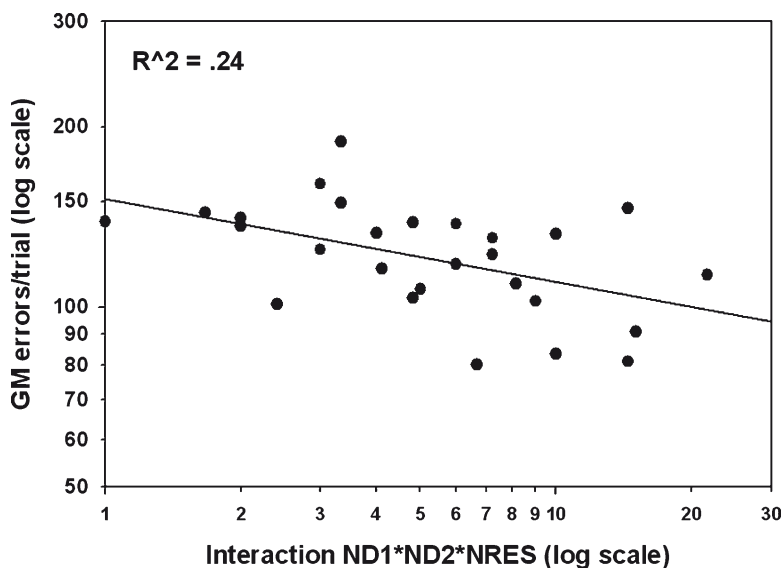


Fig. 11.3 Experiment 2: the effect of the quantitative measure of the training speed $1 \times$ training speed $2 \times$ resolution interaction on second post-test GM errors per trial

We used ratio quantification to find the relationship between the three-way interaction corresponding to each of the 27 conditions (nine speed combinations \times 3 resolutions) and the geometric mean cumulative errors on the second post-test. The result of this analysis is shown in Fig. 11.3. The linear effect plotted in Fig. 11.3 is significant ($p < .01$).

11.5.4 Individual Differences

There were high correlations across observers among the error scores in all five conditions of the experiment. Table 11.3 presents the correlations based on 162 values for each condition: the angle, radius, and distance log error scores for each of 54 observers. They are all significant at $p < .0001$. This shows that most of the variance accounted for in the experiment was due to differences among the error measures and individual differences. We also found that if, along with pre-test performance, we added first training session performance as a covariate in the factorial analysis of the first post-test, we obtained the same pattern of significant factors and interactions, with even higher levels of significance. If, along with pre-test performance, we added first and second training performance as covariates in analyzing the second post-test, we obtained the same pattern of significant factors with even higher levels of significance. Including individual differences in training

Table 11.3 Correlations between each session across individual log error scores

Test	Correlations			
Training 1	0.96007			
Post-test 1	0.94192	0.96039		
Training 2	0.92797	0.95521	0.95404	
Post-test 2	0.92623	0.93437	0.96288	0.96236
Test	Pre-test	Training 1	Post-test 1	Training 2

performance as covariates did not eliminate the effects of training; it simply made those effects clearer by removing even more of the individual variability within the training sessions.

11.6 Discussion and Conclusion

We conclude from both the experimental literature and our experiments that flight simulation training should be carried out at the highest available resolution in order to maximize transfer to an actual flight task. Based on our own experiments, we also conclude that ARTT is better than realistic training speed in improving both flight control skills, and based on the work of Guckenberger et al. (1993), that ARTT is better for both flight control and procedural skills. We did not review any experimental work, nor did we collect and analyze any experimental data, that led us to predict advantages from simulation training that starts with easy tasks and ends with above-real-time tasks, nor from training that presents easy, intermediate and hard training tasks in random order. The few experiments that have been carried out on mixed or progressive training tested few subjects and did not provide adequate statistical control for the large individual differences we demonstrated in our own work.

For the ADDNS project, we recommend the highest simulation resolution possible and we recommend flight control training at faster-than-real-time speeds. We think the advantages or disadvantages of progressive or randomly mixed training versus above-real-time training need to be studied further, using experimental designs that provide adequate experimental or statistical control for individual differences. Individual differences will be much less influential if experimental observers in future simulation studies are pilots, because they have been selected and trained to eliminate much of the variability that is the necessary consequence of testing unselected and untrained observers.



Chapter 12

Video-to-Reference Image Indexing

Vitaly Zholudev and Richard Wildes

Abstract This work is concerned with registration of data in geospatial databases, especially with registering images taken by different sensors and from different viewpoints of the same scene. This research has many applications in creating and updating maps, surveys and other geospatial data sources. While considerable work has been performed in image registration, extant approaches break down as viewpoint and/or sensor vary beyond relatively small changes. Further, extant image-based registration technology has only been demonstrated to support video-to-reference image alignment when initialized to within several hundred pixels of the correct result. Due to errors, drop outs and otherwise limited availability of telemetry, ineffective a priori knowledge of relative video/reference image alignment is a real-world problem.

The developed approach allows video-based image descriptors to index directly into a reference image database. Successful indexing into the database implies that approximate position of the video relative to the reference has been recovered; this information can then serve to initialize extant technology for video-to-reference registration. The approach provides uniform representation of video and reference imagery and a corresponding method for quantifying the goodness of match between two image samples. The representation combines image appearance, characterized in terms of texture defined regions, and image geometry, characterized in terms of spatial relationships between textured regions. By construction, the matching method is robust to a range of photometric and geometric distortions between image sources, including changes in grey-level contrast and affine geometric transformations. The developed approach has been algorithmically specified and instantiated in software. Empirical evaluations with a reference image database derived

V. Zholudev (✉)

Department of Computer Science and Engineering, York University,
4700 Keele Street, Toronto, ON, Canada, M3J 1P3
e-mail: vitaly@cse.yorku.ca

R. Wildes

Centre for Vision Research, York University, Toronto, ON, Canada

from orthoimages of distinct geographic locations that is indexed via synthetic aerial video document the promise of the approach.

Technical approach: The approach to indexing video to reference imagery involves a novel integration of regional appearance descriptors and geometric relationships between regions. Previous image indexing research has tended to concentrate on only appearance or only geometry. The philosophy of the proposed approach is that the combination of these two sources of information will yield a more discriminative approach than either one alone. Image appearance is characterized in terms of texture descriptors. Texture is employed as it is manifest in any imagery source (e.g., unlike colour, which depends on capture of multiple spectral bands). Geometric relationships are captured in terms of an affine invariant, which provides robustness to the range of variation that is anticipated in matching imagery of concern, i.e., aerial video to orthophoto reference imagery.

To combine geometry with appearance, geometry is used as a prefilter for appearance-based matching. In particular, coarse quantization of the geometric parameters, are used to define a two-dimensional (2D) look-up table. Each cell in the table is filled with the appearance vectors (i.e., oriented energy vectors) for reference images whose derived geometric parameters are covered by the cell's range. During matching, a probe indexes a cell based on its recovered geometric parameters; subsequently, the best match is based on the largest inner product between the appearance vector of the probe and all appearance vectors contained in the indexed cell.

Empirical results: Individual database entries and probes are characterized in terms of quadruples of texture-defined image regions. The appearance of individual patches is captured via their 16D oriented energy vectors. The relative geometry of patches is captured via their 2D affine invariants. The database was constructed from orthoimages corresponding to five different geographical locations. The orthoimages were captured in the visible portion of the spectrum at 8-bits/pixel intensity resolution and at approximately 1 ground meter/pixel spatial resolution. The affine parameters were uniformly quantized into bins of size 0.5, with 1.0 overlap between adjacent bins. These units were chosen based on empirical inspection so as to ensure that geometric-based match pruning would run the gambit between providing almost complete disambiguation of match to no disambiguation of match. Once populated the database consisted of 587 distinct entries (derived from subregions with the five orthoimages). The probes were constructed as synthetic flyovers of regions corresponding to the five orthoimages used to construct the database. For each region, flyovers were generated in two directions (e.g., strips flown north-south or east-west), at three spatial resolutions (where image support and resolution allowed) and two obliquities (nadir and 30° off nadir). Overall, a set of 48 probes resulted.

Overall, when the 48 probes were indexed into the database by the described method, it was found that the top ranked match was the correct match 75% of the time. Further, the correct match was within the top two ranked matches 85% of the time and within the top three ranked matches 95% of the time. Also considered were attempts to index based purely on the basis of appearance (i.e., ignoring the

geometric prefilter); in this case inferior match accuracy was observed, which underlines the importance of joint consideration of appearance and geometry in image indexing.

In this report a method was presented for matching between aerial video and corresponding reference orthoimagery, as typical of geospatial databases. The method combines image appearance, characterized in terms of texture defined regions, and image geometry, characterized in terms of spatial relationships between textured regions. By construction, the matching methods are robust to a range of photometric and geometric distortions between image sources, including changes in grey level contrast and affine geometric transformations. Empirical investigations suggest the promise of the approach. The developed approach can serve as an enabling preprocessing step for extant approaches to video-to-reference image registration. In particular, it can serve to initialize automated image registration algorithms by providing approximate spatial position of the video relative to the reference. An important direction for future research is to subject the method to additional empirical evaluation, especially evaluation that employs real operational data.

Sommaire Les travaux présentés portent sur le repérage de données dans des bases de données géospatiales, particulièrement sur la superposition d'images acquises par différents capteurs présentant différents points de vue d'une même scène. La recherche offre de nombreuses applications pour la création et l'actualisation des cartes, ainsi que pour les levés et autres sources de données géospatiales. Bien que d'importants travaux aient été effectués dans le domaine de la superposition d'images, les méthodes actuelles sont inopérantes lorsque les différences entre des points de vue ou des capteurs sont plus que mineures. En outre, la technologie existante de superposition d'images ne permet l'alignement d'images vidéo et d'images de référence qu'avec un préalignement à quelques centaines de pixels près de la correspondance exacte. Il est significatif qu'en raison d'erreurs, d'omissions et de la disponibilité limitée de mesures de télémétrie, une connaissance a priori inadéquate de l'alignement d'images vidéo et d'images de référence constitue un problème concret.

Nous avons mis au point une méthode semi-automatique d'indexation directe de descripteurs d'images vidéo dans une base de données d'images de référence. La réussite de l'indexation entraîne que la position approximative de l'image vidéo par rapport à l'image de référence a été récupérée; cette information peut ensuite servir à préaligner les images vidéo aux images de référence pour les superposer avec les outils technologiques. Notre méthode offre une représentation uniforme des images vidéo et des images de référence ainsi qu'une méthode permettant de quantifier le degré de correspondance entre les deux échantillons d'images. La représentation combine l'apparence des images, caractérisée par des régions définies selon leur texture, et la géométrie des images, caractérisée par les relations entre les régions texturées. De par sa conception, la méthode de mise en correspondance est robuste pour toute une gamme de distorsions photométriques et déformations géométriques

entre les images, notamment les variations de contraste des tons de gris et les transformations géométriques affines. La méthode élaborée est définie par des algorithmes sur lesquels est basé un logiciel. La seule intervention humaine requise est la sélection de régions d'intérêt sur les images pour focaliser le traitement. Des évaluations empiriques, utilisant une base de données d'images de référence formée de sept orthoimages indexées par des images vidéo aériennes réelles et synthétiques, ont permis de documenter le potentiel de la méthode.

Approche technique: Notre approche en matière d'indexation d'images vidéo d'après des images de référence fait intervenir une intégration innovatrice de descripteurs régionaux d'apparence aux relations géométriques entre les régions. Les recherches antérieures sur l'indexation des images tendaient à se concentrer uniquement sur l'apparence ou sur la géométrie. Notre approche repose sur le principe que la combinaison de ces deux sources d'information produit une méthode plus discriminative que l'utilisation d'une seule de ces sources. L'apparence des images est caractérisée à l'aide de descripteurs de texture. Nous utilisons la texture puisqu'elle est évidente sur toute source d'imagerie (à la différence de la couleur qui dépend de la saisie de multiples bandes spectrales). Les relations géométriques sont définies par un invariant affiné qui offre la robustesse nécessaire pour une gamme de variations prévues lorsqu'on fait coïncider les images visées: les images vidéo aériennes et les images de référence.

Afin de combiner la géométrie et l'apparence, nous utilisons la géométrie comme préfiltre pour l'appariement basé sur l'apparence. Nous utilisons notamment une quantification approximative des paramètres géométriques pour définir un tableau de recherche à deux dimensions. Chaque case du tableau contient les vecteurs d'apparence (soit les vecteurs d'énergie orientés) des images de référence dont les paramètres géométriques dérivés sont couverts par l'étendue de la case. Pendant l'appariement, un échantillon sert à indexer une case d'après ses paramètres géométriques récupérés; par la suite, la meilleure correspondance est considérée être celle pour laquelle le produit scalaire entre le vecteur d'apparence de l'échantillon et tous les vecteurs d'apparence contenus dans la case indexée est le plus élevé. L'appariement des vecteurs d'apparence candidats assure l'alignement des énergies dominantes du point de vue de l'échelle et de la rotation.

Résultats empiriques: Les entrées de la base de données et les échantillons individuels sont caractérisés par des quadruplets de régions d'images définies selon la texture, choisies manuellement. L'apparence de chacune des zones est saisie au moyen de leurs vecteurs d'énergie orientés à seize dimensions. La géométrie relative des zones est saisie à l'aide d'invariants affins à deux dimensions. Une fois constituée, notre base de données comprenait 7×12 entrées (c.-à-d. que chacune des sept régions géographiques comportait douze entrées redondantes). Pour chacun des cinq échantillons, nous avons effectué l'indexation dans le tableau de recherche fondé sur la géométrie en s'appuyant sur les paramètres affins récupérés. Pour chaque entrée dans la case indexée, nous avons établi la correspondance par rapport à tous les vecteurs d'apparence présents en conformité avec la norme sur les produits scalaires. Les dimensions de chacun des vecteurs d'apparence sont de 16×4 (seize énergies orientées, quatre régions de texture par entrée).

L'emplacement géographique correspondant au plus grand produit scalaire calculé constituait la correspondance finale.

Il faut remarquer que nous avons également examiné les résultats d'une mise en correspondance uniquement fondée sur l'apparence (c.-à-d. en ignorant le préfiltre géométrique) et que nous avons trouvé des erreurs de correspondance dans ce cas. En général, nous obtenons de bons résultats uniquement lorsque l'apparence et la géométrie sont combinées.

Dans notre rapport, nous avons présenté une méthode permettant de mettre en correspondance des images vidéo aériennes et l'orthoimagerie de référence correspondante, que contiennent habituellement les bases de données géospaciales. La méthode combine l'apparence des images, caractérisée par des régions définies selon la texture, et la géométrie des images, caractérisée par les relations entre les régions texturées. De par leur conception, les méthodes de mise en correspondance sont robustes pour toute une gamme de déformations photométriques et géométriques d'une image à l'autre, notamment les variations de contraste des tons de gris et les transformations géométriques affines. Des évaluations empiriques donnent une indication du potentiel de la méthode. La méthode élaborée peut servir à la mise en œuvre d'une étape de prétraitement pour des méthodes existantes de superposition d'images vidéo d'après des images de référence. De façon particulière, elle peut servir à initialiser des algorithmes pour la superposition automatisée d'images en fournissant une position approximative des images vidéo par rapport aux images de référence. Diverses avenues sont intéressantes pour de futures recherches. Premièrement, il est souhaitable d'automatiser davantage tous les aspects du traitement. À cet égard, il est particulièrement souhaitable d'éliminer l'obligation de sélectionner manuellement des régions d'intérêt des images, tant pour l'élaboration de la base de données d'images que pour la spécification des échantillons vidéo. Deuxièmement, dans cet ordre d'idées, il est important de soumettre la méthode et tous les travaux ultérieurs à des évaluations empiriques supplémentaires, et spécifiquement à une évaluation utilisant des données opérationnelles réelles.

Aerial video and geospatial database imagery are mutually complementary sources of information about the environment. Aerial video can provide up-to-date imagery of the environment with customized viewing parameters; however, it typically lacks in metadata (e.g., geodetic coordinates, labeling of landmarks and other features of interest). Geospatial database imagery typically is associated with abundant metadata; however, the images may not be current. If these two sources of imagery can be accurately and precisely registered, then their relative merits can be combined to good advantage (e.g., with video serving to update the database imagery, while video inherits the database metadata). While considerable work has been performed in image registration (Brown 1992), extant approaches break down as viewpoint and/or sensor vary beyond relatively small changes. Further, extant image-based registration technology has only been demonstrated to support video-to-reference image alignment when initialized to within several hundred pixels of the correct result (Wildes et al. 2001). Due to errors, drop outs and otherwise limited

availability of telemetry, ineffective a priori knowledge of relative video/reference image alignment is a real-world problem.

In this paper, an approach is presented that allows video-based image descriptors to index directly into a reference image database. Successful indexing into the database implies that at least approximate position of the video relative to the reference has been recovered; this information can then serve to initialize extant technology for precision video-to-reference registration. The approach provides uniform representation of video and reference imagery and a corresponding method for quantifying the goodness of match between two image samples. The representation combines image appearance, characterized in terms of texture defined regions, and image geometry, characterized in terms of spatial relationships between textured regions. By construction, the matching method is robust to a range of photometric and geometric distortions between image sources, including changes in grey-level contrast and affine geometric transformations. The developed approach has been algorithmically specified, instantiated in software and empirically evaluated.

A great deal of research is concerned with image-based indexing and object recognition (see Smeulders et al. 2004, or Snoek and Worring 2005 for recent reviews). The majority of such approaches concentrate on either purely appearance or geometry based methods. Appearance-based methods rely on global statistics of simple image properties (e.g., colour or texture). Such approaches can distinguish between differently coloured or textured objects without reliance on precise feature localization. However, lack of sensitivity to feature location also is a weakness: Objects that differ based primarily on the spatial distribution of components will be confused. Purely geometry-based methods endeavour to characterize an image (or objects that are depicted within an image) in terms of geometric relationships between the location of extracted features, without any explicit regard for feature appearance (beyond their image position). In practice, a major limitation of purely geometry-based indexing schemes is due to the inevitability of localization errors in feature extraction: Such errors lead to corresponding indexing errors. Previous research also has considered combination of appearance and geometry. In one of the more successful recent approaches to image-based object recognition, local descriptors of features (in terms of characteristic scale and orientation) are grouped into clusters to recognize objects and estimate their pose (Lowe 2004). Appearance and geometry also have been combined implicitly through iconic templates, practical implementations of such approaches have appealed to hierarchical, coarse-to-fine matching strategies (Burt 1988).

In the light of previous research, the approach that is presented in the current report makes use of combined information derived from image appearance and geometry to represent and match aerial video to a reference image. As with other recent approaches that have opted for such combination, the motivation comes from the fact that geometry and appearance are complementary in nature. To capture image appearance, standard methods for characterizing image texture in terms of its multiscale orientation structure are employed (Forsyth and Ponce 2003). To capture geometric relations between texture defined regions, previous work in affine invariant recognition is employed (Lamdan et al. 1990). To exploit the complementarity,

the geometric descriptors are used as an initial filter, which is then followed by appearance to complete matching between video and reference. It appears that the developed approach is novel in its particular method for combining appearance with geometry and for its application in video to reference matching.

12.1 Technical Approach

The approach to indexing video to reference imagery involves an integration of regional appearance descriptors and geometric relationships between regions. The philosophy of the proposed approach is that the combination of these two sources of information will yield a more discriminative approach than either one alone. Image appearance is characterized in terms of texture descriptors. Texture is employed as it is manifest in any imagery source (e.g., unlike colour, which depends on capture of multiple spectral bands). Geometric relationships are captured in terms of an affine invariant, which provides robustness to the range of geometric variation that is anticipated in matching imagery of concern, i.e., aerial video to orthophoto reference imagery.

12.1.1 Appearance

When observed through the field of view of any single image in a video sequence, appearance can be ambiguous due to lack of context. For example, a single frame that is completely filled by uniform forest canopy would be hard to discern from another such frame. The proposed solution to this challenge is to consider not individual video frames in isolation, but rather to mosaic collections of frames across a video into a single larger format image and use the mosaiced image as the basis for subsequent indexing to reference. In the current system, a commercially available image mosaicing software tool is used to perform the necessary processing.

For texture-based appearance descriptors, standard methods in computer vision are employed to capture image structure in terms of statistics of energy across multiple scales and orientations. In particular, given a region of interest, the image data is filtered using 2D steerable, separable filters based on the second derivatives of a Gaussian, G_2 , and its corresponding Hilbert transform, H_2 , with responses pointwise rectified (squared) and summed (Freeman and Adelson 1991). Filtering is performed across θ orientations and σ scales using a Gaussian pyramid formulation (Jahne 1991). Hence, a measure of local energy, E , is computed according to

$$E(\mathbf{x}; \theta, \sigma) = [G_2(\theta, \sigma) * I(\mathbf{x})]^2 + [H_2(\theta, \sigma) * I(\mathbf{x})]^2,$$

where $\mathbf{x} = (x, y)$ corresponds to image coordinates, I is the image and $*$ denotes convolution. Further, to increase robustness of the energy representation to changes of contrast between image sources, the filter responses are normalized as follows.

$$\hat{E}(\mathbf{x}; \theta_k, \sigma_l) = \frac{E(\mathbf{x}, \theta_k, \sigma_l)}{\sum_{i,j} E(\mathbf{x}; \theta_i, \sigma_j)},$$

with i taken across all orientations and j taken across all scales. The normalization procedure provides for invariance to local image contrast, as local energy in each band is given as a percentage of total energy across all bands, rather than as an absolute measure. The result yields a set of normalized energy images that parse the original image data according to scale and orientation with robustness to particulars of local image photometry. In the current implementation, 4 orientations and 4 scales are used in filtering; thus, each point is characterized by a 16D appearance vector.

The derived energy images capture local image appearance, i.e., they provide pointwise characterization. As texture is a regional property of an image, the pointwise measurements are grouped to provide a spatial segmentation of the imagery into regions of uniform texture. To perform the grouping, a modification of an extant approach is employed (Felzenszwalb and Huttenlocher 1998). This approach is based on segmenting images into non-intersecting regions, such that the difference between the appearances of points within a segment is minimized, while the difference between the appearances of points between segments is maximized. While the original method made use of simple grey-level images and intensity differences to calculate variation, the current approach makes use of the derived energy images and the inner product to compare local texture descriptors. These modifications are motivated by the present concern with texture-based appearance descriptors.

Finally, given the multiscale, multiorientation energy images that have been recovered for a region of interest, information is accumulated across the entire region by computing the mean response for each scale and orientation. This yields an aggregated 16D vector (4 scales \times 4 orientations) that serves to represent the appearance of the region. This same representation is used to characterize image appearance of regions in a video of interest as well as reference images. Figure 12.1 provides an overall summary of the approach to appearance representation.

12.1.2 Geometry

To capture geometric relationships between image texture-defined regions, it is desirable to choose a representation that is robust to the anticipated changes of view between the database imagery and the video to be indexed. For present purposes, affine invariant coordinates prove useful. Analytically, affine transformations capture the motion of a plane under orthography (Klein 1939). In practice, affine transformations of an image have been widely used in computer vision to compensate for viewpoint changes in cases where the relief of an object is small compared to the viewing distance, such a model is especially popular in the analysis of aerial imagery.

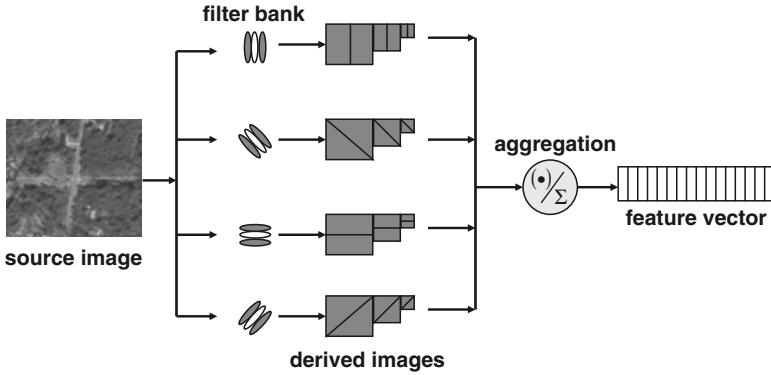


Fig. 12.1 Image appearance characterized via multiscale oriented energy image representation. Input imagery is passed through a filter bank to produce a set of derived images that decompose appearance according to orientation and scale. The filter response in the derived images is aggregated across segmented regions to produce an appearance vector for each region

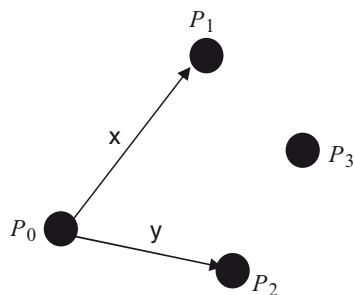
Matters can be formalized with reference to a set of four image regions and their centroids. The centroids are obtained by constructing a convex hull from the segments (O'Rourke 1994), this is done in order to be robust to minor variations in segment shape. Three of the region centroids are used to establish a 2D basis; the coordinates of the fourth centroid are projected on the basis to achieve affine invariance. With reference to Fig. 12.2, let the image coordinates of the four centroids be \mathbf{P}_0 , \mathbf{P}_1 , \mathbf{P}_2 and \mathbf{P}_3 , then the affine representation of \mathbf{P}_3 with respect to the basis established by the remaining points is given by the order pair, α , β , defined as

$$\mathbf{P}_3 = \mathbf{P}_0 + \alpha(\mathbf{P}_1 - \mathbf{P}_0) + \beta(\mathbf{P}_2 - \mathbf{P}_0) = \mathbf{P}_0 + \alpha\mathbf{X} + \beta\mathbf{Y}$$

where $\mathbf{X} = \mathbf{P}_1 - \mathbf{P}_0$ and $\mathbf{Y} = \mathbf{P}_2 - \mathbf{P}_0$. Notice that given a basis set defined in terms of three regions, it is possible to represent an arbitrary number of additional regions in terms of the established basis set. In the current work, a single additional region is employed to investigate the discriminatory power of a small set of regions.

Working with a set of four image regions leads to 12 different ways to specify a basis set for specifying the affine invariant coordinates of the remaining region (i.e., 12 ways to choose 3 distinct items from a set of 4). Following previous work (Lamdan et al. 1990), the reference image database is built with redundancy, so that for each model (i.e., reference image represented by four texture patches) 12 database entries are created, one for each choice of basis points. Subsequently, given a video probe characterized by four texture regions, any three regions can be selected to define the basis for the pair, α , β , that is to be compared to the database: All possible choices are represented in the database. In practice, the set of three centroid coordinates that most nearly yields an orthogonal frame is selected as it will be relatively resilient to centroid localization errors.

Fig. 12.2 Affine invariant representation of 2D geometry. Given four points, three can be used to define an affine invariant basis used to reference the fourth



12.1.3 Combining Appearance and Geometry

To combine geometry and appearance for indexing, three general approaches might be considered. First, geometry could be used as a prefilter for appearance. This method is found to support significant initial pruning of database entries and is the method employed. Second, appearance could be used as a prefilter for geometry. In preliminary investigations, it was found that this method yielded poor ability to provide significant initial pruning of the database. Third, appearance and geometry could be combined into a single high dimensional index. This approach will not allow for initial pruning and may therefore be less efficient than using geometry as a prefilter; however, it will avoid premature pruning of valid candidates and may be considered in the future.

Geometry is used as a prefilter for appearance as follows. Given a 4-tuple of regions (in a video or reference image) a two component descriptor is constructed. The first component consists of the geometric parameters, α , β , coarsely quantized. Coarse quantization is important as it provides robustness to imprecision in appearance region localization; as noted in Sect. 12.1, feature localization is a known fragility of purely feature-based approaches. The second component consists of the appearance vectors of all four regions concatenated and normalized to yield an overall $4 \times 16 = 64$ dimensional vector. During indexing, a two stage process unfolds. First, initial pruning of database candidates occurs by limiting further consideration to only those entries that have similar geometry. Second, the final choice is made by selecting the most similar appearance, but only amongst those entries with already selected geometry.

To facilitate two stage indexing, the reference image database is correspondingly organized. The coarsely quantized geometric parameters, α , β , are used to define a 2D look-up table. Each cell in the table is filled with the 64D appearance vectors for all reference image 4-tuples whose derived geometric parameters are covered by the cell's range. Owing to the geometric quantization, more than one database entry will populate each cell in the lookup table; however,

far fewer entries will be in any one cell than in the entire database and within cell appearance-based selection makes the final choice. The reference image is segmented according to texture-defined appearance regions. Typically, an image will yield more than four segments; each set of four adjacent segments are used to generate an independent database entry; thus, localization within a reference image is supported.

Given a video to be indexed, its two component descriptor is constructed. The geometric component is used to match to the database cell that covers the descriptor's affine parameters, α , β . Subsequently, the final match is selected by comparing the video's appearance component to all database appearance vectors within the initially indexed cell and thereby choosing the most similar. Appearance similarity is quantified by calculating the inner product between the video and database appearance vectors (i.e., given that the vectors have been normalized, by calculating the cosine between their directions); values closer to unity indicate greater similarity. Typically, a video query generates more than four segments. This leads to multiple probes into the database per video query, i.e., all sets of four adjacent segments yield a probe. To make the final selection, the consensus across probes is taken. Such redundancy in indexing is desirable as it provides robustness to variability in the generation of individual probes (e.g., due to uncompensated variability between the video and database imagery).

As defined so far, appearance matching is sensitive to rotation, as oriented texture appears differently across in plane rotations. Robustness to rotation between the videos and the database entries is important as it corresponds to view changes between the sensor platforms. The solution to this challenge lies in the fact that rotation space is cyclic, i.e., rotating the image by returns to the original. This fact is exploited to make matching robust to rotation by aligning two representations in rotation space. In particular, for two texture vectors to be compared one is systematically rotated with respect to the other and for each relative orientation the inner product is calculated. The rotation that achieves the largest inner product is taken as the aligning transformation.

Another important challenge arises with respect to scale invariance. Multiscale texture appears different under variable spatial resolution. The employed approach to scale invariance is similar to the approach to rotation invariance, except for the fact that scale is not cyclic, i.e., systematic scale change never returns to the starting scale. Still, an approach can be realized based on consideration of matches that are shifted in scale. In particular, given two texture vectors to be compared, one is systematically shifted with respect to the other in scale and the inner product is calculated. The shift that achieves the largest inner product is taken as the aligning transformation.

Significantly, no adjustments to the geometric component of matching are required to deal with changes of rotation and scale between the database and video imagery: The affine descriptor is invariant to these attributes by definition.

In summary of the overall approach, Fig. 12.3 provides an example database/video query pair with corresponding regions highlighted.

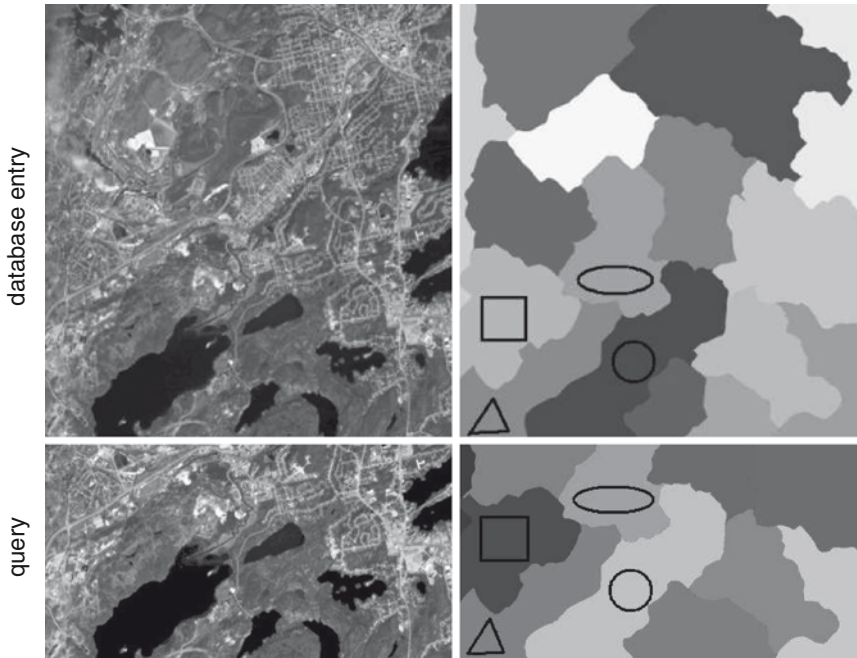


Fig. 12.3 Database/video query pair. Database reference imagery and video mosaic are shown *top-to-bottom*. The *right panels* show segmented appearance regions; overlaid symbols (*circle, ellipse, square, triangle*) highlight corresponding database entry and query probe 4-tuples

12.2 Empirical Evaluation

This section describes an experiment that evaluates the effectiveness of the proposed approach for matching video probes to a reference image database. The database was constructed from orthoimages corresponding to five different geographical locations. All locations include mixed urban and natural terrain. The orthoimages were captured in the visible portion of the spectrum at 8-bits/pixel intensity resolution and at approximately one ground meter/pixel spatial resolution. The affine parameters were uniformly quantized into bins of size 0.5, with 1.0 overlap, so that geometric-based pruning would run the gambit between providing almost complete disambiguation of match to no disambiguation of match. Once populated the database consisted of 587 distinct entries (derived from subregions with the five orthoimages).

The video queries were constructed as synthetic flyovers of regions corresponding to the five orthoimages used to construct the database. For each region, flyovers were generated in two directions (e.g., strips flown north–south, 0, or east–west, 90), at three spatial resolutions relative to the database, (half-resolution, 0.5, same resolution, 1.0 and double resolution, 2.0, as allowed by available imagery) and two

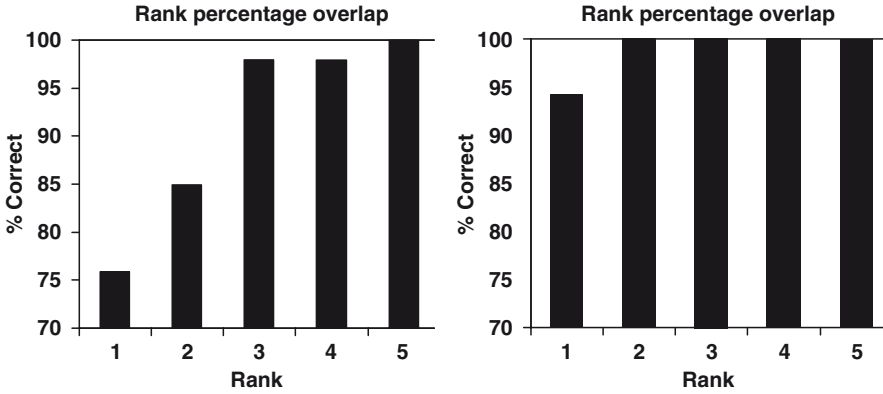


Fig. 12.4 Rank order indexing performance. Overall (*left*), nadir obliquity, 90, only (*right*)

Table 12.1 Algorithm performance as functions of rotation, scale and obliquity

Rank	Rotation		Scale			Obliquity	
	0	90	0.5	1.0	2.0	90	60
1	80%	80%	75%	80%	67%	94%	59%
2	80%	90%	88%	85%	84%	100%	71%
3	100%	90%	100%	95%	100%	100%	94%
4	100%	90%	100%	95%	100%	100%	94%
5	100%	100%	100%	100%	100%	100%	100%

obliquities (nadir, 90, and 30° off nadir, 60). Overall, a set of 34 queries resulted. Example database and video imagery are given in Fig. 12.3.

Overall performance is shown in Fig. 12.4. Results are plotted in terms of rank order, i.e., the percentage of time the correct match was within the top match (rank 1), the top two matches (rank 2), the top three matches (rank 3), etc. A breakdown of performance by major viewing parameters is given in Table 12.1.

As expected, performance is robust to rotation and shows graceful degradation with respect to scale. In contrast, performance falls rapidly with respect to obliquity. Indeed, a plot of rank order performance neglecting off-nadir viewing shows a marked improvement (Fig. 12.4): All correct matches are within rank 2, with 94% rank 1. Inspection of the oblique cases showed that video mosaicing is unreliable and the poorly constructed mosaics severely limited the ability of subsequent processing stages to recover representations from the video that match well to the database imagery. These results suggest that while the current implementation is suitable for indexing of nadir video, improvements to mosaicing are required for strong performance with oblique viewing.

Additional analysis of the results as a function of image appearance shows that locations with most varied texture, and hence largest number of segments yielded best performance. These results support the efficacy of our redundant representation,

as additional segments yield additional probes into the database and correspondingly increased indexing accuracy. Finally, also considered were attempts to index based purely on the basis of appearance (i.e., ignoring the geometric prefilter); in this case inferior match accuracy was observed, which underlines the importance of joint consideration of appearance and geometry in image indexing.

12.3 Summary

An approach has been presented for matching between aerial video and corresponding reference orthoimagery, as typical of geospatial databases. The method combines image appearance, characterized in terms of texture defined regions, and image geometry, characterized in terms of spatial relationships between textured regions. By construction, the matching methods are robust to a range of photometric and geometric distortions between image sources, including changes in grey level contrast and affine geometric transformations. Empirical investigations suggest the promise of the approach. The developed approach can serve as enabling pre-processing for extant approaches to video-to-reference image registration. In particular, it can serve to initialize automated image registration by providing approximate position of the video in the reference.



Chapter 13

AVS LIDAR for Detecting Obstacles Inside Aerosol

Xiang Zhu, Philip Church, and M. Labrie

Abstract The ability to detect obstacles or ground inside aerosols that are small particles suspended in the atmosphere is a highly sought-after feature for numerous helicopter operations. Helicopter pilots need visual assistance for safe flight and landing during white-out conditions that can be produced by natural weather phenomenon such as fog or a snow storm, or can be formed by circling snow clouds generated by a helicopter's rotorwash. Similarly helicopter pilots need to see obstacles or ground under brown-out conditions, in which dust clouds are generated by a helicopter's rotorwash during landing on a sandy area. This paper describes a specially designed AVS LIDAR for detecting obstacles inside aerosols for helicopter operation. AVS (Augmented Visionic System) is a vision assistant system designed to help helicopter pilots navigate under extremely poor weather conditions. One of the key aspects of AVS is the utilization of active sensors to detect obstacles and provide scenes behind aerosols under poor weather conditions. Since LIDAR can use the gating function of timing discrimination to suppress the effect of scattering from dust and fog, it can generally "see" farther than passive sensors such as human eyes and IR cameras, complimenting passive sensors for AVS. The challenge of using a LIDAR under aerosol conditions is not only the requirement of high laser power for penetrating aerosols, but also the requirement of high detection dynamic range and the suppression of aerosol scattering in front of the LIDAR. Although flash LIDAR or gated laser cameras can use timing-discrimination functions to suppress the signals from aerosols, they generally don't have enough optical power to penetrate aerosol and to cover a reasonably large area. Conventional scanning LIDARs have been used to detect obstacles inside aerosol with limited success. They normally use optics that has a co-aligned launching beam and returning beam before the scanning optics. Consequentially, the returning beam from a target is always imaged back onto the receiver regardless of the range to target. The scattering from aerosol in front of a LIDAR window can saturate the receiver before any

X. Zhu (✉), P. Church, and M. Labrie
Neptec Design Group Ltd, Ottawa, ON, Canada
e-mail: szhu@neptec.com

further signal processing can occur. Neptec AVS LIDAR uses an autosynchronized optical design, which utilizes a triangulation relationship to control the amount of returning beam accepted by the TOF (time-of-flight) receiver as a function of target range. The design also maintains this property during high-speed optical scanning. As a result, AVS LIDAR can suppress the return signals from nearby aerosol scattering and, at the same time, have a sensitivity and dynamic range to detect obstacles or ground inside aerosol. Neptec has conducted experiments to study the effect of atmospheric aerosols scattering on LIDAR, FLIR and human vision by using a propagation and aerosol evaluation corridor at DRDC Valcartier, Québec, Canada. Neptec has also carried out flight tests of a concept AVS LIDAR on a NRC Bell 412 helicopter around the shore of the Crash Lake located in the north of Ottawa, Canada.

Sommaire La capacité de détecter des obstacles ou le sol à l'intérieur d'aérosols est une caractéristique très recherchée aux fins de nombreuses opérations par hélicoptère. Les pilotes d'hélicoptère ont besoin d'une aide visuelle pour assurer la sécurité en vol et à l'atterrissage en présence de voiles blancs qui peuvent être dus à des phénomènes météorologiques naturels comme le brouillard et les tempêtes de neige, ou à des nuages de neige circulaires produits par le souffle du rotor d'un hélicoptère. De même, les pilotes d'hélicoptère doivent voir les obstacles ou le sol lorsque des nuages de poussière sont soulevés par le souffle du rotor de l'hélicoptère à l'atterrissage sur un sol sablonneux. Le présent document décrit un dispositif lidar AVS spécialement conçu pour détecter des obstacles à l'intérieur d'aérosols aux fins de la manœuvre d'hélicoptères. Le système AVS (système visionique augmenté) est un système d'aide à la vision conçu pour aider les pilotes d'hélicoptère à naviguer dans des conditions météorologiques extrêmement mauvaises. L'un des aspects clés de l'AVS réside dans l'utilisation de capteurs actifs pour détecter des obstacles et «voir à travers» de mauvaises conditions météorologiques. Comme le lidar peut exploiter la fonction de commande de la discrimination temporelle afin de supprimer l'effet de diffusion par la poussière et le brouillard, il peut généralement «voir» plus loin que les capteurs passifs, par exemple l'œil humain et les caméras IR; c'est un bon complément des capteurs passifs pour l'AVS. Les défis que pose l'utilisation du lidar en présence d'aérosols tiennent non seulement à la grande puissance laser nécessaire pour pénétrer les aérosols, mais aussi à la nécessité de disposer d'une haute dynamique de détection et de supprimer la diffusion par les aérosols devant le lidar. Bien que les lidars à éclats ou les caméras à commande laser puissent exploiter les fonctions de discrimination temporelle pour supprimer les signaux provenant des aérosols, leur puissance optique n'est en général pas suffisante pour pénétrer l'aérosol et couvrir une zone de taille raisonnable. Les lidars à balayage classiques ont été utilisés pour détecter des obstacles à l'intérieur d'aérosols avec un succès limité. Ils utilisent normalement un système optique avec faisceau de départ et faisceau de retour co-alignés devant le dispositif optique de balayage. Par conséquent, le faisceau de retour provenant d'une cible est toujours renvoyé au récepteur quelle que soit la distance de la cible. La diffusion par l'aérosol devant une fenêtre lidar peut saturer le récepteur et interrompre ainsi

le traitement des signaux. Le lidar AVS de Neptec offre une conception optique autosynchronisée, utilisant une relation de triangulation pour fractionner le faisceau de retour qui sera accepté par le récepteur TOF (durée de trajet) en fonction de la distance de la cible. En outre, cette propriété se maintient lors du balayage optique à haute vitesse. Ainsi, le lidar AVS peut supprimer les signaux de retour dus à la diffusion proche par l'aérosol et, en même temps, il possède la sensibilité et la dynamique nécessaires pour détecter des obstacles ou le sol à l'intérieur de l'aérosol. Neptec a mené des expériences visant à étudier l'effet de diffusion des aérosols atmosphériques sur le lidar, le FLIR et la vision humaine en utilisant un corridor de propagation et d'évaluation des aérosols à RDDC Valcartier (Québec, Canada). Neptec a également effectué des essais en vol d'un concept de lidar AVS à bord d'un hélicoptère Bell 412 du CNRC près des rives du lac Crash, au nord d'Ottawa (Canada).

Landing or taking off a helicopter from sandy or snow covered sites typically produces a dense cloud that is called "brownout" or "whiteout". Under these conditions, pilots cannot see nearby objects that provide the visual references necessary to control the aircraft near the ground. There are many reported accidents due to the brownout from both civilian and military aviation. In recent years the search for a solution has become a high priority for US military because of the global war on terror that requires the US military to conduct extensive helicopter operation in desert regions. There have been many efforts for years to develop technologies to help pilots see through dust during landing, notably the "Sandblaster" program from DARPA (the Defense Advanced Research Projects Agency) that aims to test a new landing system that promises safer flying in brownout for military helicopters in September 2008. Despite of the long-term efforts by both military and civilian, no effective solution has emerged.

AVS (Augmented Visionics System) is a project developed by CAE, which is designed to help pilots to navigate under adverse weather conditions including brownout and whiteout. The AVS platform consists of passive sensors, synthetic scene and symbology and the AVS LIDAR that provides active sensor function. The AVS LIDAR is a scanning LIDAR that is being developed at Neptec and is the focus of this paper. It is generally agreed that active sensors have to be used to see through aerosols because their ability of using certain features to discriminate aerosol response from target response. Besides LIDAR, other active sensors, e.g. MMW (millimeter-wave) radar, flash LIDAR and range-gated camera have been studied for the same purpose. Although all of them use the timing discrimination as the way to suppress aerosol signals, there are some significant differences in performance. MMW radar can penetrate deep inside aerosol clouds because of its longer wavelength, but it has poor spatial resolution; in addition, it is difficult to have scanning MMW radar in a compact package. The flash LIDAR and the range-gated camera image the full FOV (field of view) in one shot, they have high resolution and frame rate, but they lack the ability to penetrate aerosol because the light sources in these devices have to be spread into the full FOV for each shot. A scanning LIDAR offers an optimized trade-off between aerosol penetration and resolution.

The Neptec AVS LIDAR is specially designed for aerosol penetration with its unique optics and electronics. In this paper Sect. 13.2 discusses the LIDAR response to aerosols, Sect. 13.3 describes the concept of AVS LIDAR and Sect. 13.4 presents the aerosol penetration experiment and flight test of a prototype AVS LIDAR.

13.1 Aerosol Effect on LIDAR

Generally clouds of any small particles suspended in the atmosphere are called aerosols. They can be generated by the nature ranging from fog, cloud, snow, rain, snowstorm, sandstorm and etc., or man-made clouds such as industrial pollutions and smokes from burning biomass. Since aerosol plays an important role in the environment, it has been investigated by both ground based LIDAR and space based LIDAR. The fundamentals of LIDAR response to aerosol have been studied intensively and the principle is well established. However, the range of aerosol is relatively far (~1 km) and its density is small in the studies done for pollution monitoring and climate research. The aerosol generated by helicopters is dense and in short range from 0 to 100 m for helicopters; therefore it poses new requirements for LIDAR design and operation.

In a dense aerosol, the scattering is a multiple scattering process and a rigorous analysis of LIDAR return needs Monte Carlo simulation (Bissonnette 1995). For the purpose of discussing AVS LIDAR design in this paper, a single scattering process is considered for the aerosol scattering under the assumption that the light received by the LIDAR is from a single back scattering with a 180° scattering angle and light from all other scattering doesn't enter the LIDAR receiving optics and is lost. Under this assumption, the LIDAR return signal is given by:

$$P(R) = P_0 \times \frac{c\tau}{2} \times \beta(R) \times \exp[-2 \int_0^R \alpha(R') dR'] \times K(R) \times \frac{A}{R^2}$$

Where $P(R)$ is the LIDAR return from scattering at range R , P_0 is the LIDAR launch pulse power, c is the light velocity, τ is the LIDAR timing resolution, $\beta(R)$ is the back scattering coefficient, $\alpha(R)$ is the absorption coefficient, $K(R)$ is the overlap function, A is the receiving lens area and R is the range.

For a uniform aerosol cloud, $\beta(R)$ and $\alpha(R)$ are constants and independent of range. $P(R)$ is shown to be proportional to $\exp(-2\alpha R)/R^2$, it decreases very fast when range R increases. The equation shows that even for a moderate $\alpha = 0.05/m$, $P(R)$ decreases more than eight orders of magnitude from 1 to 100 m range, while the return signal decreases only four orders of magnitude at 100 m without aerosol attenuation. This has two implications for the LIDAR design: first, the initial pulse power P_0 has to be high in order to penetrate the aerosol, reach a target and has some reflected light received by the receiving optics. Second, if a LIDAR is designed to have a sensitivity to detect weak return signals from far, the scattering

from nearby aerosol will saturate the system if it does not have an extreme high dynamic range (e.g. $>10^8$).

Current technologies of high power pulse laser and detection electronics can deliver high pulse power and detect weak return signal, e.g. lasers with peak power of 10 kW at 10 kHz PRF (pulse repetition frequency) are readily available; and standard APD receivers can have a detection limit of 10 nW, the ratio of launch power to return power is 1,012 ($=10 \text{ kW}/10 \text{ nW}$). But the dynamic range of a detection electronics is normally less than 105 (50 dB). Obviously, LIDAR designed for imaging through aerosols has to use some method to suppress the return from aerosol, especially the scattering signal from nearby aerosol. AVS LIDAR uses a specially designed optics to address this requirement that will be discussed in the next section.

13.2 Concept of AVS LIDAR Design

LIDAR optics can be divided into two groups: monostatic optics and bistatic optics. In a monostatic LIDAR optical design, the launch beam and the return beam are co-aligned, for example, a launch laser beam can be introduced into a system by a small mirror in the center of a receiving telescope in most LIDAR for aerosol research. In this configuration, the scattering from the nearby aerosols can have a large signal return and saturates the detection system as described in the last section. Practically, a design using telescope optics is not suitable for high speed scanning. Monostatic design in scanning LIDAR is normally done by using a parabolic mirror with a hole in the center that allows the launch beam out as shown in Fig. 13.1, in which the detector will always track the laser spot during the scanning.

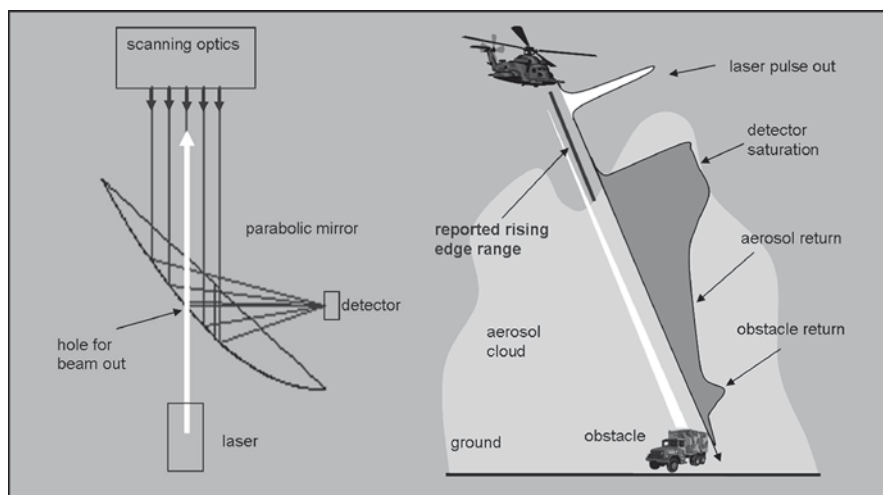


Fig. 13.1 Scanning monostatic LIDAR optics and LIDAR response to aerosol and obstacle

Figure 13.1 also shows the response of a conventional LIDAR to aerosols and obstacles, in which the receiver of the LIDAR is saturated by the nearby aerosol scattering. Since a conventional LIDAR is triggered by the rising edge of a return pulse and there is no separated pulse from the target that is buried inside the aerosol, a conventional LIDAR can only report the range of the closest aerosol under the brownout situation.

In a bistatic LIDAR optical design, the launch beam and the return beam are separated in space by two different lenses side by side. There is a triangulation relation between the spot where the light is scattered and its imaging spot on the receiving focal plane. On the focal plane of the receiving lens, the return light from a nearby object is focused off the lens optical axis; the return light from a far distance is focused on the optical axis. If a detector is aligned to receive light returned from a far distance, the returned light from the nearby aerosol scattering will not hit the detector directly and therefore is detected with a reduced sensitivity.

The standard way to acquire a 3D data set using a bistatic LIDAR is to provide external rotation (or translation) by installing the static optical assembly on a pan-tilt unit. Using external methods to move the exit beam along a path, rather than using internal swapping mirrors in the optical assembly, results in a very slow scan speed. Even though this triangulation optical property of bistatic LIDAR optics is well known, it has not been used for high-speed scanning application so far.

The AVS LIDAR uses an autosynchronized optical design, shown in Fig. 13.2. The launch beam is scanned by the X-mirror first, and is reflected by a side mirror before it is scanned by the Y-mirror. The return beam passes these components in reverse order. Because the X-mirror is passed by both the launch beam and the return beam, this configuration automatically synchronizes the detector FOV with the laser spot. The triangulation relationship is maintained during high-speed

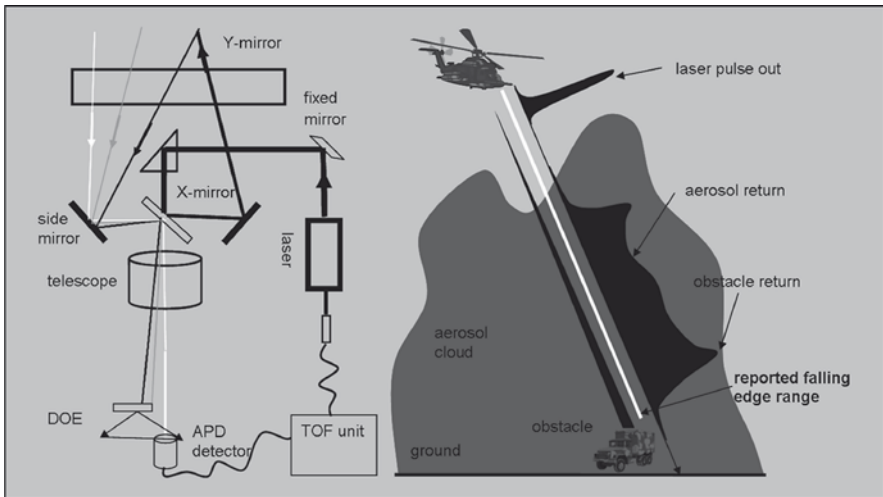


Fig. 13.2 AVS LIDAR optics and AVS LIDAR response to aerosol and obstacle

scanning operations by swapping (moving) the two internal mirrors, and not the entire optical head. Since the locations of return beams follow the triangulation relation based on the scattering range, many arrangements can be used to optically manage the LIDAR dynamic range, i.e. to suppress the return from the nearby aerosol and to detect a target at far distance with great sensitivity.

Figure 13.2 also shows the response of AVS LIDAR to aerosols and targets inside aerosols. The return signal from the nearby aerosol scattering is suppressed optically to prevent the detection system from overloading. AVS LIDAR has a detection mode to use the falling edge of a pulse to report range, in addition to standard rising edge detection. If an obstacle is buried inside an aerosol as shown in Fig. 13.2, the falling detection of AVS LIDAR reports the correct range of the obstacle.

13.3 Results from Aerosol Penetration Study and Flight Test

13.3.1 LIDAR Aerosol Penetration Experiment

To study LIDAR penetration inside aerosol and to compare LIDAR performance with passive sensors including eye, visible camera and IR camera, we conducted a series of experiments by using the aerosol research corridor at DRDC Valcartier in Québec (Fig. 13.3). The corridor has a length of 22 m. It can generate an oil-fog



Fig. 13.3 Aerosol research corridor at DRDC Valcartier, Québec

Table 13.1 Aerosol types generated by the corridor and simulated counterparts

Aerosol particle size (μm)	Aerosol material	Simulated atmospheric environment
0.3	Fog-oil	Fog, clay dust
2	Spherglass solid spheres	Clay dust
6	Spherglass solid spheres	Silt dust
25	Polymeric microspheres	Sand dust
50	Polymeric microspheres	Sand dust
90	Polymeric microspheres	Sand dust, snow flakes

with a particle size of 0.3 μm , glass bead dust with particle sizes of 2–6 μm , and dust using hollow spheres with particle sizes of 25–100 μm .

The optical depth (defined as $-\ln(T)$, T is transmission) of the aerosol clouds generated by the corridor ranges from 0.01 to over 10, but the range that can be monitored by a transmissometer is 0.01–4.

The nature of dust cloud in a brownout generated by helicopters depends on the soil type; generally the dust is classified by its particle size. The size of clay dust is 0.01–1 μm , silt dust is 1–25 μm and sand dust is 25–100 μm (Farmer 2001). Table 13.1 shows the type of aerosols generated by the corridor and their equivalent aerosols in the simulated atmospheric environment.

A trailer hosting a range of instruments (visible camera, IR camera and transmissometer) is located about 100 m away from the corridor. The target for the visible camera is a board with black and white stripes that are made according to the Johnson criteria for object reorganization. The target for IR camera is a frame with heated bars that are spaced also according to the Johnson criteria. Both targets are located behind the corridor. LIDAR with near-infrared wavelength is either located in the trailer or in front of the corridor in the experiment and the LIDAR target is about 30 m behind of the corridor.

During a typical operation, a dense aerosol cloud is generated to a steady state inside the corridor with a known density; both doors of the aerosol corridor are then opened simultaneously within 0.5 s. All sensors record data at the moment the doors are opened, and continue to collect data until the density of the aerosol cloud becomes too thin for any further measurement.

To compare the performance of LIDAR with passive sensors, we define a parameter called “detection factor” that is the ratio of aerosol density at the moment of target detection by the LIDAR or IR camera to the aerosol density at the moment of target detection by eye (visible camera).

Table 13.2 shows results in terms of the detection factor for three typical aerosols. While the IR camera works well in fog, it does not offer significant improvement compared to the eye. It performs poorly for 6 μm aerosol because of strong scattering at its wavelength. The LIDAR can ‘see’ through 50 μm sand dust that is four times denser than that an eye can see through, 6.6 times denser for the 6 μm dust case and more than 7.6 denser for fog. Since LIDAR can penetrate the densest fog generated in the corridor, the upper limit of the detection factor was not determined.

Table 13.2 Detection factor comparison between LIDAR and passive sensors for different aerosols

	Sand dust (50 μm)	Silt dust (6 μm)	Fog(0.3 μm)
Eye (visible camera)	1	1	1
IR camera	1.7	1.1	7.4
LIDAR	4	6.6	>7.6

The scattering effect of aerosols is highly wavelength dependent. The size of fog is close to visible wavelengths, and it scatters visible light strongly. As a consequence, the human eye or visible camera cannot see through fog very well. The IR camera operates in the 3–5 μm wavelength range, which is much larger than the size of a fog droplet. As a result, the IR camera can see through fog much further than the human eye. When silt dust has an average size of 6 μm , which is close to the IR camera operating wavelength, the IR camera performs only as well as the eye. Because LIDAR reduces scattering by using time discrimination, LIDAR does see through further than both the eye and the IR camera under all three aerosols and different particle sizes.

These experiments provide an understanding and comparison of LIDAR and passive sensor performance for simulated brownout and whiteout conditions. They demonstrate that LIDAR can detect obstacles in the environment of a very dense aerosol at a short distance. The results provide information for optimizing the AVS LIDAR design, including techniques of pulse mode (first pulse mode and last pulse mode) and trigger mode (rising edge and falling edge) to suppress the effect of aerosol scattering. The results are obtained for environments typical of helicopter landing, and no attempt is made to extrapolate to other conditions, meaning different ranges and visibilities. AVS LIDAR can be used to detect obstacles around a landing site during brownout or whiteout conditions.

13.3.2 Flight Test of a Prototype of AVS LIDAR

To evaluate the AVS LIDAR under regular or whiteout conditions, a helicopter flight test was carried out around the shore of the Crash Lake located in the north of Ottawa, Ontario, Canada in February, 2007. An AVS LIDAR prototype was mounted under the Bell 412 helicopter from NRC (National Research Council Canada) Flight Research Laboratory. The Bell 412 flew to a landing site: an open field with some bushes and rocks in the center, trees on the one side and a flat frozen lake on the other side. The LIDAR was used to scan the landing site under different helicopter maneuvers; a push-broom scan mode was used in fly-by or approaching operation and a raster scan mode was used in hovering operation.

The result of the pushbroom scan shows the 3D landscape of the fly-by path, and the raster scan shows the tree and obstacles on the landing site. The test shows that the AVS LIDAR design can sustain mechanical vibration during flight. LIDAR data and helicopter INS data can be fused to generate geo-referenced 3D images.

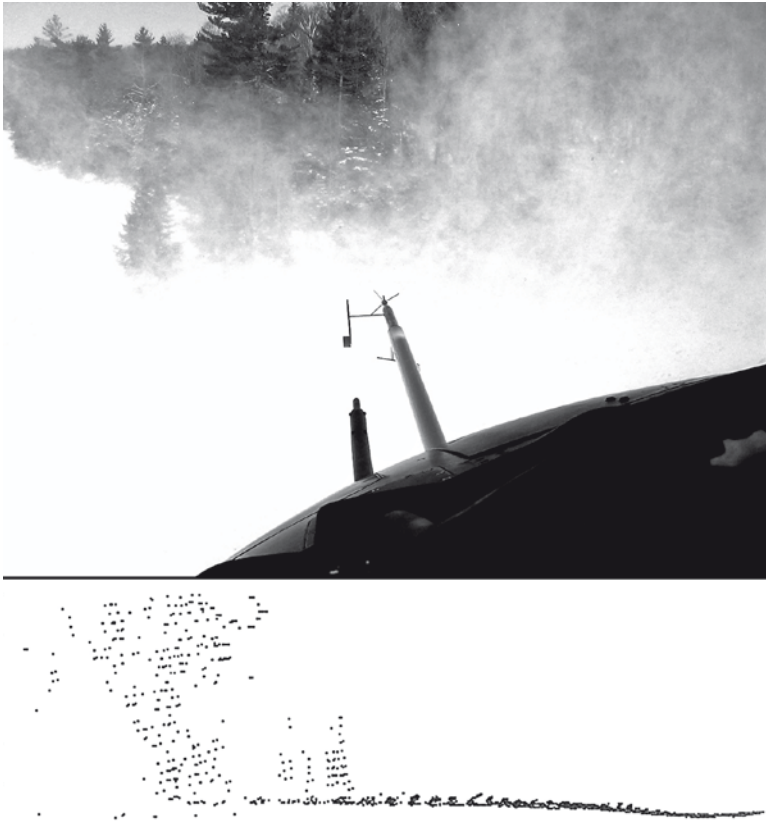


Fig. 13.4 View of the helicopter pilot and a side view of the AVS LIDAR image under whiteout conditions

To test AVS LIDAR under whiteout conditions, the helicopter hovered close to ground to generate blowing snow clouds by rotor wash. A pilot's view of this whiteout is shown in Fig. 13.4, as captured by a visible-range camera. The ground and trees of the landing site cannot be seen through the snow cloud. The bottom of Fig. 13.4 shows a side view of the 3D LIDAR scan taken simultaneously. Because of AVS LIDAR triangulation optics, the AVS LIDAR receiver is not saturated by scatter from the snow cloud near the helicopter; it is able to detect ground and trees behind the snow cloud by the last pulse mode.

13.4 Summary

With specially designed triangulation optics, the AVS LIDAR has a unique ability to detect obstacles under aerosol conditions. Some of its key features were tested in a prototype AVS LIDAR device. A study of LIDAR aerosol penetration was carried

out at the aerosol research corridor at DRDC Valcartier. This showed the effectiveness of the LIDAR as an AVS active sensor for aerosol penetration. The AVS LIDAR prototype was flight-tested on a Bell 412 helicopter. The test showed the usefulness of AVS LIDAR as an active sensor for helicopter operation, and confirmed aerosol penetration under whiteout conditions. Currently Neptec is focusing its design effort on size and weight reduction of this LIDAR, and on falling edge detection, faster sampling rates, and scan speed.

Acknowledgements We would like to thank our collaborators in this project, including Norah Link, Louis Dentigny, and Ron Kruk from CAE for their support and guidance on AVS system requirement and application. We thank Sion Jennings, Carl Swail, and Stephane Carignan from NRC Flight Research Laboratory for their support on system integration and helicopter flight tests; also Luc Bissonnette and Gilles Roy from DRDC Valcartier for their help in studying aerosol penetration.

Author Biographies

Robert Allison is interested in basic and applied aspects of stereoscopic depth perception. His areas of expertise are the psychophysical and computational investigation of depth perception, analysis of eye movements, and perceptual issues in human spatial orientation and virtual reality systems. He is currently an Associate Professor at York University (Toronto Canada) in the department of Computer Science and Engineering with appointments at the Centre for Vision Research and the graduate program in Psychology. He received his B.A.Sc. in Computer Engineering from the University of Waterloo, an M.A.Sc. in Biomedical Engineering from the University of Toronto and his Ph.D. in Biology from York University.

Jeff Clark has been active in the physics-based simulation arena for 13 years. During that time, he served in technical and project management capacities at Accurate Automation Corporation, Hughes Aircraft, Raytheon, and L-3 Communications. Today, Mr. Clark represents Renaissance Sciences Corporation, a private government services firm, where he serves as President and oversees AFRL and NAVAIR 6.3 through 6.5 research programs at the company. Mr. Clark completed a bachelor of science in computer science from Radford University and the general aviation curriculum and private pilot certificate at the University of North Dakota Aerospace. Jeff and his family of four reside in Chandler, Arizona. Mr. Clark has supported, as a contractor, the Night Operations Center of Excellence, Air Force Research Laboratory, Mesa, Arizona continuously since February, 1996 where he managed the software development of the NVTS, SensorHost, and other government technology products that have come out of the AFRL NVG program. More recently Mr. Clark and the RSC team supported the AFRL NVG image generation prototyping activity associated with Phase I of the ADDNS program.

Christine M. Covas is an Associate Research Psychologist in the Visual Systems Research Laboratory at the Air Force Research Laboratory in Mesa, Arizona. Her research focuses on visual perception, particularly as it relates to display technologies for simulation and training applications. She is in pursuit of a Ph.D. in Psychology from Arizona State University and holds a M.S. (2005) and B.S. (2003) in Applied Psychology from Arizona State University Poly.

Don Donderi is a co-founder and Principal Consultant of Human Factors North, Inc., a Toronto-based Human Factors and Ergonomics Consulting Firm. He is also an associate professor in the department of psychology at McGill University, Montreal. While in graduate school during the cold war, he worked at IBM in the United States on the visual perception demands of aerial reconnaissance. After graduating with a doctorate in experimental psychology from Cornell University in 1962, he joined the Psychology Department of McGill University in Montreal. In 1982 he co-founded Human Factors North, a Toronto-based human factors/ergonomics consulting firm. Since then he has worked with private companies and civil and military agencies involved with flight simulation, marine navigation, rail transport, road transport, air traffic control, nuclear power, and marine search and rescue. He has carried out field research on marine navigation and on search and rescue in the Canadian Arctic and in the Atlantic Ocean, in cooperation with the Canadian and US Coast Guards and the Canadian Department of Transport. He has designed wheelhouses for Canadian Coast Guard cutters. He has carried out research work and literature reviews on marine navigation displays for Defence Research and Development – Toronto. In 2004, his work on marine integrated navigation displays was published in *Displays* and *Journal of Navigation*, and in 2006 he wrote a comprehensive review of visual complexity published in *Psychological Bulletin*.

Michel Doucet received a B.Sc. degree in physics in 1988 from Université du Québec à Chicoutimi (UQAC) and a M.Sc. degree in physics (optics) in 1991 from Université Laval, Québec, Canada. He is a researcher at the National Optics Institute (INO) since 1991 where he has contributed to the development of many optical devices, including the optical design of zoom lenses, military IR lenses, and astronomical instruments. He has also been involved in the development of biomedical instruments, many sensors (3D and image) for road inspection and devices based on deformable micro-mirrors. His main research interests are in the area of optical design.

James P. Gaska is a Senior Scientist with Link Simulation and Training working at AFRL in Mesa, Arizona. Dr. Gaska received a Ph.D. degree in Biopsychology from Syracuse University in 1981. From 1981 to 2000, Dr. Gaska used single cell electrophysiological and computational techniques to explore and model the representation of the world in the visual cortices of primates. From 1995 to present, Dr. Gaska has worked as contractor to the Air Force and used vision science and visual performance models to aid in the design of camouflage, laser eye protection and, currently, flight simulation technologies.

Mary K. Kaiser is a research psychologist in the Human Systems Integration Division at NASA Ames Research Center. She received her Ph.D. in psychology from the University of Virginia, and was a postdoctoral fellow at the University of Michigan in applied experimental psychology before joining Ames in 1985. The author of over sixty articles and chapters on perceptual psychology and human factors, Dr. Kaiser also served as an associate editor of the *Journal of Experimental Psychology: Human Perception and Performance* for over a decade, and is currently Project Scientist for NASA's Space Human Factors Engineering project.

Ray M. Obidowski received a B.Sc. Degree with high distinction from the University of Toronto in 1991. He obtained a M.Sc. degree (1993) in Geomatics Engineering from the University of Calgary, as an NSERC scholar. His studies specialized in the area of precise 3D close-range measurement systems and technology. Ray became a Professional Engineer (P.Eng.) in 1997 and a Project Manager Professional (PMP) in 2006. Ray has over four years of simulation and modelling project management experience at Array Systems Computing. He is currently the project manager for a contract in Geographic Terrain Database Tools programming and integration. He is also the technical leader for projects involving support for human factors experiments, flight simulation visualization, and ADDNS Image Generator (IG) integration. Prior to joining Array Systems, Ray managed research, development and testing of a 3D laser radar (lidar) system at Metric-Vision Inc., in Virginia, USA.

Avi Parush's professional career in human factors engineering (HFE), usability engineering, and human-computer interaction (HCI) spans over 20 years. He has been involved in user-interface design and testing in a large variety of domains and projects. Parush is the founder and editor in chief of the Journal of Usability Studies. His academic background is in cognitive-experimental psychology and he is presently a professor of psychology at Carleton University, Ottawa. His main research focus are on spatial cognition, way-finding and orientations theories and systems, simulator-based learning and training, and the application of human factors and usability approaches and principles in the healthcare system with the aim of reducing errors and increasing safety.

Francis Picard received a B.Sc.A. in Physics Engineering and a M.Sc. in Physics from Laval University. For the past 18 years, he has been working at the National Optics Institute (INO) as a researcher working on thin film technology, microbolometer arrays for infrared imaging and Microsystems, including micromirrors for various applications such as projection display and laser beam shaping. He is currently the manager for the optical MEMS program at INO.

Robert Wolfe is a 1986 UBC graduate of Engineering Physics. He received a Master's degree in Electrical Engineering, specializing in image processing, from Laval University in 1996. He has been contracted to Laval University, DRDC Valcartier, DRDC Toronto, and Serenica Technologies Inc. over the last 15 years. In Québec, he was a key member of the lauded DRDC Valcartier PASSPORT project for multi-platform search and rescue software. At DRDC Toronto, Robert extended horizons in visual networked simulation to include virtual reality based ship, helicopter and dismounted infantry, and wider scale international research wargaming such as CMTR and NATO FirstWave exercises. Most recently, he has taken on the role with Serenica of integration engineer for the ADDNS project.

Mr. Vitaly Zholudev received B.Sc. Honours and Masters degrees in computer science from York University in Toronto in 2004 and 2006, respectively. As part of his M.Sc. work, Mr. Zholudev produced a thesis titled, Toward Video to Geospatial Reference Image Indexing. Since 2004, Mr. Zholudev has been a member of the

Computer Vision Laboratory at York University, from 2004 to 2006 as graduate student and from 2006 to present as a full-time research associate. His research interests include computer vision, in particular image registration and indexing.

X. Zhu is the principal optical engineer at Neptec. He received a M.Sc. in physics from Queen's University at Kingston, Canada in 1992 and a Ph.D. in engineering physics from McMaster University at Hamilton, Canada in 1996. He worked as an opto-electronic device designer at Perkin-Elmer and as an instrumentation designer at JDS Uniphase before joining Neptec. He has published numerous papers in *Applied Optics*, *IEEE Photonics Technology Letters*, *Journal of the Optical Society of America* and *SPIE proceedings* in laser technology and applications. He holds a number of patents including Neptec patents on optical design of LMS (laser metrology system) and TriDAR scanner.

Glossary

Above real time training (ARTT) A procedure in which visual-motor tasks are trained at speeds faster than those at which they are normally performed.

Aerosols (1) Clouds of small particles suspended in the atmosphere. (2) A set of solid or liquid particles in suspension in a gaseous medium.

Ambient illumination Light in an environment. More specifically, parasitic light due to the environment. It may come from light passing through curtains or under doors, or from any source of light inside the room, such as computer screens.

Angular frequency (1) The rate of change in the orientation angle during rotation, measured in radians/second. (2) A product of the frequency of a sinusoidal quantity and the factor 2π .

Anti-aliasing An operation that is applied to all pixels in an image after the image is rendered. Anti-aliasing techniques make outlines and edges appear smooth by applying an algorithm that blurs edges by averaging subpixel areas. This improves the fidelity of digital images.

Aperture A hole or an opening through which light is admitted.

Application programming interface (API) is a source code interface that an operating system or library provides to support requests for services to be made of it by computer programs.

Augmented visionics system (AVS) A vision assistant system designed to help helicopter pilots navigate under poor weather conditions.

Back reflection (1) Portion of light not transmitted nor absorbed at a target or more generally, parasitic light due to the environment. For example, light is reflected back to a projection screen from objects within the room. (2) A process in which an incident wave, meeting a surface separating two different media, gives rise to another wave called a reflected wave, which propagates away from the surface in the same medium as the incident wave and is interpretable by geometrical optics.

Bidirectional reflectance distribution function (BRDF) A mathematical representation of a surface's reflectivity properties that is used for luminance computation.

Binocular disparity The difference in images of an object seen by left and right eye resulting from the eyes' horizontal separation. The brain uses binocular disparity to extract depth information from the two-dimensional retinal images in stereopsis.

Brownout Occurs when dust clouds are generated by a helicopter's rotor wash during landing on a sandy area.

Capacitance (1) A measure of the amount of electric charge stored (or separated) for a given electric potential, measured in farads. (2) For a capacitive two-terminal element with terminals A and B, quotient of its electric charge q by the voltage u_{AB} between the terminals: $C = q/u_{AB}$ where u_{AB} is the line integral of the electric field strength from A to B and where the sign of q is determined by taking the electric current in the time integral defining the electric charge as positive if its direction is from A to B and negative in the opposite case.

Close air support Air action by fixed or rotary winged aircraft against hostile targets that are in close proximity to friendly forces, and which requires detailed integration of each air mission with fire and movement of these forces.

Collimate To bring into line or to make parallel. Collimated light is light from an indefinitely distance source, whose rays are approximately parallel.

Commercial off the shelf (COTS) Software or hardware, generally technology or computer products, that are ready-made and available for sale, lease, or license to the general public.

Common image generator interface (CIGI) An interface designed to promote a standard way for a host device to communicate with an image generator (IG) in the real-time simulation industry.

Contrast (1) The range of optical density and tone on a photographic negative or print (or the extent to which adjacent areas on a television screen differ in brightness). (2) In the perceptual sense: Assessment of the difference in appearance of two or more parts of a field seen simultaneously or successively (hence: brightness contrast, lightness contrast, colour contrast, simultaneous contrast, successive contrast, etc.)

Dead reckoning The process of estimating one's current position based upon a previously determined position, or fix, and advancing that position based upon known speed, elapsed time, and course.

Digital light processing (DLP) A projection technology that uses microscopically small mirrors laid out in a matrix on a semiconductor chip to create an image. The trademark is proprietary to Texas Instruments.

Digital to analog converter (DAC) (1) A device for converting a digital (usually binary) code to an analog signal (current, voltage or electric charge).

Digital video interface (DVI) A video interface standard designed to maximize the visual quality of digital display devices such as flat panel LCD computer displays and digital projectors. The DVI interface uses a digital protocol in which the desired illumination of pixels is transmitted as binary data.

Distributed interactive simulation (DIS) An open standard for conducting real-time platform-level wargaming across multiple host computers

Distributed rendering (DR) A method used to improve the performance of computer graphics creation software by sharing computing across more than one compute node.

Electro-optical sensor Converts light rays into an electronic signal.

Emmert's law (1) States that objects that generate retinal images of the same size will look different in physical size if they appear to be located at different distances. Specifically, the perceived size of an object increases as its perceived distance from the observer increases. (2) The apparent increase in the size of a visual afterimage when viewed against a relatively nearby surface and then against a more distant surface. The relation is one of proportionality, the explanation lies in the fact that the visual angle remains constant but the apparent distance increases, which would normally indicate an increase in size (see under visual angle). The law can be demonstrated by staring at a brightly lit image with a steady gaze for a minute and then shifting one's gaze first to a blank sheet of paper and then to a distant wall: the afterimage will appear to grow much larger. [Named after the Swiss ophthalmologist and psychologist Emil Emmert (1844–1911) who formulated it in 1881, although it had been discovered more than two centuries earlier by the Italian physicist Benedetto Castelli (1578–1643)].

Endian The ordering of bytes in a multi-byte computer datum.

Euclidean distance The “ordinary” distance between two points that one would measure with a ruler, which can be proven by repeated application of the Pythagorean theorem.

Flying spot A mechanism used to generate dynamic images from light or light-generating sources. Consists of a beam that is scanned in two orthogonal directions. The intensity of the beam is modulated as the spot moves on a screen.

Frequency A measure of the number of occurrences of a repeating event per unit time, measured per second in Hertz (Hz).

Gain (1) A measure of the ability of a circuit (often an amplifier) to magnify the power or amplitude of a signal. (2) For a linear time-invariant system with a sinusoidal input variable in steady state the ratio of the amplitude of the output variable to the amplitude of the corresponding input variable, represented as a function of the angular frequency ω . (3) An increase between two points of an electric, electromagnetic, or acoustic power. (4) The quantitative expression of a power increase by the ratio of the values at two points of a power or of a quantity related to power in a well defined manner. (5) Of a measuring instrument: The ratio of the output to the input values of quantities of the same kind in a device or system.

Galvanometer A type of ammeter; an instrument for detecting and measuring electric current. It is an analog electromechanical transducer that produces a rotary deflection, through a limited arc, in response to electric current flowing through its coil.

Genlock Occurs when two electrical video signals are synchronized in horizontal and vertical scan.

Glideslope Used in aviation as an instrument to give vertical guidance while landing. Enables pilots to touch down on the runway at the correct point.

GNU A computer operating system composed entirely of free software.

Gonioreflectometer A device for measuring a BRDF

Halo In the context of night vision devices (NVDs), halos are the phenomenon that occurs when a bright light source viewed through NVDs appears to be surrounded by a corona or halo that is much larger than predicted by the point spread function of the device.

Hermetic (1) Made airtight by fusion or sealing. (2) A seal is hermetic when the specified rate of leakage is several orders smaller than that of a barrier seal.

Heuristic A trial-and-error method of problem solving.

High Level Architecture (HLA) A general purpose communication architecture for distributed computer simulation systems. HLA is data format and computing platform independent.

Host-emulator An emulator duplicates (provides an emulation of) the functions of one system using a different system, so that the second system behaves like (and appears to be) the first system

Illuminance The illuminance at a point of a surface is the quotient of the luminous flux incident on an infinitesimal element of the surface containing the point under consideration, by the area of that surface element. (Wyszecki & Stiles, 1982).

Illumination The intensity of light falling at a given place on a lighted surface; the luminous flux incident per unit area, expressed in lumens per unit of area. See also irradiance.

Incandescence (1) The emission of visible light from a substance or object as a result of heating it to a high temperature. The color of the light emitted from solids and liquids is a function of their chemical structure and their temperature. (2) Emission of optical radiation by the process of thermal radiation.

Interferometric microscopy The concept of microscopy which is related to holography, synthetic-aperture imaging, and off-axis-dark-field illumination techniques. Interferometric microscopy allows enhancement of resolution of optical microscopy due to interferometric (holographic) registration of several partial images (amplitude and phase) and the numerical combining.

The inverse square law States that (1) The principle according to which the intensity of a wave, including a light wave or a sound wave, decreases in proportion to the square of the distance from its source. (2) Any point source which spreads its influence equally in all directions without a limit to its range will obey the inverse square law. This comes from strictly geometrical considerations. The intensity of the influence at any given radius r is the source strength divided by the area of the sphere. Being strictly geometric in its origin, the inverse square law applies to diverse phenomena. Point sources of gravitational force, electric field, light, sound or radiation obey the inverse square law.

Irradiance (1) The density of radiation incident on a given surface usually expressed in watts per square centimeter or square meter. (2) The radiant power incident on an element of a surface divided by the area of that element.

Jitter (1) An unwanted variation of one or more signal characteristics in electronics and telecommunications. Jitter may be seen in characteristics such as the interval between successive pulses, or the amplitude, frequency, or phase of successive cycles. Jitter can apply to a number of signal qualities (e.g. amplitude, phase, pulse width or pulse position), and can be quantified in the same terms as all time-varying signals. (2) A sudden, small, irregular departures from the ideal value of a characteristic of a signal, such as the phase, pulse duration or magnitude. (3) A perceptible instability of the time base of a reproduced video signal.

Johnson criteria The minimum resolution according to Johnson's Criteria are: Detection – an object is present: $2 + 1/-0.5$ pixels. Orientation - symmetrical, asymmetric, horizontal or vertical: $2.8 + 0.8/-0.4$ pixels. Recognition – the type object can be discerned, a person vs. a car: $8 + 1.6/-0.4$ pixels. Identification – a specific object can be discerned, a woman vs. a man, the specific car: $12.8 + 3.2/-2.8$ pixels. These measurements give a 50% probability of an observer discriminating an object to the specified level.

Lambertian materials (1) Materials that have a constant BRDF. Light falling on such material is scattered such that the apparent brightness of the surface to an observer is the same regardless of the observer's angle of view. (2) A Lambertian surface is an ideal surface for which the radiation coming from that surface is distributed angularly according to Lambert's cosine law.

Laser based video projectors (LBVP) Takes a video signal and modulates a laser beam in order to project a raster-based image. The systems work either by scanning the entire picture a dot at a time and modulating the laser directly at high frequency, much like the electron beams in a CRT, or by optically spreading and then modulating the laser and scanning a line at a time, the line itself being modulated in much the same way as in a DLP. See also flying spot.

Latin square An $n \times n$ table filled with n different symbols in such a way that each symbol occurs exactly once in each row and exactly once in each column.

LIDAR (light detection and ranging) An optical remote sensing technology that measures properties of scattered light to find range and/or other information of a distant target.

Linear spatial light modulator (SLM) A mechanism used to generate dynamic images from laser sources. When illuminated with a thin laser line, a linear SLM allows the display of a line of pixels.

Liquid crystal display (1) Is a thin, flat display device made up of any number of color or monochrome pixels arrayed in front of a light source or reflector. (2) A display device which uses certain liquid crystals whose reflectance or transmittance can be changed by applying an electric field.

Liquid crystal on silicon Is a “micro-projection” or “micro-display” technology. It uses liquid crystals applied directly to the surface of a silicon chip coated with an aluminized layer, with some type of passivation layer, which is highly reflective.

Luminaires A complete lighting unit that consist of a lamp or lamps.

Luminance (1) Luminance is a photometric measure of the density of luminous intensity in a given direction. It describes the amount of light that passes through or is emitted from a particular area, and falls within a given solid angle. The SI unit for luminance is candela per square metre (cd/m²). (2) The luminous flux of an elementary source in one direction, divided by the apparent surface area of this source in the direction being considered. (3) The luminance at a point of a surface and in a given direction is the quotient of the luminous intensity in the given direction of an infinitesimal element of the surface containing the point under consideration, by the orthogonally projected area of the surface element on a plane perpendicular to the given direction. (Wyszecki and Stiles 1982). See also radiance.

Luminous flux Luminous flux or luminous power is the quantity derived from radiant flux (radiant power) by evaluating the radiant energy according to its action upon a selective receptor, the spectral sensitivity of which is defined by a standard luminous efficiency function. (Wyszecki and Stiles 1982).

Luminous intensity The luminous intensity in a given direction is the quotient of the luminous flux emitted by a point source in an infinitesimal cone containing the given direction, by the solid angle of that cone. The unit of luminous intensity is the candela (cd). (Wyszecki and Stiles 1982).

Master–slave relationship A model for a communication protocol where one device or process has unidirectional control over one or more other devices. Once a master/slave relationship between devices or processes is established, the direction of control is always from the master to the slaves. In some systems a master is elected from a group of eligible devices, with the other devices acting in the role of slaves.

Micro-electro-mechanical systems (MEMS) Are semiconductors with integrated circuits and three-dimensional features as well as moving parts. MEMS generally range in size from a micrometer (a millionth of a meter) to a millimeter (thousandth of a meter).

Micromirrors Very small mirrors that manipulate light for applications such as optical switching and projection displays.

Micro-opto-electromechanical systems (MOEMS) A special class of MEMS which involves sensing or manipulating optical signals on a very small size scale using integrated mechanical and electrical systems.

Mission function server Provides intervisibility and collision detection functions which are used by a simulator host’s interface to the guest image generator.

Motion parallax The change of angular position of two observations of a single object relative to each other as seen by an observer, caused by the motion of the observer. It is the apparent shift of an object against the background that is caused by a change in the observer's position.

Next generation threat system (NGTS) Serves weapon information (e.g., trajectory, effect) to the host simulation system.

Night vision goggles (NVGs) Electro-optical devices that enhance visibility in low light. They amplify light, through a cascade effect across an electro-optic plate. NVGs are also known as image-intensifier devices. NVGs work in near darkness by detecting ordinary ambient light, usually from the moon and stars that is reflected by objects in the scene being viewed. NVGs contain an image intensifier tube that uses the photoelectric effect to amplify very weak light. As each photon of incoming light collides with a detector plate inside the intensifier tube, the plate ejects several electrons that are further amplified into a cascade of electrons. These electrons are accelerated by a strong electric field towards a phosphor screen which emits light at the point of impact of the electrons. A bright image is thus formed on the phosphor screen.

Noise (1) A random, usually unwanted, fluctuation of pixel values in an image. Image noise can originate in film grain, or in electronic noise in the input device (scanner or digital camera) sensor and circuitry, or in the unavoidable shot noise of an ideal photon detector. (2) A variable physical phenomenon apparently not conveying information and which may be superimposed on, or combined with, a useful signal.

Orthoimage An aerial photograph that has been planimetrically corrected to remove distortion caused by camera optics, camera tilt, and differences in elevation. Orthoimages have the positive attributes of a photograph such as detail and timely coverage, and the positive attributes of a map including uniform scale and true geometry.

Oscilloscope (1) A type of electronic test equipment that allows signal voltages to be viewed, usually as a two-dimensional graph of one or more electrical potential differences (vertical axis) plotted as a function of time or of some other voltage (horizontal axis). (2) An instrument intended to show, in the form of a transitory trace, instantaneous values of a quantity.

Parasitic light Light that causes signal deterioration in a system (e.g., image projection). See also ambient illumination and back reflection.

Phoria The relative directions of the eyes during binocular fixation on a given object in the absence of an adequate fusion stimulus.

Photocathode (1) A cathode, typically of a cesium or sodium compound, having the property of emitting electrons when activated by light or other radiation, via the photoelectric effect. (2) A cathode that functions by photoelectric emission. (3) A metallic or semiconducting layer designed for efficient photoemission of electrons and used in a photoelectric detector. (4) An electrode in a vacuum tube made of a material that emits electrons when light is incident upon it.

Photodiode (1) A type of photodetector capable of converting light into either current or voltage, depending upon the mode of operation. (2) A photoelectric detector in which y absorption of electromagnetic radiation in the neighbourhood of a p-n junction between two semiconductors or a junction between a semiconductor and a metal, produces a change or resistance, voltage, or ultimately photocurrent.

Photoelectron An electron emitted from a system by the photoelectric effect.

Photolithographic mask A template used in photolithography that allows selective exposure of a photosensitive surface.

Photometer An instrument that measures luminous intensity or brightness, luminous flux, light distribution, color, etc., usually by comparing the light emitted by two sources, one source having certain specified standard characteristics.

Photometry (1) The science of measurement of light, in terms of its perceived brightness to the human eye. (2) Measurement of quantities referring to radiation as evaluated according to a given spectral luminous efficiency function, e.g. $V(\lambda)$ or $V'(\lambda)$.

Photomultiplier (1) Are extremely sensitive detectors of light in the ultraviolet, visible and near infrared. These detectors multiply the signal produced by incident light by as much as 10⁸, from which single photons can be resolved. (2) A photoelectric detector comprising a photocathode, an anode and an electron multiplication device that uses the secondary emission of dynodes or channels between photocathode and anode. (3) A photosensitive vacuum tube comprising of a photoemissive cathode and an electron multiplier.

Polyimide A synthetic polymeric resin of a class resistant to high temperatures, wear, and corrosion, used primarily as a coating or film on a substrate substance.

Potentiometer (1) An instrument for measuring an unknown voltage by comparison to a standard voltage. (2) A n-terminal resistor having two outer and one or more intermediate terminals, the latter being fixed or sliding. (3) A voltage measuring instrument in which the voltage to be measured is opposed to a known voltage. (4) A resistor with one or more intermediate tappings allowing fractions of the voltage across the resistor to be obtained.

Probe card An interface between an electronic test system and a semiconductor wafer. Its purpose is to provide an electrical path between the test system and the circuits on the wafer, thereby permitting the testing and validation of the circuits at the wafer level, usually before they are diced and packaged.

Probe station Is used to physically acquire signals from the internal nodes of a semiconductor device. The probe station utilizes manipulators which allow the precise positioning of thin needles on the surface of a semiconductor device.

Quasi-transfer The criterion task is also a simulator task

Radiance (1) Is a radiometric measure that describes the amount of light that passes through or is emitted from a particular area, and falls within a given solid angle in a specified direction. It is used to characterize both emission from diffuse sources and reflection from diffuse surfaces. The SI unit of radiance is watts per steradian per square metre ($\text{W}\cdot\text{sr}^{-1}\cdot\text{m}^{-2}$). See also luminance. (2) The radiant power in a given direction and at a given point of a real or hypothetical surface transmitted by an elementary beam passing through the given point and propagating in the solid angle containing the given direction divided by the product of the value of this solid angle, the area of section of that beam containing the given point and the cosine of the angle between the normal to that section and the direction of the beam.

Radiometry (1) The field that studies the measurement of electromagnetic radiation, including visible light. (2) A measurement of the quantities associated with radiant energy.

Rectilinear Moving in, consisting of, bounded by, or characterized by a straight line or lines.

Red hat One of the larger and more recognized companies dedicated to free software / open source, and a major distributor of the Linux operating system. Red Hat was founded in 1995 and has its corporate headquarters in Raleigh, North Carolina with satellite offices worldwide.

Resolution (1) In imaging, refers to the number of distinct pixels in each dimension that can be displayed. (2) The smallest change in the measurand, or quantity supplied, which causes a perceptible change in the indication. (3) The indication characterizing the finest detail that can be recognized on the picture by a given system.

Schlieren optics The basic optical Schlieren system uses light from a single collimated source shining on, or from behind, a target object. Variations in refractive index caused by density gradients in the fluid distort the collimated light beam.

Smoothing (1) In statistics and image processing, to smooth a data set is to create a function that attempts to capture important patterns in the data, while leaving out noise and consolidating high frequency information. (2) A treatment of an image in which neighbouring values of luminance signal are replaced by one value. The same process may also be carried out for the chrominance signals.

Spatial resolution The ability to sharply and clearly define the extent or shape of features within an image.

Speckle Occurs when the wavefront of a coherent beam is randomly modified and effects image quality.

Spectral power density (SPD) (1) A positive real function of a frequency variable associated with a stationary stochastic process, or a deterministic function of time, which has dimensions of power per Hz, or energy per Hz. (2) For a quantity having a continuous spectrum and a finite mean power, limit, at any frequency, of the quotient of the power within a frequency band containing that frequency by the bandwidth when the bandwidth tends to zero. (3) The distribution as a function of frequency of the power per unit bandwidth of the spectral components of a signal or a noise having a continuous spectrum and a finite mean power.

Stochastic Pertaining to a process involving a randomly determined sequence of observations each of which is considered as a sample of one element from a probability distribution.

Telemetry (1) A technology that allows the remote measurement and reporting of information of interest to the system designer or operator. (2) A process in which measurements are made at some remote location and the results are transmitted by telecommunication.

Telephotometer An instrument for measuring the amount of light emanating from a distant object.

Veiling glare A part of the light reflected on lens interfaces (of a projector) is transmitted toward the screen. It is intrinsic to the projector, but depends on the amount of light passing through the optics and the amount of light projected on the screen. See back reflection.

Weibull psychometric function A continuous probability distribution with probability density function.

Whiteout Whiteout conditions can be caused by circling snow clouds generated by a helicopter's rotorwash, fog, or a snow storm.

References

- Ali SF, Guckenberger D, Rossi M, Williams M (2000) Evaluation of above real-time training and self-instructional strategies for airmanship tasks on a flight simulator. US Air Force Research Laboratory Technical Report No. AFRL-HE-AZ-TR-2000-0112 (Contract No. F41624-98-1-005), pp 74
- Allison RS, Howard IP (2000) Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision Res* 40(14):1869–1885
- Badano A, Flynn MJ (2000) Method for measuring veiling glare in high-performance display devices. *Appl Opt* 39(13):2059–2066
- Berkley WE (1992) Night vision goggle illusions and visual training. Paper presented at the visual problems in night operations (AGARD-LS-187), Neuilly sur Seine, France
- Bigelow AE (1996) Blind and sighted children's spatial knowledge of their home environments. *Int J Behav Dev* 19(4):797–816
- Billingham M, Weghorst S (1995) The use of sketch maps to measure cognitive maps of virtual environments. In: Proceedings of the virtual reality annual international symposium (VRAIS'95), 11–15 March 1995, pp 40–47
- Bissonnette LR (1995) Multiple scattering of narrow light beams in aerosols. *Appl Phys B: Lasers Opt* 60(4):315–323
- Blades M, Lippa Y, Golledge GM, Jacobson RD, Kitchen RM (2002) The effect of spatial tasks on visually impaired peoples' way-finding abilities. *J Vis Impair Blind* 96(6):407–419
- Bliss JP, Lampton DR, Boldovici JA (1992) The effects of easy-to-difficult, difficult-only, and mixed-gunnery practice on performance of simulated gunnery tasks. Technical Report ARITR – 948, Army Research Institute for the Behavioral Sciences, Alexandria, VA. April, 1992 (Report ADA251866, National Technical Information Service), pp 42
- Bloom DM, Tanner AH (2007) Twenty megapixel MEMS-based laser projector. *SID Symp Digest Tech Papers* 38(1):8–11, session 3: Advanced projection systems
- Boynton PA, Kelley, EF (2001) Compensation for stray light in projection display metrology. *SID Symp Digest Tech Papers*, 32(1):334–337, session 21: Improvements in optical metrology
- Bradley A, Kaiser MK (1994) Evaluation of visual acuity with Gen III night vision goggles. NASA Technical Memorandum # 108792. National Aeronautics and Space Administration, Moffett Field, CA
- Braithwaite MG, Douglass PK, Durnford SJ, Lucas G (1998) The hazard of spatial disorientation during helicopter flight using night vision devices. *Aviat Space Environ Med* 69(11):1038–1044
- Brown LG (1992) A survey of image registration techniques. *ACM Comput Surv (CSUR)* 24(4):325–376
- Burt P (1988) Smart sensing within a pyramid vision machine. *Proc IEEE* 76(8):1006–1013
- CAT, The Journal for Civil Aviation Training (2007) Annual training sourcebook, Issue 4. Halldale Media Group
- CIE (1983) The basis of physical photometry, Publication CIE, No 18.2 (TC-1.2). Commission internationale de l'éclairage, CIE Central Bureau, Kegelgasse 27, A-1030 Vienna, Austria
- Craig G, Macuda T, Thomas P, Allison R, Jennings S (2005) Light source halos in night vision goggles: psychophysical assessments (conference proceedings paper, March 29). In: Rash CE, Reese CE (eds) *Helmet- and head-mounted displays X: technologies and applications*. Proceedings of SPIE – the international society for optical engineering, vol 5800, no. 1, Orlando, FL, 19 May 2005, pp 40–44
- Cutting JE, Millard RT (1984) Three gradients and the perception of flat and curved surfaces. *J Exp Psychol: Gen* 113(2):198–216
- Deering S, Cheriton D (1990) Multicast routing in datagram internetworks and extended LANs. *ACM Trans Comput Syst* 8(2):85–110
- DeLucia PR, Task HL (1995) Depth and collision judgment using night vision goggles. *Int J Aviat Psychol* 5(4):371–386

- DeVilbiss CA, Ercoline WR, Antonio JC (1994) Visual performance with night vision goggles (NVGs) measured in US Air Force aircrew members. In: *Helmet- and head-mounted displays and symbology design requirements*. Paper presented at the proceedings of the SPIE – the international society for optical engineering, vol 2218. Orlando, FL, 5–7 April 1994, pp 64–70
- Doucet M, Picard F, Niall KK, Jerominek H (2005) Operation modes for a linear array of optical flexible reflective analog modulators. In: Hopper DG, Forsythe EW, Morton DC, Bradford CE, Girolamo HJ (eds) *Cockpit and future displays for defense and security*. Proceedings of the SPIE defense & security symposium, Orlando, FL, 28–31 March 2005, vol 5801. SPIE – The International Society for Optical Engineering, Bellingham, WA, pp 219–233
- Duncan CP (1958) Transfer after training with single versus multiple tasks. *J Exp Psychol* 55:63–72
- Durham L (2006) Software design document for the CIGI Host Emulator. v. 3.1.x. (Document ID: TST05I030, Rev. A). Boeing Company Training and Support Technology Research & Development
- Emmert E (1881) *Großenverhältnisse der Nachbilder*. *Klin Monbl Augenheilkd* 19:443–450
- Farmer WM (2001) *The atmospheric filter*. JCD Publishing, Winter Park, FL
- Federal Aviation Administration (1991) *Airplane simulator qualification advisory circular 120-40B*
- Federal Aviation Administration (1995) *Airplane simulator qualification advisory circular 120-40C (Draft)*
- Felzenszwalb P, Huttenlocher D (1998) Image segmentation using local variation. In: *Proceedings of the IEEE computer society international conference on computer vision and pattern recognition*, pp 98–104
- Flach JM, Warren R, Garness SA, Kelly L, Stanard T (1997) Perception and control of altitude: splay and depression angles. *J Exp Psychol: Hum Percept Perform* 23(6):1764–1782
- Forsyth D, Ponce J (2003) *Computer vision: a modern approach*. Prentice Hall, Upper Saddle River, NJ
- Freeman W, Adelson E (1991) The design and use of steerable filters. *IEEE Trans Pattern Anal Mach Intell* 13(9):891–906
- Galanis G, Jennings A, Beckett P (2001) Runway width effects in the visual approach to landing. *Int J Aviat Psychol* 11(3):281–301
- Gauthier MS, Parush A, Macuda T, Tang D, Craig G, Jennings S (2008) The impact of night vision goggles on way-finding performance and the acquisition of spatial knowledge. *Hum Factors* 50(2):311–321
- Geri GA, Martin EL, Wetzel PA (2002) Head and eye movements in visual search using night vision goggles. *Aviat Space Environ Med* 73(8):779–786
- Geri GA, Winterbottom MD, Pierce BJ (2004) Evaluating the spatial resolution of flight-simulator visual displays. AFRL-HE-AZ-TR-2004-0078, June 2004. Link Simulation and Training, Mesa, AZ, pp 41
- Giachritsis CD, Harris MG (2005) Global versus local image expansion in estimating time-to-contact from complex optic flow. *Perception* 34(5):577–585
- Gibb RW (2007) Visual spatial disorientation: revisiting the black hole illusion. *Aviat Space Environ Med* 78(8):801–808
- Gibbon E (1776) *History of the decline and fall of the Roman empire*, vol. 1. Various editions, including Bury JB (ed) (1906). Fred de Fau & Company, New York. Available June 2009 from <http://www.gutenberg.org/etext/731>
- Gibson JJ (1950) The perception of visual surfaces. *Am J Psychol* 63:367–384
- Gillner S, Mallot HA (1998) Navigation and acquisition of spatial knowledge in a virtual maze. *J Cogn Neurosci* 10(4):445–463
- Goldin SE, Thorndyke PW (1982) Simulating navigation for spatial knowledge acquisition. *Hum Factors* 24(4):457–471
- Gray R, Regan D (1999) Motion in depth: adequate and inadequate simulation. *Percept Psychophys* 61(2):236–245
- Guckenberger D, Uliano KC, Lane NE (1993) Teaching high-performance skills using above-real-time training. National Aeronautics and Space Administration, NASA Contractor Report 4528
- Halsted CP (1993) Brightness, luminance, and confusion. *Inf Disp* 9(3):21–24

- Harris MG, Giachritsis CD (2000) Coarse-grained information dominates fine-grained information in judgments of time-to-contact from retinal flow. *Vis Res* 40(6):601–611
- Hilgard E, Bower GH (1966) *Theories of learning*, 3rd edn. Appleton-Century-Crofts, New York
- Hughes PK, Zalevski AM, Gibbs P (2000) Visual acuity, contrast sensitivity, and stereopsis when viewing with night vision goggles. Technical Report DSTO-TR-1012, Defence Science and Technology Organization. Aeronautical and Maritime Research Laboratory – Air Operations Division, Melbourne
- Hung ES, Senturia SD (1999) Extending the travel range of an analog-tuned electrostatic actuator. *J Microelectromech Syst* 8:497–505
- IESNA (2000a) *IESNA lighting handbook*, 9th edn. IES, Illuminating Engineering Society of North America, New York
- IESNA (2000b) *Roadway lighting*. ANSI/IESNA RP-8-00, New York: IES, Illuminating Engineering Society of North America, available at iesna.org. (ANSI/IES RP-8-00, 1983, reaffirmed 2005: Roadway lighting; American National Standards Institute/Illuminating Engineering Society of North America)
- Jahne B (1991) *Digital image processing*. Springer, Berlin
- Jennings S, Craig G (2000) Effects of field-of-view on pilot performance in night vision goggles flight trials: preliminary findings. In: *Proceedings of the SPIE*, vol. 4021, Paper presented at *Helmet-and Head-Mounted Displays V*, Orlando, FL, pp 335–342
- Kilgore JI (1989) The planes that never leave the ground. *Am Herit Invent Technol* 4(3):56–63
- Klatzky RL (1998) Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections [Electronic version]. In: Freska C, Habel C, Wender KF (eds) *Spatial cognition: An interdisciplinary approach to representing and processing of spatial knowledge* (Lecture Notes in Computer Science, vol. 1404). Springer, Berlin, 1–17
- Klein F (1939) *Elementary mathematics from an advanced standpoint: geometry* (Hedrick EH, Noble CA Trans.). Dover, New York
- Knight KK, Apsey DA, Jackson WG, Dennis RJ (1998) A comparison of stereopsis with ANVIS and F4949 night vision goggles. *Aviat Space Environ Med* 69(2):99–103
- Knill DC (1998) Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture. *Vis Res* 38(17):2635–2656
- Knill DC, Saunders JA (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vis Res* 43(24):2539–2558
- Koonce JM, Bramble WJ Jr (1998) Personal computer-based flight training devices. *Int J Aviat Psychol* 8(3):277–292
- Lamdan Y, Schwartz J, Wolfson H (1990) Affine invariant model-based object recognition. *IEEE Trans Robot Autom* 6(5):578–589
- Lee B, Sagong D, Jeong G (2001) LCDs: How fast is enough? SID (society for information display) symposium digest of technical papers, session 41: Spatial temporal artifacts and issues, 32(1), June, 1106–1109
- Levi DM, Klein SA, Aitsebaomo P (1984) Detection and discrimination of the direction of motion in central and peripheral vision of normal and amblyopic observers. *Vis Res* 24:789–800
- Lieberman D (1990) *Learning: behavior and cognition*. Wadsworth, Belmont, CA
- Lin K, Wang M, Wang J, Schab D (1995) The smoothing of dead reckoning image in distributed interactive simulation. In: *Proceedings of the 1995 American institute of aeronautics and astronautics flight simulation technologies conference*, AAIA-1995-3398, pp 83–87
- Lintern G (1991) An informational perspective on skill transfer in human-machine systems. *Hum Factors* 33(3):251–266
- Lintern G, Koonce JM (1992) Visual augmentation and scene detail effects in flight training. *Int J Aviat Psychol* 2(4):281–301
- Lintern G, Walker MB (1991) Scene content and runway breadth effects on simulated landing approaches. *Int J Aviat Psychol* 1(2):117–132
- Loftin RB, Kenney P (1995) Training the Hubble space telescope flight team. *IEEE Comput Graph Appl* 15(5):31–37
- Low DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110

- Macuda T, Allison R, Thomas P, Craig G, Jennings S (2004) Detection of motion- defined form under simulated night vision conditions. In: *Helmet- and head-mounted displays IX: technologies and applications*. Proceedings of the society of photo-optical instrumentation engineers (SPIE), vol 5442. pp 36–44
- Maida JC, Bowen CK, Pace JW (2006). Enhancing lighting techniques and augmented reality to improve human task performance. NASA Technical Report NASA/TP-2006-213724; S-981. Johnson Space Center, July, pp 34
- Maida JC, Bowen CK, Pace JW (2007) Improving robotic operator performance using augmented reality. Proceedings of the human factors and ergonomics society 51st annual meeting, pp 1635–1639
- McKee SP, Nakayama K (1984) The detection of motion in the peripheral visual field. *Vis Res* 24:25–32
- McKee SP, Welch L, Taylor DG, Bowne SF (1990) Finding the common bond: Stereoacuity and the other hyperacuities. *Vis Res* 30(6):879–891
- McLean WE, Rash CE, McEntire J, Braithwaite MG, Mora JC (1997) A performance history of AN/PVS-5 and ANVIS image intensification systems in U.S. Army aviation. In: *Proceedings of the society of photo-optical instrumentation engineers (SPIE), Head-Mounted Displays II*, vol 3058. pp 264–298
- McNamara TP (1986) Mental representation of spatial relations. *Cogn Psychol* 18(1):87–121
- Menendez RG, Bernard JE (2000) Flight simulation in synthetic environments. In: *Proceedings of the 19th digital avionics systems conferences*, vol 1. pp 2A5/1–2A5/6
- Menzio M, Napflin U, Krueger H (1999) CRT versus LCD: A pilot study on visual performance and suitability of two display technologies for use in office work. *Displays* 20:3–10
- Mestre DR, Masson GS, Stone LS (2001) Spatial scale of motion segmentation from speed cues. *Vis Res* 41(21):2697–2713
- Michaels DD (1985) *Visual optics and refraction: a clinical approach*, 3rd edn. Mosby, St. Louis
- Montello DR (2005) Navigation. In: Shah P, Miyake A (eds) *The Cambridge handbook of visuospatial thinking*. Cambridge University Press, Cambridge, pp 257–294. Accessed June 2009 at <http://www.geog.ucsb.edu/~montello/pubs/navigate.pdf>
- Nakayama K, Tyler CW (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vis Res* 21(4):427–433
- Niall KK, Reising JD, Martin EL (1999) Distance estimation with night vision goggles: A little feedback goes a long way. *Hum Factors* 41(3):495–506
- Ode MC (2003) Understanding Luminaires and Lamps, *Electrical Contractor*, June. <http://www.ul.com/global/eng/documents/offering/industries/lighting/lightingindustryservices/articles/0603.pdf>. Accessed June 2009
- O'Rourke J (1994) *Computational geometry in C*. Cambridge University Press, Cambridge
- Palmisano S, Gillam B (2005) Visual perception of touchdown point during simulated landing. *J Exp Psychol: Appl* 11(1):19–32
- Parush A, Berman D (2004) Navigation and orientation in 3D user interfaces: the impact of navigation aids and landmarks. *Int J Hum Comput Studies* 61:375–395
- Parush A, Ahuvia-Pick S, Erev I (2007) Degradation in spatial knowledge acquisition when using automatic navigation systems. In: Winter S, Duckham M, Kulik L, Kuipers B (eds) *Spatial information theory (Lecture notes in Computer Science)*, vol 4736. Springer, New York, pp 238–254
- Pavlov IP (1928) *Lectures on conditioned reflexes*. International Publishers, New York
- Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C (1995) A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. *Brain Cogn* 28(1):39–58
- PHELPS WB (2002) Interface control document for the common image generator interface (CIGI) Version 2.0. Boeing Corporation, St. Louis, MO. Report TST02I015, available June 1009 from <http://cigi.sourceforge.net>. (Version 2.1 dated January 28, 2003)
- Picard F, Campillo C, Niall KK, Larouche C, Jerominek H (2002) MEMS-based light valves for ultra-high resolution projection displays. In: *Proceedings of the interservice/industry training, simulation & education conference Orlando, FL, 2–5 December 2002*

- Picard F, Campillo C, Pope TD, Niall KK, Peppler P, Larouche C, Jerominek H (2003) Flexible micromirror linear array for high resolution projection displays. In: Urey H (ed) MOEMS display and imaging systems. Proceedings of the micromachining symposium of photonics west, San Jose, CA, 28–29 January 2003, vol 4985. SPIE – The International Society for Optical Engineering, Bellingham, WA, pp 44–55
- Picard F, Doucet M, Niall KK, Larouche C, Savard M, Crisan S, Thibault S, Jerominek H (2004) Recent advances in a linear micromirror array for high-resolution projection. In: Cien L-C, Wu MH (eds) Liquid crystal materials, devices, and applications X and projection displays X. Proceedings of electronic imaging: science and technology, San Jose, CA, 21–22 January, 2004, vol 5289, no. 28. SPIE – The International Society for Optical Engineering, Bellingham, WA, pp 284–293
- Poor A (2003) Liquid-crystal displays: big and bigger. *Inf Disp* 19(9):16–19
- Rabin J, Wiley R (1994) Switching from forward-looking infrared to night-vision goggles: transitory effects on visual resolution. *Aviat Space Environ Med* 65(4):327–329
- RCA (1974) RCA electro-optics handbook. Technical series EOH-11, 2nd edn. [RCA solid state division, electro-optics and devices, RCA Corp., Lancaster] RCA/Commercial Engineering, Harrison, NJ
- Roessingh JJM (2005) Transfer of manual flying skills from PC-based simulation to actual flight-comparison of in-flight measured data and instructor ratings. *Int J Aviat Psychol* 15(1):67–90
- Salazar G, Temme L, Antonio JC (2003) Civilian use of night vision goggles. *Aviat Space Environ Med* 74(1):79–84
- Sheehy JB, Wilkinson M (1989) Depth perception after prolonged usage of night vision goggles. *Aviat Space Environ Med* 60(6):573–579
- Shima T, Itakura T, Yamada S, Minamizaki H, Ishioka T (1995) Principle and applications of an autocharge-compensated sample and hold circuit. *IEEE J Solid-State Circuits* 30:906–912
- Simtech, Inc. (1998) Air force distributed mission training technical analysis. Aeronautical systems center training systems program office Wright-Patterson AFB, OH 45433, Contract No. F33657-94-D-2253
- Slater M, Covas CM (2007) Effects of long-haul network connectivity on the visual fidelity of real-time flight simulations. In: Proceedings of the simulation interoperability standards organization 2007. Fall simulation interoperability workshop, September 2007 (Paper No. 07F-SIW-023)
- Smeulders A, Huang T, Gevers T (eds) (2004) Special issue on content-based image retrieval. *Int J Comput Vis* 56(1–2):5–6
- Snoek L, Worring M (2005) Multimodal video indexing: A review of the state-of-the-art. *Multimed Tools Appl* 25(1):5–35
- Sweet BT (2006) Visual cueing for vehicular control. Presented at the AIAA modeling and simulation technologies conference (Panel Session: MSTC Motion Working Group), Keystone, CO, August 22
- Sweet BT, Kaiser MK (2006) Integration of size and binocular disparity visual cues in manual depth-control tasks. AIAA modeling and simulation technologies conference, Keystone, CO, 21–24, August. AIAA Technical Papers (AIAA 2006-6628), Washington, DC
- Tanaka K, Fukada Y, Saito HA (1989) Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 62(3):642–656
- Task HL (2001) Night vision goggle visual acuity assessment: Results of an interagency test. In: Proceedings of SPIE – The international society for optical engineering: helmet- and head-mounted displays VI, Vol 4361. Orlando, FL, pp 130–137
- Taylor HL, Lintern G, Hulin CL, Talleur DA, Emanuel TW Jr, Phillips SI (1999) Transfer of training effectiveness of a personal computer aviation training device. *Int J Aviat Psychol* 9(4):319–335
- Thomas PJ, Allison RS, Carr P, Shen E, Jennings S, Macuda T et al (2005) Physical modeling and characterization of the halo phenomenon in night vision goggles. Paper presented at the helmet- and head-mounted displays X: technologies and applications, 29 March 2005, Orlando, FL, USA. Proceedings of the SPIE – The International Society for Optical Engineering, vol 5800, no. 1, pp 21–3, USA

- Thorndyke PW, Hayes-Roth B (1982) Differences in spatial knowledge acquired from maps and navigation. *Cogn Psychol* 14(4):560–589
- Transport Canada (1993) Aerodrome standards and recommended practices, 4th edn (TP No. 312E). Transport Canada, Ottawa, ON. <http://www.tc.gc.ca/publications/EN/TP312/PDF%5CHR/TP312E.pdf> Accessed June 2009
- Tyler CW, Torres J (1972) Frequency response characteristics for sinusoidal movement in the fovea and periphery. *Percept Psychophys* 12:232–236
- van Ee R, van Dam LC, Erkelens CJ (2002) Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *J Vis* 2(9):597–607
- van Raalte JA (1970) A new Schlieren light valve for television projection. *Appl Opt* 9(10):2225–2230
- Video Electronics Standards Association (2001) Flat panel display measurements standard, version 2.0. Video Electronics Standards Association (VESA), pp 76–77
- Watson AB, Ahumada AJ Jr (2004) The spatial standard observer [abstract]. *J Vis* 4(8):51, 51a
- Wentzel KC (1995) An informative guide to qualification of flight simulators for commercial aircraft. AIAA flight simulation technologies conference, Baltimore, MD, 7–10 August. AIAA Technical Papers (A95-39235 10-01)
- Wildes R, Hirvonen D, Hsu S, Kumar R, Lehman B, Matei B, Zhao W (2001) Video georegistration: Algorithm and quantitative evaluation. In: Proceedings of the eighth IEEE international conference on computer vision, pp 343–350
- Winterbottom MD, Geri GA, Morgan WD, Pierce BJ (2004) An integrated procedure for measuring the spatial and temporal resolution of visual displays. In: Proceedings of the interservice/industry training, simulation and education conference, ADA430109, Orlando, FL, Paper No. 1855
- Witmer BG, Bailey JH, Knerr BW, Parsons KC (1996) Virtual spaces and real world places: transfer of route knowledge. *Int J Hum Comput Stud* 45:413–428
- Wyszecki G, Stiles WS (1982) Color science, concepts and methods, quantitative data and formulae, 2nd edn. Wiley, New York
- Yamamoto T, Aono Y, Tsumura M (2000) Guiding principles for high quality motion picture in AMLCDs applicable to TV monitors. *Soc Inf Disp Symp Dig Tech Pap* 31(01):456–459
- Young MJ, Landy MS, Maloney LT (1993) A perturbation analysis of depth perception from combinations of texture and motion cues. *Vis Res* 33:2685–2696
- Zalevski A, Meehan JW, Hughes PK (2000) Size estimation with night vision goggles. Technical Report DSTO – RR – 0201. Defence Science and Technology Organisation (DSTO Air Operations Division, Aeronautical and Maritime Research Laboratory, Fishermans Bend, VIC, Australia), pp 22
- Zalewski EF (1995) Radiometry and photometry. In: Blass M (ed) *Handbook of optics*, vol II, 2nd edn. McGraw-Hill, New York

Index

A

Above-real-time training (ARTT), 143, 145–148, 155, 156, 161
Aerial video, 164, 165, 167–169, 176
Aerosol, 177–187
Aerosol penetration, 179, 180, 183–187
Affine, 163–174, 176
Affine invariants, 164, 168–172
Affine transformations, 170
Air Force Research Laboratory (AFRL), 88, 89, 98, 99
American National Standards Institute (ANSI), 54, 55
Analysis of covariance (ANACOVA), 155, 156, 159
Analysis of variance (ANOVA), 94, 106, 134
Antialiasing, 91
Appearance matching, 173
Appearance similarity, 173
Augmented reality, 8
Augmented visionics system (AVS), 177–187
Automatic gain control (AGC), 127, 129
Aviator night vision imaging system (ANVIS), 114, 126
Azimuth, 145, 148

B

Background, 28, 40, 51, 56, 65, 91, 127–129
Bidirectional reflectance distribution function (BDRF), 28, 32–34
Bistatic optics, 181
Blur, 97–101, 103, 105–109
Brownout, 179, 182, 184, 185

C

Calibration, 32, 101, 152
Cathode ray tube (CRT), 14, 31, 48, 91, 97, 98, 100, 104–109
Centroids, 171
CF-18, 61, 63–65, 80
Collimated, 10, 33
Color separation, 103
Comet, 7
Commission international de l'éclairage (CIE), 32, 55
Common image generator interface (CIGI), 61–72, 75–77, 82–83, 85
Contrasts, 1, 16, 22, 27–48, 50, 65, 101, 102, 106, 125, 135, 144, 149, 155, 163, 165–170, 175, 176
Criterion, 96, 134, 141–145, 148, 150–152, 154–156
Criterion performance, 143, 148
Cultural lighting, 51–60

D

Dark ride, 5, 6
Database, 1, 57, 59, 60, 66, 70, 72, 83–85, 121, 163–165, 167, 168, 170–176
Day for night, 4–6
Dead reckoning, 95, 96
Defence Research & Development Canada (DRDC), 61, 64, 183, 187
Detection, 27, 29, 32, 44, 47, 64–66, 89, 97, 109, 177, 181, 183–185, 187
Detection threshold, 87–96
Digital light processor (DLP), 97–100, 104–106, 108, 109
Digital video interface (DVI), 13–26, 77, 78, 81
Digital-to-analog converter (DAC), 14, 16, 22–24, 26

Direction, 28, 29, 31–38, 44, 46, 47, 51, 52, 54, 66, 95, 102, 103, 111, 112, 115–120, 126, 127, 137, 164, 165, 173, 174

Display, 1, 14, 27, 52, 61, 75, 91, 97, 113, 131, 144

Display characterization, 32

Display technology, 100

Distance, 6, 8, 34–36, 42, 45, 52, 55, 56, 60, 63, 65, 87–96, 105, 113, 115, 116, 118–135, 137, 139, 146, 149, 155, 156, 160, 170, 178, 179, 182, 183, 185

Distributed rendering (DR), 75, 82–83, 85

Distributed simulation, 64, 83

Distributive interactive simulation (DIS), 65, 66, 83, 95

Dust, 177, 179, 184, 185

E

Electro-optical (EO), 52, 111, 113

Electronics Industry Alliance (EIA), 80

Expert, 4, 8–10, 122, 144

Extravehicular (EVA), 8

F

FAA certification, 9

Familiarization trials, 145, 150, 152–154, 156

Federal Aviation Administration (FAA), 2, 9, 11, 144

Fidelity, 2, 5, 8–11, 13, 50–53, 56, 58, 87, 123, 125, 142

Field of view (FOV), 82, 103, 113, 114, 120, 121, 125–127, 149, 169, 179, 182

Filters, 32, 46, 47, 169, 171

Flexible reflective analog modulators (FRAM), 13–26

Flight, 5, 9, 11, 50, 51, 53, 75, 82, 84, 87, 88, 90, 95, 96, 130–133, 141–161, 177, 178, 180, 183–186

Flight simulation, 1–3, 5, 6, 11, 16, 50–52, 61, 64, 75, 83, 87–109, 141, 143, 144, 147, 148, 155, 161

FlightIG, 61, 62, 64, 66, 67, 69, 72, 75–77, 80, 82, 83

Flying spot, 31, 41, 44, 46

Foreshortening, 131, 135

Formation flight, 90, 95, 96

Frames, 4, 5, 7, 14, 22, 23, 26, 44, 46, 58, 61, 65–67, 72, 77, 84, 90, 92, 97, 100, 103, 108, 123, 125, 132, 169, 171, 179, 184

G

Gamma function, 101

Gap width, 101–103, 106–109

Genlock, 61, 66, 77, 80, 81, 131, 196

Genlock card, 80, 81

Geographic database, 1

Geometric descriptor, 169

Geospatial database, 163, 165, 167, 176

Glideslope, 124, 131, 136–140

Graphics, 2–4, 60, 62, 64, 67, 72, 75, 77, 78, 80, 81, 83, 85, 91, 121, 131

Graphics card, 72, 75, 77, 80, 81, 91

Graphics processor units (GPUs), 77

H

Halo, 123–127, 130–140

Hardware, 61, 64, 65, 72, 75, 77, 80–81, 84, 149

Helicopter, 130–132, 145, 177–180, 184–187, 191

High-level architecture (HLA), 83

Hold-time, 97, 98, 100, 103–109

I

Identification, 68, 97, 109

Illuminating Engineering Society of North America (IESNA), 54, 55

Image appearance, 163–165, 168–171, 175, 176

Image generation, 53, 61, 64, 65, 67

Image generator (IG), 10, 61–77, 87

Image geometry, 163, 165, 168, 176

Image indexing, 163–176

Image intensifier, 123, 125, 126

Image quality, 10, 16, 43, 108

Indexing, 163–176

Individual differences, 148, 160–161

Infrared (IR), 7, 145

International Commission on Illumination (CIE), 55

Inverse video, 7

Irradiance, 33, 45, 57

Irradiance map, 33, 53, 57, 59

J

Jitter, 87, 92

L

Lambertian, 34

Laser, 16, 17, 27–49, 61, 62, 64, 65, 69, 70, 72, 75–78, 82–85, 100, 177, 178, 181, 182

Laser based video projector (LBVP), 27–48

Latin square, 150, 152
 Light detection and ranging (LIDAR),
 177–202
 Light source, 47, 49–60, 113, 124–128, 132,
 137, 179
 Line pair, 102, 103, 106–109
 Linux, 66, 70, 77, 79, 80, 82, 131
 Liquid crystal display (LCD), 83, 100
 Liquid crystal on silicon (LCoS), 97, 100
 Look-up table (LUT), 22, 164, 172
 Low light, 113, 125
 Luminaire, 54–56, 59
 Luminance, 27–29, 32, 33, 37, 38, 41, 45–48,
 52, 91, 99, 100, 103, 104, 109

M

Map, 33, 116, 118, 119, 122
 Methodology, 2, 5, 10, 119, 120
 Micro optical mechanical system (MOEMS),
 15, 16
 Micromachining, 16
 Micromirror, 16–18, 21–23
 Micromirror curvature, 16
 Modeling, 10, 51–59, 120, 125
 Monostatic optics, 181
 Motion parallax, 10, 130, 132
 Motion platform, 5
 Multi-task trainer (MTT), 64, 67
 Multiscale, 168, 170, 171, 173
 Multivariate repeated-measures analysis
 of variance (MANOVA), 134

N

Navigation, 3, 4, 53, 111–113, 117–122, 190
 Network, 21, 61, 64–66, 70, 72, 75, 77, 78,
 82–85, 90, 95, 96
 Networked, 87–96, 191
 Night vision, 1, 7, 111–140
 Night vision device (NVD), 7, 123–140
 Night vision goggle (NVG), 51–53, 56–58,
 111–122, 124
 NVG simulation, 52, 53, 124

O

Object recognition, 168
 Observer, 10, 28, 32, 35, 37, 40, 45, 48, 52,
 56, 58, 88, 90–96, 101–103, 105, 114,
 120, 121, 126, 127, 129–138, 143, 145,
 147–156, 159–161
 Obstacle, 177–187
 Onset, 100, 101, 108

Optic flow, 131, 136, 137, 139
 Optical damage, 17
 Out-the-window (OTW), 52, 53, 57, 58, 60,
 149–151

P

Parametric, 55
 Parasitic light, 28, 39–40, 45–47
 Perceptual, 8, 90, 98, 101–103, 109, 119,
 124–126, 129, 130, 132
 Performance criteria, 10
 Personal computer (PC), 22–24, 26, 76, 77,
 79–85, 91, 131, 142–144, 148
 Perspective, 9, 10, 57, 124, 130–133, 135,
 136, 139
 Photocathode, 130
 Photometric, 55, 165, 168, 176
 Physics-based, 1, 7, 10, 52–55, 58–60, 131
 Pixel, 7, 14–16, 18–24, 26, 28–33, 35, 40–42,
 44–46, 56, 64, 72, 76, 77, 88, 89, 91, 95,
 98, 100–103, 105, 107, 127, 148, 149,
 152, 154, 156, 158, 164, 165, 167, 174
 Pixel count, 100, 102
 Point source, 51–56, 59, 126, 128–130
 Positional error, 88, 90, 96
 Post-test, 142, 148, 149, 151, 152, 154–160
 Pre-test, 148–152, 154–156, 160
 Prefilter, 164, 165, 172, 176
 Probe, 14, 19–21, 26, 164, 171, 173, 174, 176
 Processor, 71, 72, 76, 77, 79, 80, 84, 100
 Progressive training, 147, 148, 161
 Projection, 4, 14–19, 21, 24, 26, 28,
 30–32, 40, 57, 76, 77, 100, 121,
 130, 131, 137
 Projector, 4, 13–48, 64, 72, 77–85, 91,
 97–109, 131
 Psychometric function, 92
 Psychophysics, 10, 11, 123–140

Q

Quasi-transfer experiment, 144
 Query, 173, 174

R

Radiance, 7, 28, 29, 32–37, 41, 48, 52, 54,
 56–58, 60
 Range, 11, 16, 21, 38, 46, 51, 52, 56, 70, 117,
 126–128, 133, 137, 147, 148, 164, 165,
 168, 169, 172, 176, 179–185
 Range-gated camera, 179
 Realism, 51–53, 58, 88, 90

- Reflected light, 40, 51–53, 56–57, 180
- Resolution, 1, 5, 11, 14, 16, 23, 56, 57, 64, 65, 72, 76–78, 85, 95–109, 124, 125, 127, 131, 136, 141–161, 164, 173, 174, 179, 180
- Rotation, 41, 47, 66, 117, 150, 154, 166, 173, 175, 182
- S**
- Sampling, 27, 28, 32, 38–43, 45, 46, 187
- Scale invariance, 173
- Scattering, 28, 43, 53, 177, 178, 180–185
- Schlieren optics, 14, 16, 20
- Screen, 4, 10, 24, 27, 28, 31–40, 43–45, 47, 48, 61, 64, 65, 92, 120, 121, 131, 132, 137, 141–161
- Simulation, 1–11, 16, 17, 49–85, 87–90, 92, 94–109, 120, 124, 125, 131–133, 136, 139, 143–148, 150, 154, 155, 161, 180
- Slope, 123, 126, 128, 132–134, 136
- Smoothing, 87, 88, 90, 92–96
- Snellen acuity, 117
- Software, 5, 14, 21, 22, 26, 55, 56, 60–62, 64–67, 69, 71, 72, 74, 75, 77, 79–81, 83, 85, 91, 109, 149, 152, 155, 163, 168, 169
- Spatial cognition, 111–122
- Spatial displacement, 90, 92
- Spatial light modulator (SLM), 16, 31, 40, 44, 46
- Spatial resolution, 95, 96, 99–101, 105, 109, 164, 173, 174, 179
- Spatial standard observer, 10
- Speckle, 40, 42–43
- Spectral characteristics, 125, 130
- Spectral response, 32, 37, 46
- Street lighting, 55
- Substrate, 16–18
- Symposium, 1, 7, 10, 11, 62
- T**
- Target aircraft, 88, 94, 148–150, 153–155
- Target speed, 98, 106, 109, 145, 147
- Telephotometer, 31, 32, 38, 43
- Temporal resolution, 97–109, 144
- Temporal response, 44, 98–101, 103, 104, 106, 109
- Terrain, 2, 4, 6, 50, 53, 57, 58, 65–70, 72, 83–85, 113, 130, 131, 137, 174
- Texture, 56–58, 66, 80, 124, 130–133, 135, 136, 138, 139, 149, 163–171, 173, 175, 176
- Texture defined region, 163, 165, 168, 170, 176
- Texture gradient, 124, 130, 132, 135, 136, 139
- Threshold, 17, 87–96, 98, 102, 106, 132
- Training condition, 145–148, 154–156
- Training session, 120, 133, 134, 143, 145–147, 152, 154, 159–161
- Training speed, 145, 148–152, 155–161
- Training system, 50–53, 73, 77
- Training trials, 145, 146, 150, 152–156
- Transfer of training, 143–144, 148
- Transfer performance, 144
- U**
- Ultra-high resolution, 1, 74, 76, 77
- Uniformity, 18, 20–22, 40, 41
- Urban, 51–53, 55, 56, 174
- V**
- Variance, 133–135, 137, 148, 155, 156, 158–160, 170, 171, 173
- Vector, 2, 35–38, 164, 170–173
- Veiling glare, 28, 29, 39, 40, 45–47
- Veridical, 10, 133
- Vestibular cues, 5
- Video, 1, 6–8, 14, 16, 22, 23, 26–48, 72, 73, 76, 77, 80, 81, 84, 97, 99, 101, 131, 132, 147, 152, 163–176
- Video imagery, 1, 173, 175
- Video-to-reference, 163–176
- Visual, 2–8, 10, 11, 20, 32, 37, 46, 50, 52–56, 58–65, 67, 72, 74, 83, 84, 87–96, 107, 113, 114, 116, 117, 121, 122, 127, 128, 130–132, 135, 136, 142–145, 149, 150, 179
- Visual acuity, 64, 83, 91, 117, 149
- Visual flight attachment (VFA), 6
- Visual system, 2, 4, 5, 8, 10, 32, 37, 46, 50, 52, 88, 96, 135, 136
- W**
- Wavelength, 38, 54–56, 130, 179, 184, 185
- Way-finding, 111–122
- Whiteout, 179, 185–187