CHOOSING YOUR POISON: OPTIMIZING SIMULATOR VISUAL SYSTEM SELECTION AS A FUNCTION OF OPERATIONAL TASKS

Barbara T. Sweet
Aerospace Engineer/AST Human-Machine Systems
NASA Ames Research Center

Mary K. Kaiser
Research Psychologist
NASA Ames Research Center

ABSTRACT

Although current-technology simulator visual systems can achieve extremely realistic levels, they do not completely replicate the experience of a pilot sitting in the cockpit looking at the outside world. Some differences in experience are due to visual artifacts, or perceptual effects that would not be present in a naturally viewed scene. Others are due to features or cues that are missing from the simulated scene. The depiction of depth in displays is especially prone to such artifacts and cue conflicts. In this paper, the differences between natural and simulated scenes will be defined, and discussed in terms of the capabilities and limitations of various visual system technologies. The significance of these differences will be examined as a function of several particular operational tasks. A framework to facilitate the choice of visual system characteristics based on operational task requirements will be proposed.

INTRODUCTION

One of the greatest challenges for simulator visual systems is to recreate the spatial relationship of the pilot's vehicle to the surrounding environment. In addition to perceiving "what" is surrounding the aircraft, it is critical that the pilot can extract "where" objects are. In the natural environment, "where" information is provided by a number of visual cues working in concert. In simulator visual systems, it can be difficult to recreate all these depth cues; in fact, it is not uncommon for anomalous (and hence conflicting) depth cues to be present in simulator displays.

Depth cues have been traditionally grouped by the underlying sources of information. Thus, this taxonomy differentiates between Primary (a.k.a. Physiological) cues, which are derived from the physiological mechanisms of accommodation, convergence, and stereopsis, and Secondary (a.k.a. Pictorial or Psychological) cues. Generally speaking, the Primary/Physiological cues map to the physical and optical characteristics of the simulator visual system's displays, whereas the Pictorial/Psychological cues are created by the system's image generator. More recently, the traditional taxonomy has been expanded to include the motion-specified depth information contained in optic flow, particularly motion parallax and radial expansion.

Beyond enumerating the various cues and their underlying sources, psychologists and simulator engineers have become increasingly interested in how humans integrate these sources of depth information to navigate through 3-D space. Thus, it has become increasingly important to understand what information each cue can provide *within a particular operational context*, and how useful that cue is in concert with other available cues. This includes a discussion of the salience of cues as a function of distance, as well as a consideration of the type of information provided by these cues.

We then consider how depth perception is impacted when cues become impoverished, absent, or put in conflict with one another. While impoverishment and absence can occur in the natural environment (e.g., due to darkness or fog), cue conflict is most often the artifact of synthetic displays created by artists, psychologists, and display technologists such as simulator engineers. For as diligently as we strive to create high-fidelity simulators that recreate the perceptual experiences of the operational flight environment, limitations in current display and image generation technologies inevitably result in anomalous depth cues that then conflict with the veridical cues. Our challenge, then, is to understand the normative use of these cues in order to design simulation systems that minimize

Presented at the IMAGE 2013 Conference
Scottsdale, Arizona – June 2013

artifactual impact on control task performance, both in the simulator, and when transferring skills to the actual flight environment. Let us begin, then, with a consideration of how we humans recreate our 3-D world from the light that enters our eyes.

VISUAL CUES

There are multiple taxonomies that can be used to classify visual cues. For the purposes of this discussion, we will classify cues into two categories: static cues (i.e., cues that can be used to derive depth information without any visual motion); and dynamic cues (i.e., cues that are specified by the changing visual characteristics that result from relative motion between the scene features and the observer.) As we describe in our other paper presented at this conference (Kaiser & Sweet, 2013), this division of visual depth cues can be linked to separate neurological pathways in the human brain. Ultimately, however, the human visual system transparently integrates both sets of cues into a seamless construction of the surrounding environment.

Static Cues

Traditionally, static depth cues have been further divided into two categories: physiological, and pictorial. The *physiological* cues include: lens accommodation (i.e., the distance at which the lenses in the eyes are focused); convergence (i.e., the inward rotation of the eyes); and binocular disparity (i.e., the difference in images between the eyes).

Pictorial cues are taken from the 2-D retinal image. A list of these cues includes:

- Retinal Image Size (relative size)
- Height in Field
- Linear Perspective
- Occlusion
- Texture Gradient
- Shadows and Shape from Shading
- Aerial Perspective

Artists have utilized pictorial depth cues since the Renaissance, but 19th century psychologists were the first to systematically study the cues' structure and efficacy.

Dynamic Cues (Optic Flow)

Of course, psychologists of the 19th century recognized that human observers were capable of creating and perceiving motion. However, moving stimuli were considered a complication for the visual system rather than a unique source of information – motion was derived, not perceived. It wasn't until James Gibson's pioneering work in the 1940s that psychologists seriously considered the depth information provided by motion. Although geometrically related, perceptual psychologists tend to differentiate between depth information resulting from observers' movements perpendicular to their line of sight (i.e., horizontal or vertical motion parallax) from the depth information revealed by movement along their line of sight (i.e., radial flow). Both are subclasses of the optic flow generated when observers move about their environment. Nevertheless, our visual system employs different receptor pathways to process these two kinds of motion. (Optic flow can also contain motion components that result from rotation of the observer's head. However, although these components may complicate flow perception, they themselves do not provide any depth information.)

Motion Parallax: In principle, the depth information provided by motion parallax is similar in content to that provided by binocular disparity. Both reveal depth information via differences in image distances between inter-object boundaries when the eye-point is displaced. However, binocular disparity contains only the inter-object depth information that can be extracted from the static disparity provided by the two discrete images sampled by our left and right eyes. In contrast, motion parallax reflects continuous dynamic transformations of inter-object boundaries. Although stereo images can be (and have been) created by sampling two "frames" of motion parallax (i.e., capturing a second image when the eye-point has moved a distance equal to the inter-ocular separation), motion parallax captures a much larger (and continuous) range of eye-points.

While we tend to emphasize the fact that motion parallax provides useful depth information because the extent of motion is directly related to depth, it is important to note another essential cue provided by features in the scene that *do not* exhibit any motion parallax. Objects that are at great distances (e.g., near the horizon) provide an inertial frame to orient the observer.

Radial Outflow: The second component of optical flow is radial flow (or expansion). A forward-looking observer's locomotion creates image motion that radiates outward from his/her line of sight (i.e., his/her track vector). The radial velocity of an object's image depends on both its distance from this focus of expansion and its distance from the observer. For example, if two objects are at the same distance from the observer, the image of the one further from the observer's track vector will have a greater radial velocity. Likewise, if two objects have the same initial angle relative to the observer's track vector but lie at different distances, the closer object will have the greater radial image velocity. Thus, a metric depth map can be recovered from the objects' radial flow rates relative to the objects' bearings (i.e., angles relative to the track vector). In fact, if the observer is moving at a constant velocity, the bearing angle divided by its rate of change can inform the observer of the amount of time until the object will pass his or her viewing plane (Kaiser & Mowafy, 1993). (Unfortunately for collision avoidance, an object directly along the track vector will not have radial flow, so time-to-collision is specified only by the object's image expansion – a far less salient source of information.)

Sensitivity

In comparing binocular disparity and motion parallax, we noted that motion parallax holds two informational advantages. First, motion parallax captures the dynamic boundary-distance transformations rather than two static disparity samples. Second, and perhaps more significant, binocular disparity is constrained to a single level of disparity – that created by the distance between our two eyes (i.e., the inter-ocular distance, or IOD). Simple geometry dictates that the disparity caused by our IOD decreases as the comparison objects become more distant. Thus, there are distance limits on our functional stereopsis. Similarly, the neuromuscular information provided by accommodation and convergence falls off even more quickly with distance. Conversely, aerial perspective is useful only at relatively large distances, while the utility of occlusion, image size, and shadowing are not impacted by distance.

Nagata (1991) provided an informative meta-analysis of the depth-cue psychophysical literature and plotted out which depth cues are useful as a function of distance. (By "useful" we mean that human observers are sensitive to the cues, i.e., that they are "supra-threshold" at these distances.) Cutting and Vishton (1995) expanded on Nagata's analysis by introducing the notion that humans divide their spatial environment into three regions: 1) Personal Space, in which we tend to manipulate objects with our hands and perform fine motor tasks; 2) Action Space, in which we walk, run, and perform gross motor tasks; and 3) Vista Space, in which we orient to the larger environment and plan paths and routes. Cutting and Vishton then superimposed these regions onto a consolidation of the Nagata plots to determine which depth cues were likely to play prominent roles in each space (Fig. 1).

This approach helped psychologists move beyond the notion of determining which cue (or cues) played the dominant role in depth perception. It led toward a more nuanced understanding that our visual system is adaptive and dynamic, and will exploit those cues that are most functional in the space of current interest.

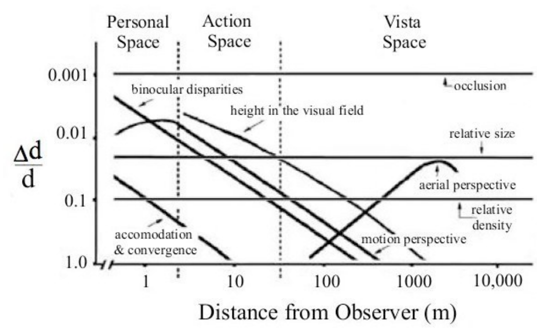

Figure 1. Sensitivity to depth cues. Meta-analysis of psychophysical data by Nagata (1991) maps out the difference in depth (Δd/d) required for depth cues to be supra-threshold as a function of distance from the observer. These sensitivities were then mapped into a three-space model by Cutting & Vishton (1995).

It should be noted that the distance ranges associated with the three regions in Fig. 1 were developed for humans in a "natural" environment, unaided by vehicles or other technologies. Under such conditions, a person's "Action Space" – the region that can be reached in ~6-8 seconds of locomotion or the distance an object can be accurately thrown – is constrained to ~30 m. When humans are asked to shoot guns, or to drive automobiles or fly aircraft, their functional Action Space expands much further, and the depth cues that inform them in this space will likewise shift to those that are useful at these greater distances.

Essentially, one's Action Space for locomotion will be defined by the "look ahead" needed to plan and execute impending control actions, and thus will be a function of vehicle speed and dynamics.

<u>Information</u>

In addition to varying in strength and reliability at varying distances, different cues provide different kinds, or "levels," of information. Occlusion, for example, is a highly reliable cue and functions equally well at near and far distances, but it provides only ordinal information. That is, it definitely specifies that Object B is behind Object A but, as an isolated cue, gives no information regarding the distance of either object. Other cues, such as texture gradient and image size, provide relative depth information (e.g., the distance from the viewer to Object B is twice as large as the distance to Object A, but neither can specify whether those distances are 5 and 10 m, or 20 and 40 m).

The human visual system seamlessly integrates the information and constraints from the various cues to derive an optimized depth-map solution (Landy, Maloney, Johnston, & Young, 1995). Ambiguities and uncertainties arise only when the cues are insufficient to specify a unique solution, or when cues conflict with one another. Cue conflict rarely occurs in natural vision, but (as we will discuss) can be a common artifact of visual displays.

VISUAL SYSTEM TECHNOLOGY

When considering how to simulate these depth cues, and the contributions of different parts of a simulator visual system, we can separate the cues into the functions of the visual system components: the image generator, and the display.

<u>Image Generator</u>

As shown in Table 1, the image generator correctly portrays most of the pictorial cues, as well as optic flow resulting from vehicle motion. Relative size, height in field, linear perspective, occlusion, and aerial perspective (through fog and shaders) are easily achieved. Shadows and shading can be obtained, although it requires more programming to add them to a particular application such as aerial refueling. The methods used to generate *texture gradients* in image generators can cause some artifacts, as will be discussed later.

<table>
<tr><th colspan="2">Image Generator Depth Cues</th></tr>
<tr><td>Type</td><td>Cue</td></tr>
<tr><td rowspan="7">Pictorial</td><td>Retinal Image Size</td></tr>
<tr><td>Height in Field</td></tr>
<tr><td>Linear Perspective</td></tr>
<tr><td>Occlusion</td></tr>
<tr><td>Texture Gradient</td></tr>
<tr><td>Shadows</td></tr>
<tr><td>Aerial Perspective</td></tr>
<tr><td rowspan="2">Optic Flow (from Vehicle Motion)</td><td>Motion Parallax</td></tr>
<tr><td>Radial Outflow</td></tr>
</table>

Table 1: Depth Cues provided by the Image Generator. Most cues are rendered correctly; texture gradients are not veridical, and shadows, when provided, are typically 'simple'.

<u>Display</u>

There are numerous methods used to display visual system imagery. Particular distinctions in display characteristics that will be made are 1) collimated versus real-image, 2) stereoscopic versus non-stereoscopic, and 3) head-tracked versus non head-tracked. We will discuss those depth cues that are a function of the display type: accommodation, convergence, binocular disparity, and head-related motion parallax. A summary of what cues are provided, as a function of display type, is shown in Table 2.

Accommodation is fixed for any type of display. With collimated displays, accommodation is fixed at the collimation distance provided by the optics (typically optical 'infinity', 50 ft or more). With real-image displays, accommodation is fixed at the depth of the display surface. In our personal lives, we frequently experience accommodation mismatch with pictorial displays, such as when watching television or a movie. As discussed above, accommodation is a relatively weak depth cue, and the inability to provide veridical accommodation is likely not to be a major issue in flight simulation.

The second cue we will consider is convergence – the rotational movement of the eyes inward so that both point at a proximal object of interest. In non-stereo displays, convergence will match accommodation. In a real-image display, the eyes will converge to the display surface; in a collimated display, they will converge to the collimation distance. In a stereo display, convergence will occur to the depth of the object. This places the convergence cue in potential conflict with the accommodation cue, which will still be specifying either the display surface (for a real-image display) or optical infinity (for a collimated display).

The last physiological cue, binocular disparity, is only available through a stereo display. It is generally provided correctly as long as certain boundary conditions are not violated (as, for example, where the stereo cue specifies an object between the observer and the screen, but the rendering is truncated at the screen's edge). As noted, binocular disparity is a relatively strong cue related to forming depth judgments. However, mismatches between accommodation and convergence can make it impossible to achieve clear binocular vision (fusion of the images), and/or can create discomfort.

Motion parallax induced by head movement has approximately the same level of sensitivity that binocular disparity has. Head-related motion parallax can be provided only via a head-tracked display.

Type	Real-Image				Collimated			
Stereo	☒		☑		☒		☑	
Head-tracked	☒	☑	☒	☑	☒	☑	☒	☑
Accommodation								
Convergence								
Binocular Disparity								
Head-related Motion Parallax								

Table 2: Ability of visual system display to provide depth cues, as a function of display type and functions. ☒ indicates a characteristic (i.e. stereo, head-tracking) is absent, ☑ indicates a characteristic is present. Dark grey shaded regions indicate that the cue is not present in this display combination; light green shaded regions indicate that the cue is present.

POTENTIAL PROBLEMS

Several particular examples of flight operations in which issues with providing adequate cueing environments have been observed are reviewed here.

<u>Aircraft Stabilization</u>

Unlike real-image displays, distant scene features in a collimated display do not transform with observer head movement. Thus, a collimated display (like a natural scene) provides an important spatial orientation cue. In flight operations, pilots often use distant features (e.g., the horizon) to inform their control movements. The primary author well remembers her experience learning to fly a helicopter – her instructor repeatedly directed her to look at the horizon (although her natural instincts were to look at the ground since it posed a FAR greater threat than the horizon). In fact, looking at the horizon provides very good attitude information, which is the first 'control loop' the pilot 'closes' in a helicopter. The lack of stability of these distant features due to small head movements can cause problems with a real-image display.

A relatively close real-image display was developed for one of the cabs on the Vertical Motion Simulator at NASA Ames. When one of the test pilots was evaluating this display, he was quite certain that the simulated helicopter dynamics had been changed, specifically that it was under-damped. After some investigation, it was determined that the math model for the helicopter was identical to what the test pilot had been accustomed to flying in a cab with collimated displays. This led to a study in which the collimating and real-image displays were compared[1] (Chung, Kaiser, Sweet & Lewis, 2003). The task in the study was to maintain a hover. The primary performance difference between the real-image and collimated displays was superior pitch attitude and velocity (fore-aft and side-to-side) control with the collimated display. Improved pitch attitude control was likely supported by the image stability provided by the collimated display.

The specific aircraft dynamics are likely to be a factor in the suitability of a real-image display. While many helicopters have sophisticated

[1] Field-of-view and display resolution were also manipulated.

control stabilization systems, the 'raw' helicopter dynamics are neutrally stable (at best), making access to accurate orientation information important for the pilot to achieve stabilization. In normal operations, airplanes tend to have good static and dynamic stability; this will likely lead to less dependence on visual orientation cues.

It would be expected that simulated operations of statically stable aircraft would be less affected by a real-image display (relative to a collimated display) than simulated operations of unstable aircraft (e.g., helicopters with unaugmented control laws). Training for the purpose of establishing or renewing skill-based behaviors are likely aided by providing image stability through collimation.

Studies of manual-control performance using real-image and collimated displays under whole-body vibration indicate that manual control tracking performance was superior with a collimated display in comparison to a real-image display (McLeod & Griffin, 1990; Wilson, 1974). This effect was attributed to the stabilization of the retinal image location relative to translational head movements. Thus, in flight simulations with motion, collimation will potentially provide better performance.

Close Contact

A significant number of flight operations are performed in close proximity to other aircraft, terrain, structures, and natural obstacles. Providing a simulated visual scene with sufficient cues, while minimizing cue conflict, can be challenging. The two ends of the close-proximity operations spectrum will be considered – cases where there is little relative motion between one's aircraft and the visual features of the proximal object(s), and cases where there is significant own-vehicle motion in relatively close contact to visual scene elements.

Low Relative Motion: Helicopter taxi/hover and landing, formation flight, and aerial refueling are all examples of operational tasks in which there is little or low relative motion between some or all of the near-scene elements. In these cases, the pilot flying an actual aircraft would derive significant depth information from both binocular disparity and head-related motion parallax. When these cues are not present, some interesting cue conflicts arise. Specifically, the lack of these cues appears to create conflict with pictorial cues.

Researchers at the Air Force Research Laboratories in Mesa, AZ confirmed differences in both size and velocity perception related to display type (Pierce & Geri, 1998). For an aerial formation task, they determined that the perceived relative size of objects was smaller with a real-image display[2] than with a collimated display. Similarly, velocities were judged to be lower in the real-image display than the collimated display. In this case, the lead aircraft was rendered at distances ranging from 500 to 12,000 ft. The targets rendered on the real-image display appeared to be approximately 20% smaller than on the collimated display. (This finding resulted in a recommendation to magnify the target aircraft to match perceived image size with a collimated display, resulting in more detail being perceptible on the formation aircraft.)

Velocity perception was also studied. In a laboratory study using moving arrays of bright dots in an otherwise dark field, perceived velocity was lower with the real-image displays: approximately 12% for a display distance of 0.5 m, and approximately 5% for a display distance of 1.2 m. However, when typical simulated ground textures were used as a stimulus in a constant-altitude flight task, there was little difference in velocity perception associated with the display type.

If a real-image display is associated with misperception of the size of distant objects, perhaps collimated displays can create a misperception of near objects in the absence of stereo disparity or head-related motion parallax. The primary author had the opportunity to 'fly' in a level-D Sikorsky S-76 simulator. A notable feature was that when on the ground, or hovering/taxiing, the height never seemed 'right'. Another notable characteristic of this display was related to size perception – even known-size objects, at close range, appeared much larger. For example, a 3-inch aircraft tie down anchor appeared to be the size of a dinner plate.

A recent study (Lloyd & Nigus, 2012) compared collimated and real-image displays in the presence/absence of stereo and head tracking for an aerial refueling boom operator training simulation. The real-image display was set at a distance of 1.4 m; the collimation distance was set

[2] The real-image display was viewed from a distance of 28 inches.

to the distance of the operator from the nozzle/receptacle at contact (20 m). Inter-pupillary distance (IPD) was set to be 64 mm, the average for military personnel.

Operators were asked to judge the distance between the nozzle and receptacle. Distance estimates were significantly better with the stereoscopic/collimated display than with the non-stereoscopic/non-collimated display; most of this difference could be attributed to the stereo component. Head tracking provided limited improvement, although Lloyd (2011) has reported that when encouraged to move their heads, operators typically indicated that they didn't tend to move their heads when performing the task in the aircraft. Additionally, operators reported better 'geometric stability' with the stereoscopic/collimated display than the other combinations. Consistent with the literature on accommodation/convergence conflict, the real-image stereo display was reported to be less comfortable than the other display conditions. The difference in distance between the real-image display and the convergence point in this case was 18.4 m, a discrepancy equivalent to 0.66 diopters. Shibata, Kim, Hoffman, and Banks (2011) performed a study to relate uncoupled accommodation and convergence cues to viewer comfort when using a stereoscopic display. Although the collimated stereoscopic display described in the boom operator study was well within the "zone of comfort" defined by Shibata, the real-image stereo display was not within this zone. (Banks, Read, Allison, and Watt [2012] provides an excellent reference for the use of stereoscopy.)

Binocular disparity and head-related motion parallax provide similar levels of saliency, although specific aspects of head tracking can limit its practical implementation. Observers are extremely sensitive to latency in a head-tracked system; research has shown that observers can detect head-tracking latencies as low as 17 ms (Adelstein, Lee, & Ellis, 2003). It has also been demonstrated that head-tracking latencies will affect the way in which an observer interacts with the displays. In the 1990s, one possible cockpit configuration considered for the NASA High Speed Research (HSR) Program would eliminate forward-looking cockpit windows. This would improve aerodynamics and eliminate the need for an articulated nose for landing[3]. A test-bed was developed to test this concept for taxi, incorporating a head-tracked display system to allow the pilot look-around capability with a limited field-of-regard (Kaiser, 1998). It was observed that pilots greatly limited their head movements with this system; the only time they moved their heads to look around was when the vehicle was not in motion. It is likely that the ~200 ms latency inherent in this system made it very difficult to disambiguate vehicle-related scene movement from artifactual head-related scene movement.

In 2004, Carmel Applied Technologies reported on the development of a Landing Signal Enlisted (LSE) training system that incorporated both stereo and head-tracking (Holmes & Franz, 2004). An interesting finding is that although two very strong depth cues (binocular disparity and head-related motion parallaz) were available, it was felt that the baseline configuration did not provide sufficient depth cues in comparison to actual flight deck operation. The solution to this was to employ 'hyper-stereo', by moving the eye-points to a much greater displacement than the human IPD. Although the degree of hyper-stereo was not defined, it was noted that some viewers were unable to fuse the stereo images into a single image. This would imply that the hyper-stereo configuration exceeded the zone of clear single binocular vision (ZCSBV) been determined in the optometric research community (Saladin & Sheedy, 1978). At a viewing distance of 50 feet, this would imply a hyper-stereo IPD displacement of over 3 feet, which is also the primary author's recollection from discussion following the talk.

Head tracking was considered to be an important feature in this display due to the need for the LSE to have a complete 360° field-of-regard for situational awareness, and to adjust his position to maintain eye contact with the pilot flying. Although latency for this configuration was not reported, the technologies commonly used at the time of implementation (2004) would suggest that excessive latency was present. Such latencies reduce the capability of head-tracked displays to

[3] A 'droop snoot' was incorporated in both the Concorde and Russian Tupolev Tu-144 supersonic transports to allow pilots to see the runway during the required high angle-of-attack landing approaches.

provide the veridical depth cue from tightly coupled head-related motion parallax.

High Relative Motion: Takeoffs and landings, and nap-of-the-earth/terrain-following flight, are examples of flight operations with significant visual motion (depending, of course, on the aircraft; a 747 pilot experiences far less visual motion on landing and takeoff than a T-38 pilot).

Just as the static details in a simulated scene are difficult to match to reality, dynamic qualities of a simulated visual scene also do not veridically match the experience of viewing a real scene. Spatiotemporal aliasing and motion-induced blur can degrade the effectiveness of a simulator visual scene for certain types of operations.

Spatiotemporal aliasing can negatively affect the perception of motion in a visual scene (Watson, Ahumada & Farrell, 1986; Sweet, Stone, Liston and Hebert, 2008; Sweet & Kato, 2012). When a simulated visual scene is sampled (updated) at an insufficient rate, the motion can appear jerky or even incoherent. For a given image motion, this effect is worsened with increasing levels of scene detail. Spatiotemporal aliasing can also create the appearance of a double exposure in portions of the image.

As part of a study examining the effect of motion-cueing on simulated autorotation performance in the Blackhawk helicopter, visual cueing was studied to examine the possible role of spatiotemporal aliasing (Dearing et al., 2001). In the autorotation, the pilot pulls into a very nose-high attitude (like a quick-stop) to reduce forward velocity, before leveling in order to touch down. When the helicopter is in the nose-high position, the only visual cues are available through the 'chin' window.

In the study, three levels of ground texture were applied in the database: a coarse, medium, and fine texture, each differing from the next by a factor of two. The texture sizes were chosen such that the finest texture displayed some level of visible spatio-temporal aliasing during the nose-up portion of the maneuver. As had been anticipated, the control of rate-of-descent during the nose-up condition was best with the 'medium' texture, and worst with the 'fine' texture that exhibited spatio-temporal aliasing.

Another artifact that occurs with high image motion is motion-induced blur (Sweet & Hebert, 2007). The amount of blur is proportional to the image motion, and can result in a loss of detail. This can negatively impact detection and identification of discrete targets (such as an aircraft) in the visual scene. Motion-induced blur can also be reduced by increasing the update rate (Sweet & Kato, 2012), as well as through shuttering to reduce the persistence of the image.

<u>Flight Path Estimation/Off-Site Landing</u>

As discussed previously, locating the center-of-expansion in the visual scene provides flight path information to the pilot. This is particularly useful when landing; student pilots are taught how to detect this feature, and to adjust the path to put the center of expansion on the runway threshold. In addition to these expansion cues, with experience, pilots also learn to identify the pictorial cues provided by the linear perspective of the runway and surround. However, there is evidence that the pictorial cues are not sufficient to inform flight path guidance (Perrone, 1983) and perform the landing flare (Mulder, Pleijsant, van der Vaart, & Wieringen, 2000. When pilots do not have sufficient cues to detect the center of expansion, flight path estimation is compromised.

In typical flight simulation databases, a great level of detail is provided in the airport environment, particularly for runway thresholds and heliports. Pictorial/perspective cues with significant visual detail and texture in the landing zone afford sufficient cues to judge the touchdown location. In contrast, due to texture memory limitations for the entire database, off-site landing areas (impromptu, pilot-chosen) not specifically designed for landing frequently do not have great level of detail and hence fail to provide sufficient visual cues for landing operations.

Another challenging helicopter operation is performing a rooftop landing. One factor is the difficulty in visually judging the flight path—unless the center of expansion is located ON the rooftop, expansion cues are difficult to find. In the real world, this difficulty is mitigated by the fact that the visual environment typically has eye-limiting details available on the landing surface. In a flight simulator, it is difficult to incorporate sufficient detail (or provide sufficient resolution) to perceive the center of expansion, even on the rooftop itself. Training for military operations involving landings on unprepared sites would likely benefit from having high levels of database detail available for all potential landing locations.

TASK/OPERATION –BASED VISUAL SYSTEM SELECTION

As detailed in the preceding discussion, specific choices for simulator visual system characteristics can affect its suitability for particular types of operations and tasks.

Development/Renewal of Manual Flying Skills

This operational requirement is best met with a collimated display because of image stability; this allows the pilot to have an inertially fixed frame of reference to judge vehicle orientation[4]. Intrinsic aircraft dynamics, as well as the degree of control augmentation for stabilization, will likely affect the importance of providing image stability. Thus, less stable/more maneuverable aircraft simulations (i.e. helicopters, fighters) would likely benefit more from collimation.

Close Contact/Low Relative Motion

Operations with low relative motion and close contact to scene features (e.g., helicopter taxi/hover and landing, formation flight, aerial refueling, NVG scanning) can benefit from stereoscopic displays. Care must be taken to avoid excessive conflicts between accommodation and convergence. This is difficult to achieve with a real-image display due to the large variation in scene content depth. A collimated stereo display can provide comfortable viewing of scene features that range from very near (2 m) to infinitely far.

Helmet-Mounted Displays (HMDs) typically have some level of collimation, and are a good candidate for stereoscopic presentation. Although HMDs are typically driven using head-tracking information, any latency or noise in the head position and orientation measurement compromises the intrinsic image stability associated with fixed collimating displays. Head tracking *is* potentially another useful cue in close contact/low relative velocity situations, but only when implemented with extremely low latencies. The relative importance of binocular disparity and head-related motion parallax are likely to be very heavily dependent on the particular characteristics of the task (e.g., scene features, actions required).

Close Contact/High Relative Motion

Effective perception of smooth motion is likely important in the training and execution of operations with high relative motion and close contact to scene features (e.g., terrain following, nap-of-the-earth, helicopter autorotation). Spatiotemporal aliasing (STA) can degrade the perception of smooth motion (Watson, et al., 1986). Care should be taken in developing the textures used for these situations, and should be evaluated to determine whether there is a compelling sense of motion. Reducing scene/texture detail will reduce the saliency of STA. Increasing update rate will also reduce STA, with the added benefit of reducing motion-induced blur (MIB). While shuttering will decrease the appearance of MIB, it will not improve STA; in fact, by reducing the blur, shuttering could increase the saliency of STA..

Flight Path Estimation/Off-Site Landing

Scene detail contributes to the ability of a pilot to detect the focus-of-expansion to determine the flight path for landing. While scene cueing in the airport environment is typically fairly detailed, practical limitations on image generator capabilities, particularly memory for run-time database loading, can make it difficult to provide sufficient detail to enable effective off-site landings throughout a database. For off-site landings, particularly with significant elevation variation (such as a roof-top landing), attention should be paid to providing sufficient detail to potential landing sites, particularly in the surrounding visual scene.

CONCLUSIONS

There is a paucity of research literature to guide selection of simulator visual system requirements for particular aircraft types and operational tasks. However, consideration of existent operational research findings in conjunction with the body of knowledge in perceptual psychology can provide insight to guide the design of simulator visual systems with an eye towards developing the most cost-effective system for a particular set of operational tasks.

[4] In aircraft, vehicle orientation is intrinsically tied to the generation of forces and moments that produce rotation and displacement of the vehicle. Thus, perception and control of vehicle orientation is a primary piloting skill that must be mastered to enable effective vehicle control.

The conclusions drawn from this sparse body of literature should be taken as starting points for investigation, not as a validated design reference. A significant amount of research should be done to develop comprehensive guidance for visual system characteristics selection. Particular areas of need are:

- Research to determine the extent to which pilots move their heads during eyes-out operations in actual flight (i.e., how much head-related motion parallax *do* pilots experience when performing a task using external visual references?).
- Research to determine the effectiveness of collimated for near-field (~ 3 m), low relative velocity operations (cf. the Lloyd & Nigus, 2012 study used collimated stereo at a distance of 20 m).
- Research to determine the relative performance of collimated and real-image displays in manual flying skill acquisition and retention.
- Research to determine the effect of update rate on eyes-out tasks with high relative motion (e.g., nap-of-the-earth, landing, significant maneuvering).
- Research to determine the relationship between latency and the perception of head-related motion parallax.

REFERENCES

Adelstein, B. D., Lee, T. G., and Ellis, S. R. (2003). Head tracking latency in virtual environments: psychophysics and a model. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting – 2003,* 2083-2087.

Banks, M. S., Read, J. R., Allison R. S., and Watt, S. J. (2012). Stereoscopy and the human visual system. *SMPTE Mot. Image Journal*, *121*(4), 24-43.

Chung, W. W. Y, Sweet, B. T., Kaiser, M. K., and Lewis, E. (2003). Visual cueing effects investigation for a hover task. *AIAA Modeling and Simulation Technologies Conference and Exhibit,* AIAA 2003-5524.

Cutting, J. E., and Vishton, P. M. (1995). Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth. In W. Epstein and S. Rogers (Eds.), *Handbook of Perception and Cognition: Perception of Space and Motion (*pp. 69-117). Academic Press, San Diego, CA.

Dearing, M., Schroeder, J., Sweet, B, and & Kaiser, M. (2001). Effects of visual texture, grids, and platform motion on unpowered helicopter landing. *AIAA Modeling and Simulation Technologies Conference and Exhibit,* A01-37387.

Holmes, B. J., and Franz, T. (2004). Vertical flight deck training system. *Proceedings of the IMAGE 2004 Conference,* The IMAGE Society.

Kaiser, M. K. (1998). The surface operations research and evaluation vehicle (SOREV): A testbed for HSCT taxi issues. *World Aviation Conference,* Paper 985558.

Kaiser, M. K. and Mowafy, L. (1993). Optical specification of time-to-passage: observers' sensitivity to global tau. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1028-1040.

Kaiser, M. K. and Sweet, B. T. (2013). Relating standardized visual perception measures to simulator visual system performance. *Proceedings of the IMAGE 2013 Conference,* The IMAGE Society.

Landy, M. S., Maloney, L. T., Johnston, E. B., and Young, M. (1995). "Measurement and modeling of depth cue combination: In defense of weak fusion". *Vision Research, 35*, 389-412.

Lloyd, C. J., and Nigus, S. G. (2012). Effect of stereopsis, collimation and head tracking on air refueling boom operator performance. *Proceedings of the IMAGE 2012 Conference*, The IMAGE Society.

Lloyd, C. J. (2011). Effects of spatial resolution and antialiasing on stereoacuity and comfort. *AIAA Modeling and Simulation Technologies Conference*, AIAA 2011-4572.

Mcleod, R. W., and Griffin, M. J. (1990). Effects of whole-body vibration waveform and display collimation on the performance of a complex manual control task. *Aviation, Space, and Environmental Medicine. 61*(3), 211-219.

Mulder, M., Pleijsant, J., van der Vaart, H., and Wieringen, P. (2000). The effects of pictorial detail on the timing of the landing flare: results of a visual simulation experiment. *The International Journal of Aviation Psychology. 10*(3), 291-315.

Nagata, S. (1991). How to reinforce perception of depth in single two-dimensional pictures. In , S. Ellis, M. Kaiser, and A. Grunwald (Eds.), *Communication in Virtual and Real Environments* (pp. 527-545). Taylor & Francis, London.

Perrone, J. (1983). Visual slant misperception and the 'black hole' landing situation. NASA TM 85866.

Pierce, B. J., and Geri, G. A. (1998). The implications of image collimation for flight simulator training. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 42*(20), 1383-1387.

Pierce, B. J., Geri, G. A., and Hitt J. M. III. (1998). Display collimation and the perceived size of flight simulation imagery. AFRL-HE-AZ-TR-1998-0058.

Saladin J. J., and Sheedy, J. E. (1978). Population study of fixation disparity, heterophoria, and vergence. *American Journal of Optometry & Physiological Optics, 55*, 744–750.

Schroeder, J., Dearing, M., Sweet, B., Kaiser, M., and Atencio, A. (2001), Runway texture and grid pattern effects on rate-of-descent perception, *AIAA Modeling and Simulation Technologies Conference and Exhibit*, AIAA 2001-4307.

Shibata, T., Kim, J., Hoffman, D. M., and Banks, M. S. (2011). The zone of comfort: predicting visual discomfort with stereo displays. *Journal of Vision, 11*(8), 1-29.

Sweet B. T., and Hebert, T. M. (2007). The impact of motion-induced blur on out-the-window visual system performance. *Proceedings of the IMAGE 2007 Conference*, The IMAGE Society.

Sweet, B. T., Stone, L. S., Liston, D. B., and Hebert, T. M. (2008). Effects of spatio-temporal aliasing on out-the-window visual systems. *Proceedings of the IMAGE 2008 Conference*, The IMAGE Society.

Sweet, B. T., and Kato, K. J. (2012). 120 Hz: the new 60 for flight simulation?. *Proceedings of the IMAGE 2012 Conference*, The IMAGE Society.

Sweet, B. T., and Kaiser, M. K. (2011). Depth perception, cueing, and control. *AIAA Modeling and Simulation Technologies Conference*, AIAA 2011-6424.

Watson, A. B., Ahumada A. J. Jr., and Farrell, J. E., (1986). Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A, 3*(3), 300-307.

Wilson, R. V. (1974). Display collimation under whole-body vibration, *Human Factors, 16*(2), 186-195.

AUTHOR BIOGRAPHIES

Dr. Barbara Sweet is a Human-Machine Aerospace Technologist in the Human Systems Integration Division at NASA Ames Research Center. She received B.S. and M.S. degrees in Aeronautical and Astronautical Engineering from Purdue University in Indiana in 1982 and 1986, respectively. In 1999, Dr. Sweet received a Ph.D. in Aeronautics and Astronautics from Stanford University, for the development of models of human control behavior that account for perspective scene viewing. Since joining NASA in 1984, she has worked in helicopter handling-qualities research, research simulation facility development and management, and human-factors research. Dr. Sweet is a pilot, with commercial ratings in both airplanes and helicopters, and formerly worked as a flight instructor and instrument flight instructor in airplanes.

Dr. 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 and a Fellow of the Association for Psychological Science, 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. She also holds one patent and has a second patent pending for innovative display technologies.