Spatial Awareness in Synthetic Vision Systems: Using Spatial and Temporal Judgments to Evaluate Texture and Field of View

Matthew L. Bolton and **Ellen J. Bass,** University of Virginia, Charlottesville, Virginia, and **James Raymond Comstock, Jr.,** NASA Langley Research Center, Hampton, Virginia

Objective: This work introduced judgment-based measures of spatial awareness and used them to evaluate terrain textures and fields of view (FOVs) in synthetic vision system (SVS) displays. **Background**: SVSs are cockpit technologies that depict computer-generated views of terrain surrounding an aircraft. In the assessment of textures and FOVs for SVSs, no studies have directly measured the three levels of spatial awareness with respect to terrain: identification of terrain, its relative spatial location, and its relative temporal location. **Methods**: Eighteen pilots made four judgments (relative azimuth angle, distance, height, and abeam time) regarding the location of terrain points displayed in 112 noninteractive 5-s simulations of an SVS head-down display. There were two between-subject variables (texture order and FOV order) and five within-subject variables (texture, FOV, and the terrain point's relative azimuth angle, distance, and height). **Results**: Texture produced significant main and interaction effects for the magnitude of error in the relative angle, distance, height, and abeam time judgments. FOV interaction effects were significant for the directional magnitude of error in the relative distance, height, and abeam time judgments. **Conclusion**: Spatial awareness was best facilitated by the elevation fishnet (EF), photo fishnet (PF), and photo elevation fishnet (PEF) textures. **Application**: This study supports the recommendation that the EF, PF, and PEF textures be further evaluated in future SVS experiments. Additionally, the judgment-based spatial awareness measures used in this experiment could be used to evaluate other display parameters and depth cues in SVSs.

INTRODUCTION

Controlled flight into terrain (CFIT), in which a fully functional aircraft is inadvertently flown into the ground, water, or other terrain obstacle, is the largest source of fatalities in commercial aviation (Boeing Commercial Airplanes, 2006). CFIT accidents are characterized by a loss of situation awareness (SA) in low-level flight and low visibility conditions (Khatwa & Roelen, 1999). By using onboard terrain and obstacle databases and Global Positioning System data, synthetic vision system (SVS) displays give pilots the information necessary to proactively avoid CFIT by creating a synthetic, clear-day view of the world in front of ownship regardless of the actual visibility conditions.

Measuring Spatial Awareness

Spatial awareness (SpA), an aspect of SA (Wickens, 2002a), encompasses the extent to which pilots notice objects in the surrounding environment (Level 1), their understanding of where these objects are relative to ownship (Level 2), and their prediction of these objects' future position (Level 3; Wickens, 2002b). SpA is integral to CFIT prevention because it encompasses a pilot's knowledge about the relative spatial and temporal location of terrain. SVS can help prevent CFIT by enhancing pilot SpA.

Evaluations of SVS have employed a variety of metrics. Performance measures include cross track error (Alexander, Wickens, & Hardy, 2005; Arthur, Prinzel, Kramer, Parrish, & Bailey, 2004; Schnell

Address correspondence to Ellen J. Bass, University of Virginia, Department of Systems and Information Engineering, 151 Engineer's Way, Charlottesville, VA 22904; ejb4n@virginia.edu. ***HUMAN FACTORS,*** Vol. 49, No. 6, December 2007, pp. 961–974. DOI 10.1518/001872007X249848.

& Lemos, 2002), the number of correct identifications made when matching video of actual terrain to SVS displays (Schnell & Lemos, 2002), the reproduction of highlighted terrain points from a blanked SVS display on an out-the-window display (Alexander et al., 2005), and the detection and anticipation of CFIT incidents (Arthur et al., 2004; Uhlarik, Peterson, & Herold, 1998). Subjective awareness measures have also been used in SVS research: Situation Awareness Rating Technique (SART; Hughes & Takallu, 2002), Situation Awareness Subjective Workload Dominance (SA-SWORD; Arthur et al., 2004; Hughes & Takallu, 2002), and terrain awareness (Bailey, Parrish, Arthur, & Norman, 2002; Glaab & Hughes, 2003).

Other research domains have used SpA measures. Yeh (1992) used ordinal distance judgments to assess spatial perception for stereoscopic and perspective displays. Several studies have utilized azimuth and elevation angle measures of the relative position of objects (Alm, Lif, & Öberg, 2003; Barfield & Rosenberg, 1995; Dorighi, Grunwald, & Ellis, 1992; McGreevy & Ellis, 1986). Wells, Venturino, and Osgood (1988) asked participants to replicate, from memory, the position of objects presented in helmet-mounted displays. Barfield, Rosenberg, and Furness (1995) had participants replicate, on maps, the positions of targets shown out the window during flight. Fracker (1990) had participants replicate the display position of enemy aircraft during pauses in simulated flight. Marshak, Kuperman, Ramsey, and Wilson (1987) had participants make judgments about the location of targets shown on map displays during pauses in simulated flight.

None of these studies directly measured all three levels of SpA. In addition, studies that employed general SA measures did not explicitly measure the SpA component, potentially biasing the rating (Carmody, 1994). Additionally, subjective awareness measures, such as SART, SA-SWORD, and terrain awareness may assess confidence in performance rather than awareness. Third, performance measures, such as cross track error, require a relationship between awareness and performance that cannot always be assumed.

Influences on SpA

A variety of SVS display features impact SpA. The two investigated in this work are terrain texture and field of view (FOV).

Terrain texture. Terrain texture (Figure 1) refers to the imagery drawn on the synthetic terrain of SVS displays. SVS displays have used the following base texture concepts:

- fishnet (F): a grid of interlinked 500- × 500-foot squares drawn on a solidly colored terrain;
- elevation (E): distinct bands of color drawn on the terrain representing regularly spaced intervals of terrain elevations corresponding to color schemes used by visual flight rule sectional charts; and
- photo (P): satellite photos of the actual terrain superimposed on the synthetic terrain.

Terrain texture is important to SpA because it facilitates different depth and motion cues. Because SVS displays are two-dimensional (2-D) representations of a 3-D space, only pictorial depth cues (cues that can be represented in a 2-D picture) are relevant. Texture supports the following pictorial depth cues (Goldstein, 2002).

- Familiar size: When an object's size is known by the observer, he or she can determine its relative distance by observing how big it is in the 2-D display.
- Relative size: An object that is the same size as another, but is farther away from the observer, will be smaller in the 2-D display.
- Texture gradient: As objects and patterns that are equally spaced get farther away from the observer, they get closer together in the 2-D display.
- Linear perspective: Parallel lines will converge in the 2-D display as their distance from the observer increases.

The three base texture concepts (fishnet, elevation, and photo) convey different information. The fishnet texture's pattern contributes to familiar size, relative size, texture gradient, and linear perspective: The pilot will know how big each square actually is, contributing to familiar size; the size of the squares will decrease as they get farther away, contributing to relative size; squares will get closer together as they get farther away, contributing to texture gradient; and the parallel lines that result from the grid will converge as they get farther away, contributing to linear perspective.

Elevation texturing facilitates the texture gradient depth cue by applying prerendered terrain shading in addition to its elevation color intervals. This provides a texture gradient that is denser than that of fishnet texture. Additionally, the coded elevation information may facilitate pilot comprehension of terrain height.

The photo texture provides a more natural view of surrounding terrain and contributes to the familiar size, relative size, and texture gradient depth

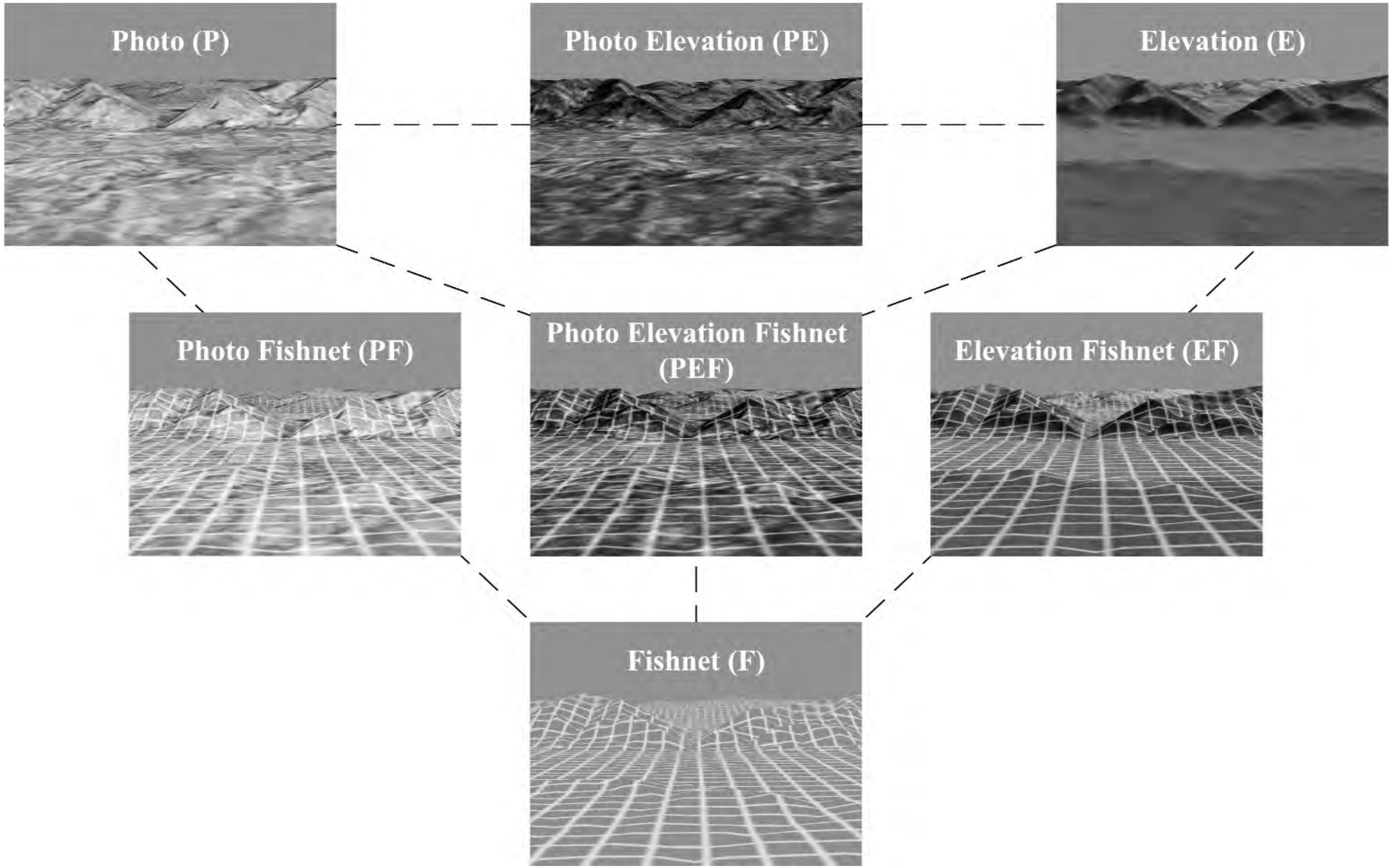

Figure 1. The terrain textures evaluated in the experiment.

cues. The familiar size and relative size depth cues are both achieved through the presentation of familiar terrain features that get smaller as they get farther away. The photo texture has a denser texture gradient than do the fishnet and elevation textures.

All textures contribute to the global optic flow motion cue by texturing the invariant structures in the optical ambient array (Gibson, 1986).

In a comparison of six textures, Schnell and Lemos (2002) found that elevation, photo, and fishnet-like grid textures (a checkerboard pattern) produced the highest percentages of correct identifications for video terrain matching and the least cross track error in flight simulator studies. Takallu, Wong, Bartolone, Hughes, and Glaab (2004) found that photo and elevation texturing produced significantly higher SART scores than did no texturing and fishnet texturing. Hughes and Takallu (2002) found that there were significantly larger SA-SWORD scores associated with elevation and photo textures than with no texture.

Field of view. FOV refers to the angular boundaries of the volume of space represented in the display. FOV is important to SpA because larger FOVs result in a larger geographical region being presented on the display. On displays with larger FOVs, objects of the same physical size and relative position will appear smaller than on displays with smaller FOVs.

SVS simulation tests conducted by Comstock, Glaab, Prinzel, and Elliott (2002) allowed transport pilots to change among FOVs of 30°, 60°, and 90°. Pilots preferred the 30° FOV for large (9.25- × by 8-inch; 23.5- × 20.3-cm) displays. Glaab and Hughes (2003) found that general aviation pilots who were allowed to select among 22.5°, 30°, 60°, and 90° FOVs during flight tests preferred the 60° FOV. Arthur et al. (2004) found that pilots preferred a 30° FOV during approach and a 60° FOV during departure.

Objectives and Hypotheses

This research investigated how SpA was affected by the three leading texture types in all combinations (F, E, P, EF, PF, PE, and PEF; Figure 1) and the two leading FOVs (30° and 60°), using judgments with respect to a terrain point that probed all three levels of SpA. Identifying the terrain point probed Level 1 SpA. Judgments of the relative azimuth angle, distance, and height of the terrain point to ownship probed Level 2 SpA (providing a 3-D perspective of the terrain's location). An abeam time judgment (the time it would take the pilot to fly to the point of closest approach to the terrain point) probed Level 3 SpA.

Because each of the base textures conveys different spatial information, it was hypothesized that SpA would be best facilitated by the combination of all three texture types. This was because each base concept contained unique information that would be preserved as a result of combination. Whereas the fishnet texture facilitates the linear perspective depth cue and has a coarse texture gradient (assisting in course distance and angle estimates), the addition of the elevation or photo textures would provide gradient information between the fishnet's grid lines (helping refine angle and distance estimates). Additionally, the elevation texture contains coded elevation information, not present in the other two, which may facilitate terrain height judgments. Finally, the photo texture presents a more natural and familiar view of the terrain.

Because combining textures has the potential to clutter the displays and thus negate the effectiveness of different depth cues, experimentation was necessary to determine what benefit to SpA, if any, composite textures would have.

METHODS

Participants

Eighteen general aviation pilots participated in the study. All had fewer than 400 hr of flight experience ($M = 157$, $SD = 75$). They were familiar with the out-the-window view from a cockpit but not with SVS displays.

Apparatus

Experiments were run in a windowless, constantly lighted laboratory. Workstations displayed each simulation and collected participant judgments. Simulations depicted SVS head-down displays with the symbology in Figure 2. In simulations, terrain point location was indicated using a yellow inverted cone ($d = 500$ feet, $h = 500$ feet) rendered as part of the SVS environment. The tip of the cone intersected the terrain at the terrain point. All simulations showed SVS displays in straight, level flight at 127 knots. They were displayed as 5-s, 836 × 728 pixel, 30 frames/s, Windows Media Video (WMV) files. Custom software played the WMV files and collected participant responses (Bolton, Bass, & Comstock, 2006).

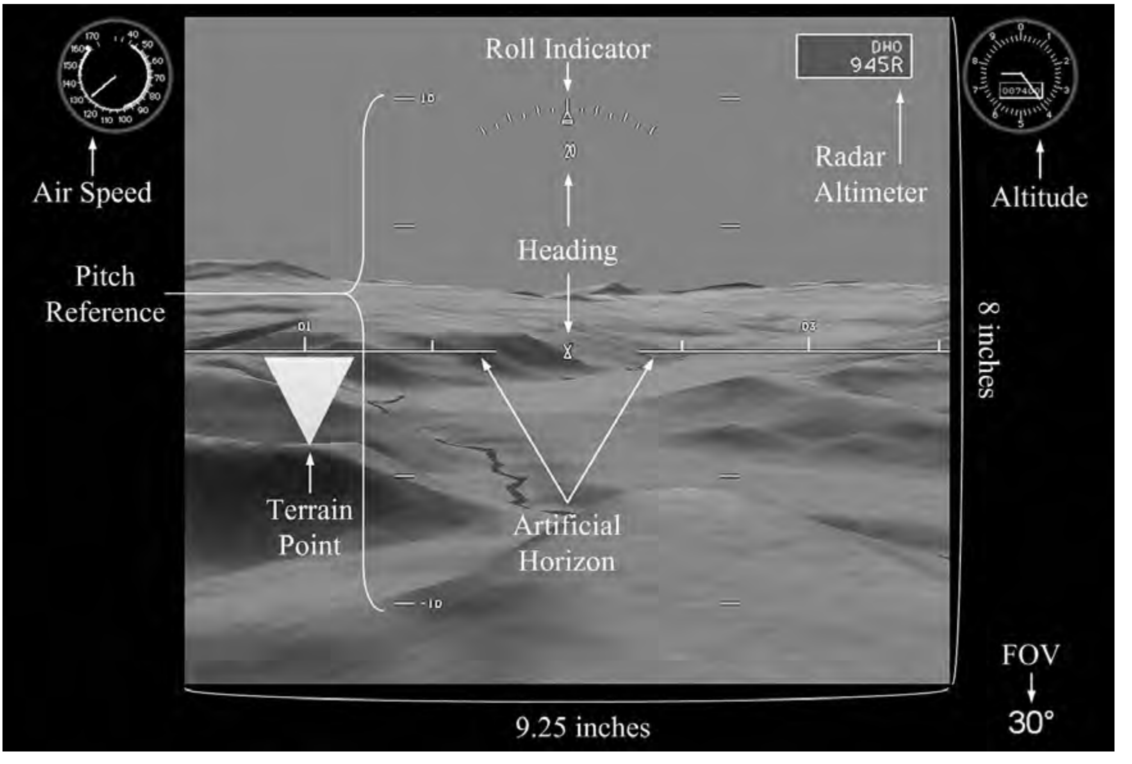

Figure 2. The SVS display and symbology used in the experiment (labels added). SVS displays were presented to participants at an eye distance of approximately 30 inches (76.2 cm) and a horizontal visual angle of approximately 18° (display dimensions of 9.25 × 8 inches = 23.5 × 20.3 cm).

Independent Variables

Within-subject variables. The five within-subject variables were texture, FOV, and three scenario geometry variables: the relative azimuth angle (R_a), relative distance (R_d), and relative height (R_h) of the terrain point to ownship. Seven textures were used in the experiment: three base textures (F, P, and E) and four derivative textures (EF, PF, PE, and PEF). Two FOVs (30° and 60°) were used in the SVS displays. The terrain point location varied based on its relative position at the end of a scenario by changing the three scenario geometry variables, each with two levels (Table 1).

Between-subject variables. There were two between-subject variables: FOV order and texture order. A participant either saw all of the 30° FOV trials first or all of the 60° FOV trials first. Thus, FOV order had two levels: 30° FOV first or 60° FOV first.

Textures used to derive other textures always appeared before their derivatives in order to avoid

TABLE 1: Terrain Point Relative Position (Scenario Geometry) Levels

Independent Variable	Range	Distribution	Level
R_a (relative azimuth angle)	[0°, 6.5°] [8.5°, 15°]	N(μ = 3.75, σ = 1.25) N(μ = 11.25, σ = 1.25)	Small Large
R_d (relative distance)	[1 nmi, 3.25 nmi] [3.75 nmi, 6 nmi]	N(μ = 2.25, σ = 0.417) N(μ = 4.75, σ = 0.417)	Near Far
R_h (relative height)	[–1000 ft, –100 ft] [100 ft, 1000 ft]	U(–1000, –100) U(100, 1000)	Below Above

Note. nmi = nautical mile, ft = feet.

complications associated with presenting derivative textures before participants had seen their bases. Each participant saw two of the base textures, the combination of them, the third texture, and the rest of the combinations. Three texture orders were created so that no base texture was introduced in more than one ordered position: {P, E, PE, F, PF, EF, PEF}, {E, F, EF, P, PE, PF, PEF}, and {F, P, PF, E, EF, PE, PEF}.

Dependent Measures

Eight dependent measures were calculated from the four judgment values (relative angle in degrees, relative distance in nautical miles, relative height in feet, and abeam time in seconds) using the three judgment tasks (Table 2). Two dependent measures (one for directional error and one for absolute error) resulted from each judgment value. Each directional error term represented both the direction and magnitude of the judgment error. When a participant overestimated a judgment, the corresponding directional error term was positive. When the participant underestimated a judgment, it was negative. Absolute error terms represented the magnitude of the judgment error and were calculated as the absolute value of their corresponding directional error term.

Procedure

Each experimental session lasted fewer than 4 hr. The participants completed consent forms and were briefed about the experiment. For each trial, participants viewed a 5-s simulation of an SVS head-down display in flight (Figure 2). At the end of the 5 s, the simulation paused for 1 s, and the screen was cleared.

For each trial, using the interface in Figure 3, participants made four judgments based on the relative position of the terrain point: relative (azimuth) angle, relative distance, relative height, and abeam time. For the relative distance and angle judgments, participants placed a yellow X in the upper left section of the display corresponding to the lateral location of the terrain point relative to the aircraft. Values for relative angle (in degrees) and distance (in nautical miles) were displayed next to the X.

For the relative height judgment, the participant placed a yellow X on a vertical scale in the upper right of the display corresponding to the relative height of the terrain point. The relative height was displayed in feet next to the X as it was moved. For the abeam time judgment, participants entered the time judgments in minutes and seconds using the keyboard. To support this time judgment, a yellow dot on the relative distance and angle judgment collection interface indicated the abeam point (point of closest approach) based on the relative distance and angle judgment.

Participants were asked to perform these tasks as quickly and accurately as possible. For training trials, participants were given feedback relating to the accuracy of their judgments. (See Bolton et al., 2006, for more information about the experimental apparatus.)

Each participant experienced 112 experimental trials (7 textures × 2 FOVs × 2 R_a levels × 2 R_d levels × 2 R_h levels = 112) and 72 training trials, each with a unique terrain configuration. For the first texture experienced for each FOV, there were 12 training trials. For the other textures, for each FOV, there were 4 training trials. Thus, each

TABLE 2: Dependent Measure Formulations

Terrain Point Position Measure	Actual Value	Judgment Value	Directional Error Dependent Measure	Absolute Error Dependent Measure
Azimuth angle	A_a	A_j	$A_e = \begin{cases} A_j - A_a & \text{if } A_a > 0 \\ -A_j + A_a & \text{otherwise} \end{cases}$	$\lvert A_e \rvert$
Distance	D_a	D_j	$D_e = D_j - D_a$	$\lvert D_e \rvert$
Height	H_a	H_j	$H_e = \begin{cases} H_j - H_a & \text{if } H_a > 0 \\ -H_j + H_a & \text{otherwise} \end{cases}$	$\lvert H_e \rvert$
Abeam time	τ_a	τ_j	$\tau_e = \tau_j - \tau_a$	$\lvert \tau_e \rvert$

Note. All terrain point position measures were made relative to ownship. A_a and A_j were measured relative to the aircraft's vector of displacement with angles in the clockwise direction being positive and angles in the counterclockwise direction being negative. H_a and H_j were measured relative to the aircraft's height with positive heights above the aircraft and negative heights below.

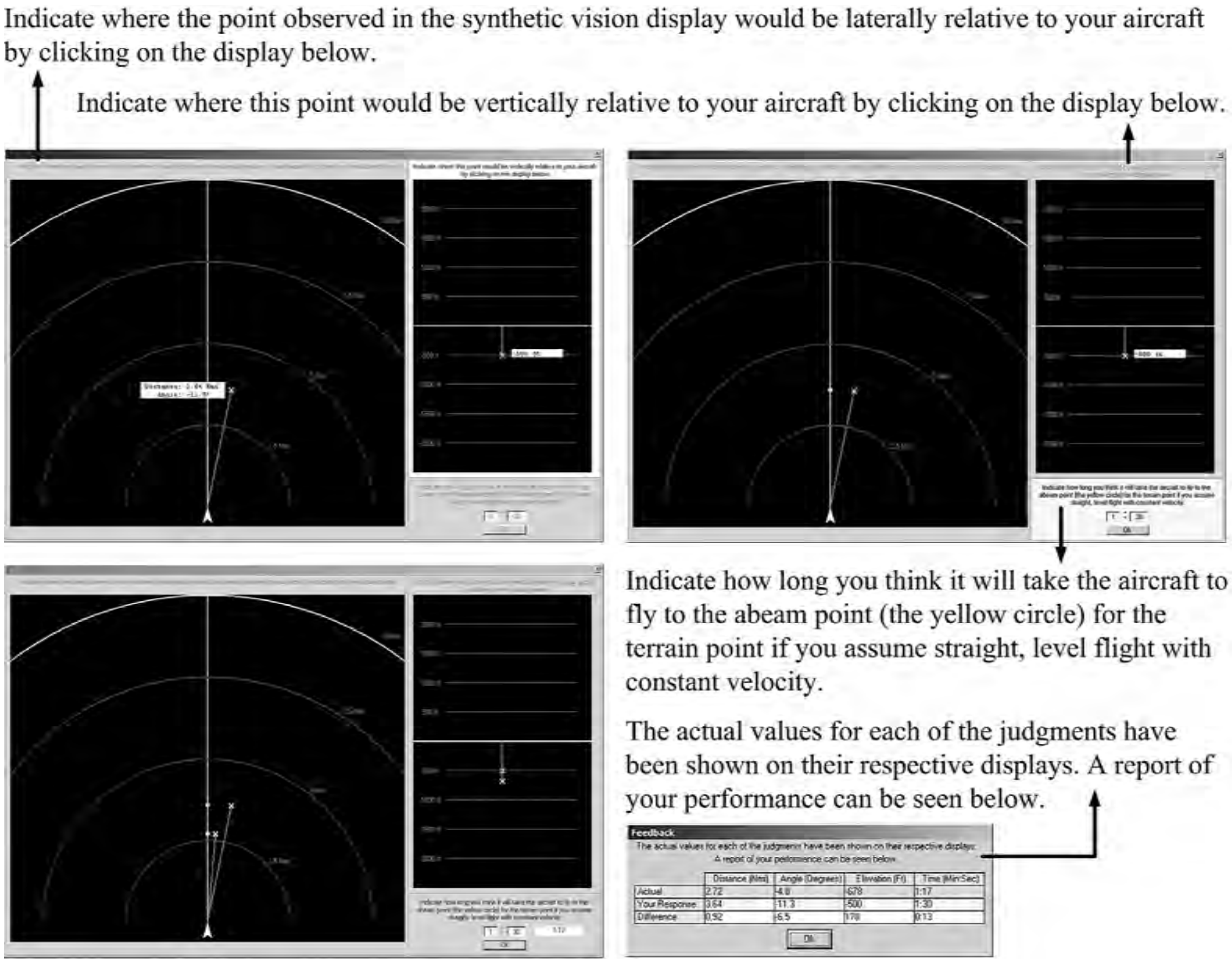

Figure 3. The judgment collection interface in multiple modes of operation (clockwise from the upper left): the relative distance, angle, and height judgments; the abeam time judgment; training feedback on the judgment collection interface; numerical training feedback shown concurrently with feedback on the judgment collection interface.

participant saw a total of 72 training trials, resulting in a total of 184 trials. The order in which the eight scenario geometry levels were presented was unique for each texture and FOV combination. Thus there were 14 scenario geometry presentation orders. Scenario geometry variable levels were counterbalanced between presentation orders so that each combination of variable levels appeared in each ordered slot twice and directly followed every other combination twice.

On completion of all of the trials for each texture for each FOV, subjective demand (Taylor, 1990), awareness (Glaab & Hughes, 2003), and clutter (Bailey, Kramer, & Prinzel, 2006) ratings were collected. After all of the trials for a FOV were completed, participants made SA-SWORD pairwise comparisons (Vidulich & Hughes, 1991) for all of the textures seen with that FOV (for more detail see Bolton & Bass, 2007).

Experimental Design and Data Analysis

The experiment employed a repeated measures design with 18 participants. Three participants were randomly assigned to each of the six combinations of the between-subject variables (2 FOV orders × 3 texture orders = 6).

This study was concerned with the effect of texture and FOV on spatial awareness. However, because spatial biases had the potential to impact spatial awareness between levels of the scenario geometry variables (Bolton & Bass, in press; Wickens, 2002b), this study was concerned not only with the texture and FOV main effect and their interaction but also with their two-way interactions with scenario geometry variables. The effect of these within-subject factors and the between-subject factors on the dependent measures were assessed using repeated measures MANOVAs (one for directional error and one for absolute error) with Wilks's lambda (Brace, Kemp, & Snelgar, 2003).

Effects found to be significant in the MANOVA were evaluated against each of the dependent measures using a univariate repeated measures ANOVA. When a Mauchly's test of sphericity was violated ($p < .05$) a Greenhouse-Geisser epsilon correction factor was applied (Brace et al., 2003). For variables with more than two levels, post hoc analyses were used to identify significant differences between levels. When sphericity was violated, a Bonferroni multiple comparison method was used (Stevens, 2002). A Tukey's HSD was used otherwise (Stevens, 2002).

RESULTS

This section presents results in which a priori significance levels were $\alpha = .01$. The MANOVA results for the main effect and two-way interaction model for both absolute and directional error revealed that FOV order, $\Lambda = .55$, $F(4, 9) = 1.86$, $p = .201$, $\Lambda = .47$, $F(4, 9) = 2.51$, $p = .116$, texture order, $\Lambda = .47$, $F(8, 18) = 1.02$, $p = .457$, $\Lambda = .46$, $F(8, 18) = 1.08$, $p = .418$, and their interaction, $\Lambda = .17$, $F(8, 18) = 3.18$, $p = .020$; $\Lambda = .35$, $F(8, 18) = 1.53$, $p = .215$, were not significant for absolute or directional error, respectively. Thus, all three were eliminated from all subsequent models.

The resulting MANOVA analyses revealed that for absolute error, all of the following were significant: Texture, $\Lambda = .63$, $F(24, 346.58) = 2.07$, $p = .003$, Texture × R_a, $\Lambda = .51$, $F(24, 346.58) = 3.03$, $p = .001$, Texture × R_d, $\Lambda = .63$, $F(24, 346.58) = 2.02$, $p = .004$, Texture × R_h, $\Lambda = .63$, $F(24, 346.58) = 2.07$, $p = .003$.

For the directional error MANOVA, all of the following were significant: Texture × R_a, $\Lambda = .49$, $F(24, 346.58) = 3.31$, $p = .001$, Texture × R_d, $\Lambda = .62$, $F(24, 346.58) = 2.12$, $p = .002$, Texture × R_h, $\Lambda = .49$, $F(24, 346.58) = 3.29$, $p = .001$, and FOV × R_a, $\Lambda = .25$, $F(4, 14) = 10.44$, $p = .001$.

ANOVA results for these variables appear in Tables 3 and 4.

Texture

There was a significant main effect of texture for $|D_e|$ (Table 3). A Tukey's HSD (Figure 4a) revealed that significantly less distance error was committed with PEF than with F.

A Tukey's HSD was performed for each of the Texture × Scenario Geometry interactions that were significant for an absolute error dependent measure (Table 3). This revealed significant differences between textures for all of the following: $|A_e|$ when R_a was large (Figure 4b); $|H_e|$ for both R_a levels (Figures 4c and 4d); and $|\tau_e|$ when R_d was far (Figure 4e). No significant differences were observed in mean $|H_e|$ between textures for the Texture × R_h interaction.

Examining the main and interaction effect results together revealed that there were three textures (EF, PF, and PEF) among the homogeneous

TABLE 3: Results of the Repeated Measures ANOVA for Absolute Error Dependent Measures

Independent Variable	$\lvert A_e \rvert$	$\lvert D_e \rvert$	$\lvert H_e \rvert$	$\lvert \tau_e \rvert$
Texture	$F(3.74, 65.87) = 3.21^a$ $p = .021$	$F(6, 102) = 3.43$ $p = .004$*	$F(6, 102) = 1.39$ $p = .225$	$F(2.15, 65.87) = 2.95^b$ $p = .061$
Texture × R_a	$F(6, 102) = 3.25$ $p = .006$*	$F(6, 102) = 1.45$ $p = .205$	$F(6, 102) = 6.45$ $p < .001$*	$F(6, 102) = 2.87$ $p = .013$
Texture × R_d	$F(6, 102) = 2.07$ $p = .064$	$F(6, 102) = 2.62$ $p = .021$	$F(3.4, 74.36) = 1.73^c$ $p = .166$	$F(6, 102) = 3.23$ $p = .006$*
Texture × R_h	$F(2.82, 61.6) = 1.16^d$ $p = .335$	$F(6, 102) = 2.48$ $p = .028$	$F(6, 102) = 3.88$ $p = .002$*	$F(6, 102) = 2.25$ $p = .045$

Note. Superscripts indicate that a Greenhouse-Geisser ε correction was applied to the degrees of freedom because of a violation of sphericity: $^aW = .1$, $\chi^2(20) = 33.48$, $p = .033$, $\varepsilon = .62$; $^bW = .01$, $\chi^2(20) = 62.54$, $p < .001$, $\varepsilon = .36$; $^cW = .07$, $\chi^2(20) = 39.25$, $p = .007$, $\varepsilon = .57$; $^dW = .04$, $\chi^2(20) = 49.43$, $p < .01$, $\varepsilon = .47$.

*$p < .01$.

subsets that produced the least average absolute error across the five error categories (Figure 4).

There were also significant differences found between textures for directional error (Table 4). Post hoc analyses revealed that for the Texture × R_a interaction, participants significantly underestimated distances more for PEF as compared with EF, PF, and PE when R_a was small (Figure 5a); underestimated heights more for PEF than for PE when R_a was small (Figure 5b); underestimated heights more for E and P compared with F, EF, and PEF, and for PE compared with EF and PEF when R_a was large (Figure 5c); and underestimated times significantly more for PEF than for F, E, EF, PF, and EF when R_a was small (Figure 5e). For the Texture × R_d interaction, participants underestimated heights more for E and P than for EF when R_d was far (Figure 5d). There were no significant differences in mean H_e between textures for the Texture × R_h interaction.

FOV

FOV was not a significant main effect for any of the error terms. However, the FOV × R_a interaction was significant for D_e, H_e, and τ_e (Table 4). A Tukey's HSD revealed that for both levels of R_a, participants underestimated distances for the 30° FOV and overestimated them for the 60° FOV (Figures 6a and 6b). When R_a was large, they underestimated relative heights for both FOVs, but more so for the 60° FOV (Figure 6c). When R_a was small, participants underestimated abeam times for the 30° FOV and overestimated them for the 60° FOV (Figure 6d).

DISCUSSION

The primary objective of this experiment was to apply a new judgment-based means of measuring SpA to determine how different textures and FOVs facilitate SpA in SVS displays. This is discussed for both texture and FOV.

Texture

We hypothesized that because the three base textures conveyed different spatial information, their combinations would result in increased SpA. For the absolute error terms for which texture or a Texture × Scenario Geometry parameter was significant, the data do suggest that combinations of the base texture concepts enhance SpA. In all cases, the EF, PF, and PEF textures were in the homogeneous subset of textures that produced the smallest magnitudes in error (Figure 4).

Whereas the main effects analysis (Figure 4a) differentiated between textures at only the second level of SpA (significant differences in $\lvert D_e \rvert$), the EF, PF, and PEF textures were shown to facilitate the least error for Level 2 ($\lvert A_e \rvert$ and $\lvert H_e \rvert$) and Level 3 ($\lvert \tau_e \rvert$) SpA judgments in the interaction effects (Figures 4b–4e). This suggests that SpA is best facilitated by one of these textures. As the PEF texture was always the last texture seen by participants, there is a potential confound associated with its low error values. Future work should investigate whether the low mean errors observed for the PEF texture were a result of observation order.

A common feature in the E, P, and PE textures is the terrain form they convey via shading. In

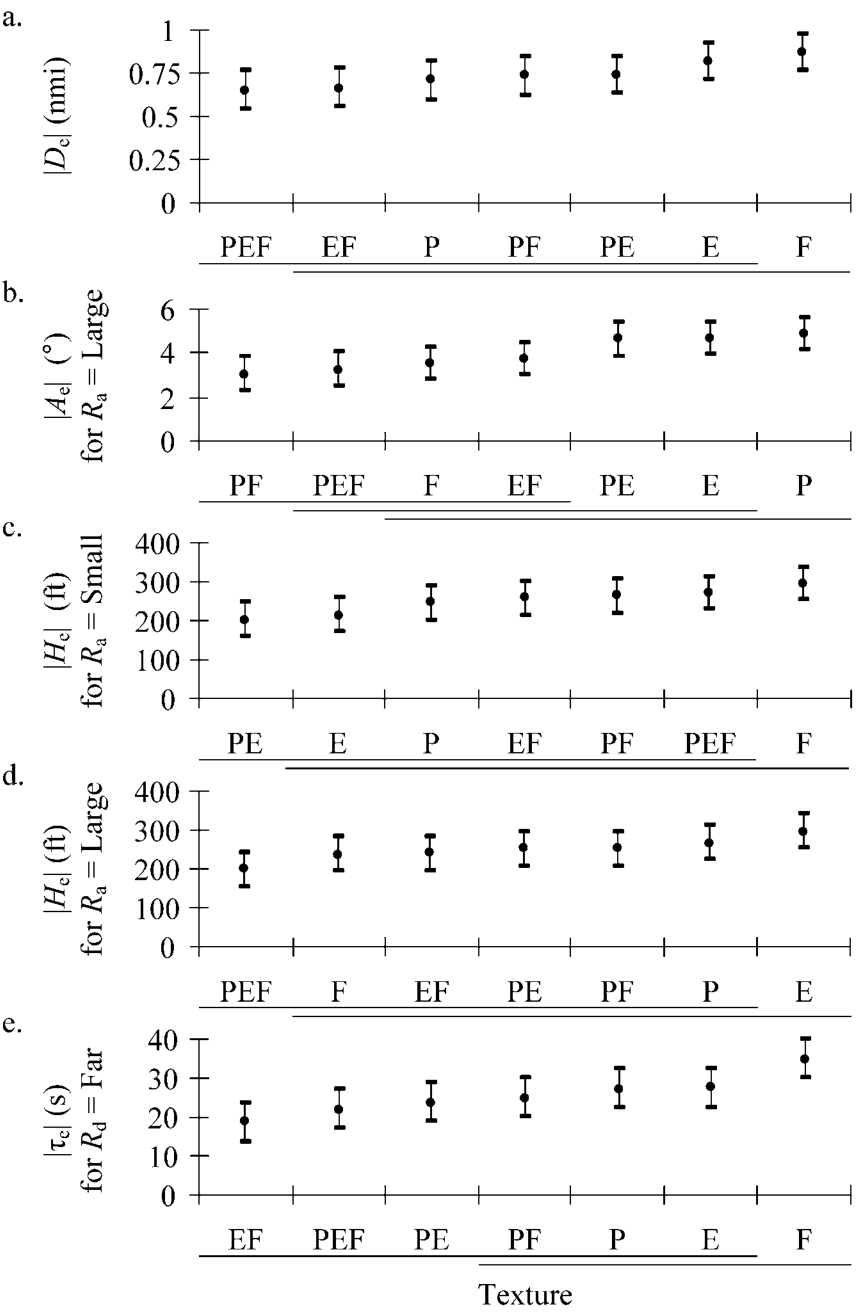

Figure 4. Post hoc analyses results for the texture main effect and Texture × Scenario Geometry interactions that were significant for absolute error dependent measures for which significant differences were observed between textures. Filled circles indicate means (sorted in ascending order by error magnitude). Bars around means indicate Tukey's intervals from a Tukey's HSD with α = .01. Lines under textures indicate homogeneous subsets (in which no significant differences were observed between textures for the given graph's category of error) as indicated by the Tukey's HSD; ft = feet, nmi = nautical miles.

TABLE 4: Results of the Repeated Measures ANOVA for Directional Error Dependent Measures

Independent Variable	A_e	D_e	H_e	τ_e
Texture × R_a	$F(6, 102) = 1.33$ $p = .249$	$F(6, 102) = 3.77$ $p = .002$*	$F(6, 102) = 9.25$ $p < .001$*	$F(6,102) = 3.49$ $p = .004$*
Texture × R_d	$F(6, 102) = 1.42$ $p = .212$	$F(6, 102) = 2.36$ $p = .035$	$F(6, 102) = 3.18$ $p = .007$*	$F(6, 102) = 2.85$ $p = .013$
Texture × R_h	$F(2.99, 74.77) = 1.57^a$ $p = .207$	$F(6, 102) = 2.05$ $p = .065$	$F(3.71, 74.77) = 10.47^b$ $p < 0.001$*	$F(6, 102) = 2.13$ $p = .056$
FOV × R_a	$F(1, 17) = 4.8$ $p = .043$	$F(1, 17) = 8.43$ $p = .010$*	$F(1, 17) = 20.38$ $p < .001$*	$F(1, 17) = 17.39$ $p < .001$*

Note. Superscripts indicate that a Greenhouse-Geisser ε correction was applied to the degrees of freedom because of a violation of sphericity: $^aW = .05$, $\chi^2(20) = 45.35$, $p = .001$, $\varepsilon = .5$; $^bW = .11$, $\chi^2(20) = 32.43$, $p < .044$, $\varepsilon = .62$.
*$p < .01$.

order to convey terrain form without dynamic shading (see Schnell & Lemos, 2002), E, P, and PE had terrain shading represented on the texture itself. The F texture did not. Thus, results comparable to those observed for EF, PF, and PEF may be achievable with a texture that uses a fishnet pattern over a shaded, solidly colored terrain.

If only the three base textures (F, E, and P) are considered, the data suggest that SpA is best facilitated by P as it was among the textures that produced the least absolute error for more main and interaction effects as compared with the F and E textures (Figure 4). This is interesting because the textures that resulted in the least judgment error (EF, PF, and PEF) were all derivatives of F, not P. Thus, benefits afforded by the P texture are dominated by the benefits achieved through the combination of E and F or P and F.

The directional error terms with significant interaction effects, to which texture contributed, convey insight into how the choice of texture biases spatial perception in SVS displays. It is clear that any overestimation is potentially hazardous as it constitutes an underestimation of the terrain's threat. However, it is not clear to what extent underestimation is undesirable. Thus, a heuristic would need to be developed in order to address what magnitude of underestimation is detrimental and what the relative importance of error is between judgments.

FOV

FOV did not contribute to main or interaction effects for the absolute error dependent measures. Thus, it is not clear which FOV provided superior SpA. For the significant differences seen in directional error between FOVs (Figure 6), varying degrees of over- and underestimations were observed. This suggests a need to investigate FOVs between 30° and 60°. However, no value judgment can be made without a heuristic for assessing the relative importance of directional error.

Conclusions

Because SVS displays are designed to address CFIT, they must ensure that they support accurate pilot SpA. This work introduced new measures of SpA that allowed the accuracy of pilot SpA to be evaluated for textures and FOVs in SVS displays. In general, SpA was best supported by the EF, PF, and PEF textures. Either a 30° or 60° FOV could be used. A 30° FOV would promote underestimation of relative distances and abeam times. A 60° FOV would promote overestimation of relative distances and abeam times and underestimation of relative heights. However, given the artificiality of this study's procedure, these results should be replicated in higher fidelity simulations before being used to make any design decisions.

Although the procedure used 5-s videos, this time period should be investigated. Even though the experiment produced significant results, future work should go into investigating how long videos should be displayed in order to best inform pilot instrumentation design.

The method used for assessing Level 3 SpA (the abeam time judgment) is partially confounded by the Level 2 SpA judgments. Thus, future experiments may want to investigate alternative ways of assessing Level 3 SpA.

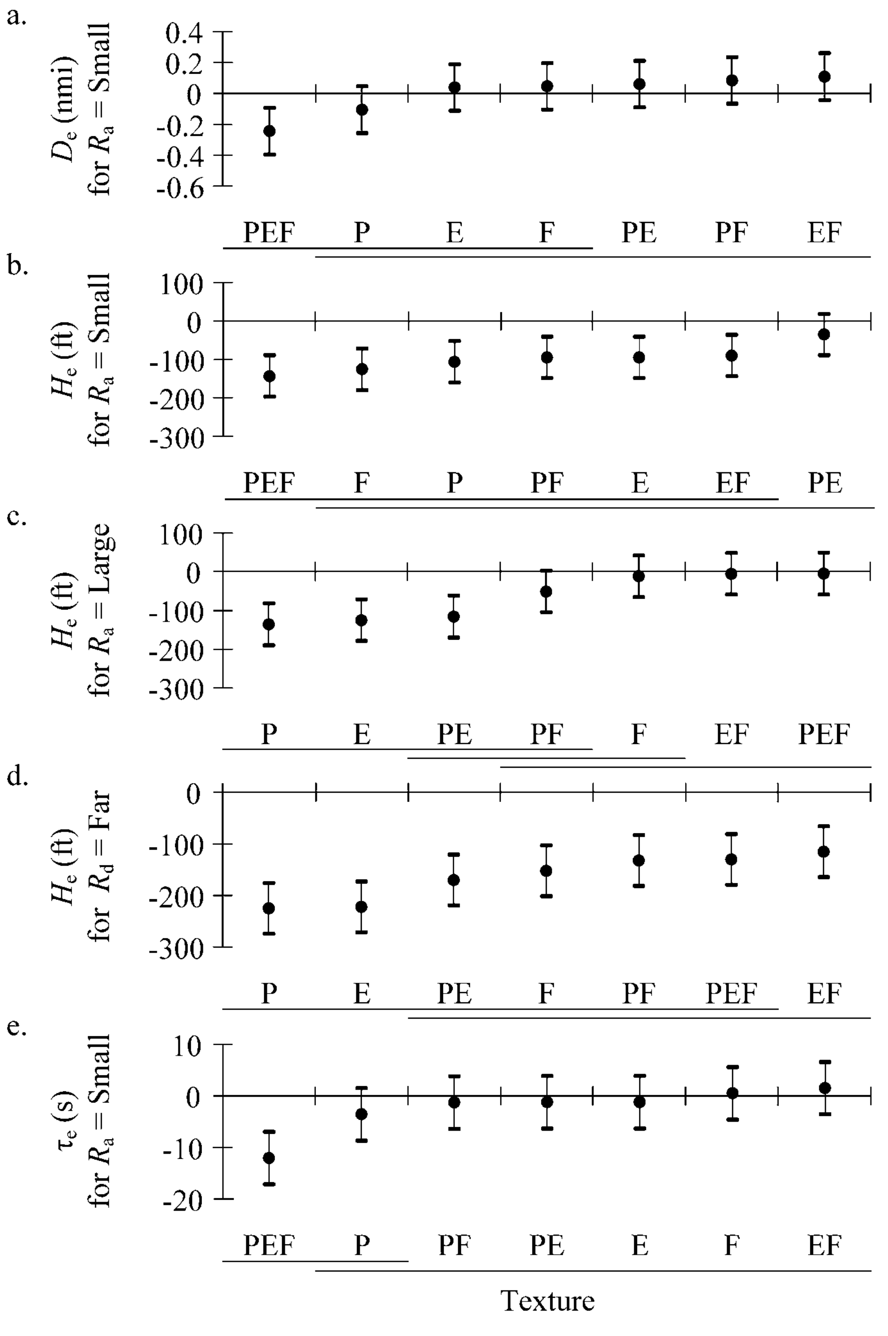

Figure 5. Post hoc analyses results for the Texture × Scenario Geometry interactions that were significant for directional error dependent measures for which significant differences were observed between textures. Filled circles indicate means. Bars around means indicate Tukey's intervals from a Tukey's HSD with $\alpha = .01$. Lines under textures indicate homogeneous subsets (in which no significant differences were observed between textures for the given graph's category of error) as indicated by the Tukey's HSD; ft = feet, nmi = nautical miles.

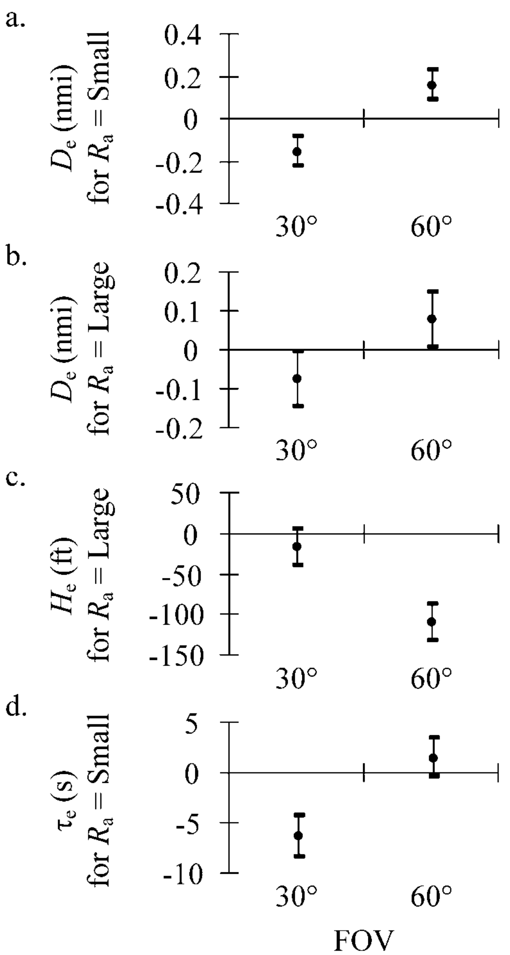

Figure 6. Post hoc analyses results for the direction error terms for which significant differences were found between the FOVs in the FOV × Angle interaction. Filled circles indicate means. Bars around means indicate Tukey's intervals from a Tukey's HSD with $\alpha = .01$; ft = feet, nmi = nautical miles.

A variety of other display parameters and pictorial depth cues could also impact pilot SpA in SVS. Such display parameters include additional FOVs, textures (see Schnell & Lemos, 2002), display sizes (see Comstock et al., 2002), and instrumentation. Depth cues could also be evaluated. Atmospheric perspective is a cue wherein an object that is farther away appears less sharp than a closer one because of particulates in the air (Goldstein, 2002). This depth cue could be added to SVS so that its impact on SpA could be evaluated. Familiar objects such as buildings, trees, towers, and runways could be added to the SVS display to facilitate relative size and familiar size depth cues.

There are also parameters in the textures themselves that may affect SpA. As the three best textures from this experiment (EF, PF, and PEF) contained grid patterns, research could investigate what grid sizes result in the most accurate spatial judgments. Other parameters of interest are the thickness of the fishnet gridlines, the colors used to represent elevation intervals, and texture resolution.

Although the focus of this research was centered on SpA with SVS displays, there are other potential applications. Although the results collected will likely prove useful for perspective aviation displays, given that the relevance of depth cues is often dependent on the relative distance of the objects (Goldstein, 2002), the methodology used for measuring SpA could be useful in any domain in which accurate SpA is important. Such domains include driving, robot control systems, air traffic control, virtual environments, and unmanned air vehicles.

ACKNOWLEDGMENTS

This work was supported in part by the National Institute of Aerospace and the National Aeronautics and Space Administration. The authors would like to thank those making suggestions for this study, including Trey Arthur, Daniel Burdette, Sheila Conway, Stephanie Guerlain, Regina Johns, Lynda Kramer, Kara Latorella, Neil O'Connor, Hayes Press, Valerie Plyler, Lawrence Prinzel, Paul Schutte, Jason Sweeters, Mohammad Takallu, and Steve Williams. The authors also thank the three anonymous reviewers.

REFERENCES

Alexander, A. L., Wickens, C. D., & Hardy, T. J. (2005). Synthetic vision and the primary flight display: The effects of guidance symbology, display size, and field of view. Human Factors, 47, 693–707.

Alm, T., Lif, P., & Öberg, M. (2003). Evaluation of monocular depth cues in 3D aircraft displays. In *Proceedings of the 12th International Symposium on Aviation Psychology* (CD-ROM). Dayton: Ohio State University.

Arthur, J. J., Prinzel, L. J., III, Kramer, L. J., Parrish, R. V., & Bailey, R. E. (2004). *Flight simulator evaluation of synthetic vision display concepts to prevent controlled flight into terrain (CFIT)* (NASA Tech. Rep. NASA/TP-2004-213008). Hampton, VA: NASA Langley Research Center.

Bailey, R. E., Kramer, L. J., & Prinzel, L. J., III. (2006). Crew and display concepts evaluation for synthetic/enhanced vision systems. In *2006 Proceedings of International Society for Optical Engineering Conference on Enhanced and Synthetic Vision* (pp. 62260G.1–62260G.18). Bellingham, WA: International Society for Optical Engineering.

Bailey, R. E., Parrish, R. V., Arthur, J. J., & Norman, R. M. (2002). Flight test evaluation of tactical synthetic vision display concepts in a terrain-challenged operating environment. In J. G. Verly (Ed.),

Enhanced and synthetic vision 2002 (pp. 178–189). Bellingham, WA: International Society for Optical Engineering.
Barfield, W., & Rosenberg, C. (1995). Judgments of azimuth and elevation as a function of monoscopic and binocular depth cues using a perspective display. *Human Factors, 37*, 173–181.
Barfield, W., Rosenberg, C., & Furness, T. A. I. (1995). Situation awareness as a function of frame of reference, computer-graphics eyepoint elevation, and geometric field of view. *International Journal of Aviation Psychology, 5,* 233–256.
Boeing Commercial Airplanes. (2006). *Statistical summary of commercial jet airplane accidents: Worldwide operations 1959–2005.* Retrieved June 13, 2007, from http://www.boeing.com/news/techissues/pdf/statsum.pdf
Bolton, M. L., & Bass, E. J. (2007). Spatial awareness: Comparing judgment-based and subjective measures. In *Proceedings of the 2007 IEEE Conference on Systems, Man and Cybernetics* (CD-ROM). Los Alamitos, CA: Institute of Electrical and Electronics Engineers.
Bolton, M. L., & Bass, E. J. (in press). Using relative position and temporal judgments to identify biases in spatial awareness for synthetic vision systems. *International Journal of Aviation Psychology*.
Bolton, M. L., Bass, E. J., & Comstock, J. R. (2006). A toolset to support the development of spatial and temporal judgment experiments for synthetic vision systems. In *Proceedings of the 2006 IEEE Systems and Information Engineering Design Symposium* (pp. 55–60). Los Alamitos, CA: Institute of Electrical and Electronics Engineers.
Brace, N., Kemp, R., & Snelgar, R. (2003). *SPSS for psychologists: A guide to data analysis using SPSS for Windows*. Mahwah, NJ: Erlbaum.
Carmody, M. A. (1994). *Current issues in the measurement of military aircrew performance: A consideration of the relationship between available metrics and operational concerns* (Naval Air Warfare Center Aircraft Division Tech. Rep. NAWCADWAR-94139-60). Warminster, PA: Naval Air Warfare Center.
Comstock, J. R., Glaab, L. J., Prinzel, L. J., & Elliott, D. M. (2002). Can effective synthetic vision system displays be implemented on limited size display spaces? In *Proceedings of the 11th International Symposium on Aviation Psychology* (CD-ROM). Dayton: Ohio State University.
Dorighi, N. S., Grunwald, A. J., & Ellis, S. R. (1992). Perspective format for a primary flight display and its effect on pilot spatial awareness. In *Proceedings of the IEEE/AIAA 11th Digital Avionics Systems Conference* (pp. 307–312). Los Alamitos, CA: Institute of Electrical and Electronics Engineers.
Fracker, M. L. (1990). Attention gradients in situation awareness. In *Situational awareness in aerospace operations* (AGARD-CP-478, Conference Proceedings #478, pp. 6.1–6.10). Neuilly-sur-Seine, France: NATO Advisory Group for Aerospace Research and Development.
Gibson, J. J. (1986). *The ecological approach to visual perception.* Hillsdale, NJ: Erlbaum.
Glaab, L. J., & Hughes, M. F. (2003). Terrain portrayal for head-down displays flight test. In *Proceedings of the 22nd Digital Avionics Systems Conference* (pp. 9.E.2-1–9.E.2-15). Los Alamitos, CA: Institute of Electrical and Electronics Engineers.
Goldstein, E. B. (2002). *Sensation and perception* (6th ed.). Pacific Grove, CA: Wadsworth.
Hughes, M. F., & Takallu, M. A. (2002). Terrain portrayal for head-down displays experiment. In *Proceedings of the International Advanced Aviation Conference* (pp. G-1–G-12). Anchorage, AK: University of Alaska.
Khatwa, R., & Roelen, A. (1998). An analysis of controlled-flight-into-terrain (CFIT) accidents of commercial operators, 1988 through 1994. *Flight Safety Digest, 17,* 166–212.
Marshak, W. P., Kuperman, G., Ramsey, E. G., & Wilson, D. (1987). Situational awareness in map displays. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 533–535). Santa Monica, CA: Human Factors and Ergonomics Society.
McGreevy, M. W., & Ellis, S. R. (1986). The effect of perspective geometry on judged direction in spatial information instruments. *Human Factors, 28,* 439–456.
Parrish, R. V., Busquests, A. M., Williams, S. P., & Nold, D. E. (1994). *Spatial awareness comparisons between large-screen, integrated pictorial displays and conventional EFIS displays during simulated landing approaches* (NASA Tech. Paper 3467). Hampton, VA: NASA Langley Research Center.
Schnell, T., & Lemos, K. (2002). *Terrain sampling density and texture requirements for synthetic vision systems* (Rockwell Collins Final Rep.). Cedar Rapids, IA: Rockwell Collins.
Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Erlbaum.
Takallu, M. A., Wong, D. T., Bartolone, A. P., Hughes, M. F., & Glaab, L. J. (2004). Interaction between various terrain portrayals and guidance/tunnel symbology concepts for general aviation synthetic vision displays during a low en-route scenario. In *Proceedings of the 23rd Digital Avionics Systems Conference* (pp. A.B.3-1–A.B.3-12). Los Alamitos, CA: Institute of Electrical and Electronics Engineers.
Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. *Situational Awareness in Aerospace Operations (AGARD-CP-478)* (pp. 3/1–3/17). Neuilly-sur-Seine, France: NATO Advisory Group for Aerospace Research and Development.
Uhlarik, J. J., Peterson, D. A., & Herold, J. M. (1998). Terrain awareness using a terrain enhanced primary flight display. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 102–106). Santa Monica, CA: Human Factors and Ergonomics Society.
Vidulich, M. A., & Hughes, E. R. (1991). Testing a subjective metric of situation awareness. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1307–1311). Santa Monica, CA: Human Factors and Ergonomics Society.
Wells, M. J., Venturino, M., & Osgood, R. K. (1988). Using target replacement performance to measure spatial awareness. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1429–1433). Santa Monica, CA: Human Factors and Ergonomics Society.
Wickens, C. D. (2002a). Situation awareness and workload in aviation. *Current Directions in Psychological Science, 11,* 128–133.
Wickens, C. D. (2002b). *Spatial awareness biases* (NASA Tech. Rep. ARL-02-6/NASA-02-4). Moffett Field, CA: NASA Ames Research Center.
Yeh, Y. (1992). Spatial judgments with monoscopic and stereoscopic presentation of perspective displays. *Human Factors, 34,* 583–600.

Matthew L. Bolton is a Ph.D. student in the Department of Systems and Information Engineering at the University of Virginia, where he earned his M.S. in systems engineering in 2006.

Ellen J. Bass is an assistant professor in the Department of Systems and Information Engineering at the University of Virginia. She earned her Ph.D. in systems engineering from the Georgia Institute of Technology in 2002.

James Raymond Comstock, Jr., is the team lead for crew systems in the Crew Systems and Aviation Operations Branch at the NASA Langley Research Center. He earned his Ph.D. in industrial/organizational psychology from Old Dominion University in 1983.

Date received: June 16, 2006
Date accepted: June 9, 2007