Crew and Display Concepts Evaluation for Synthetic / Enhanced Vision Systems

Randall E. Bailey*, Lynda J. Kramer, and Lawrence J. Prinzel III
NASA Langley Research Center, M/S 152, Hampton, VA 23681-0001

ABSTRACT

NASA's Synthetic Vision Systems (SVS) project is developing technologies with practical applications that strive to eliminate low-visibility conditions as a causal factor to civil aircraft accidents and replicate the operational benefits of clear day flight operations, regardless of the actual outside visibility condition. Enhanced Vision System (EVS) technologies are analogous and complementary in many respects to SVS, with the principle difference being that EVS is an imaging sensor presentation, as opposed to a database-derived image. The use of EVS in civil aircraft is projected to increase rapidly as the Federal Aviation Administration recently changed the aircraft operating rules under Part 91, revising the flight visibility requirements for conducting operations to civil airports. Operators conducting straight-in instrument approach procedures may now operate below the published approach minimums when using an approved EVS that shows the required visual references on the pilot's Head-Up Display. An experiment was conducted to evaluate the complementary use of SVS and EVS technologies, specifically focusing on new techniques for integration and/or fusion of synthetic and enhanced vision technologies and crew resource management while operating under the newly adopted FAA rules which provide operating credit for EVS. Overall, the experimental data showed that significant improvements in SA without concomitant increases in workload and display clutter could be provided by the integration and/or fusion of synthetic and enhanced vision technologies for the pilot-flying and the pilot-not-flying.

1. INTRODUCTION

The Integrated Intelligent Flight Deck Technologies (IIFDT) project, under NASA's Aviation Safety Program (AvSP), is comprised of a multi-disciplinary research effort to develop flight deck technologies that mitigate operator-, automation-, and environment-induced hazards. Towards this objective, IIFDT is developing crew/vehicle interface technologies that reduce the risk of pilot error, improve aircraft safety for current and future civilian and military aircraft, and proactively overcome aircraft safety barriers that would otherwise constrain the full realization of the next generation air transportation system (NGATS). Part of this research effort involves the use of synthetic and enhanced vision systems and other interface modalities as enabling technologies to meet these safety challenges.

Synthetic vision is a computer-generated image of the external scene topography that is generated from aircraft attitude, high-precision navigation, and data of the terrain, obstacles, cultural features, and other required flight information. A synthetic vision *system* (SVS) enhances this basic functionality with real-time integrity to ensure the validity of the databases, perform obstacle detection and independent navigation accuracy verification, and provide traffic surveillance. Over the last 5 years, NASA and its industry partners have developed and deployed SVS technologies for commercial and business aircraft which have been shown to provide significant improvements in terrain awareness and reductions in the potential for Controlled-Flight-Into-Terrain incidents / accidents compared to current generation cockpit technologies.[1-4]

An Enhanced Vision System (EVS) (or Enhanced Flight Vision System) is an electronic means to provide a display of the external scene by use of an imaging sensor, such as a Forward-Looking InfraRed (FLIR) or millimeter wave radar. The intended use of EVS mirrors SVS – both strive to eliminate low-visibility conditions as a causal factor to civil aircraft accidents and replicate the operational benefits of clear day flight operations, regardless of the actual outside visibility condition. The methodologies by which this capability is achieved, however, are significantly different. While

*randall.e.bailey@nasa.gov; phone 1 757 864-8682; fax 1 757 864-7793

some may consider the technologies to be competing; they are, in fact, complementary.[5]

SVS, by virtue of being weather-independent and unlimited in field-of-regard, holds many advantages over enhanced vision sensor systems for providing terrain, path, and obstacle awareness, particularly during flight phases, such as approach, which may be obscured by clouds and precipitation of which an EVS sensor cannot penetrate. Recognition of terrain and cultural features may also be improved over an EVS since the display presentation is optimized by the display designer, not the product of the sensor and its environment. Pilot recognition of EVS terrain and cultural features depends upon the reflected, emitted, and / or refracted energy at the spectral frequencies of the EVS and the ability of the pilot to (correctly) interpret this image. Atmospheric effects, time of day, and sensor characteristics can be important factors in the quality of the EVS imagery.

On the other hand, EVS is an imaging sensor which provides a direct view of the vehicle external environment; consequently, EVS is completely independent of the derived aircraft navigation solution and is independent of a database. Very little stands between the EVS image shown to the pilot and the real-world; thus, an EVS pilot gets an extremely high degree of confidence in the system. Under conditions of smoke, haze, and night, a FLIR/EVS provides orders-of-magnitude improvement over the pilot's natural vision; greatly enhancing the pilot's situation awareness and reducing the pilot's workload. The comparison of SVS and EVS in Figure 1 on a night visual meteorological conditions approach into an airfield highlights the similarities and differences in these two technologies.
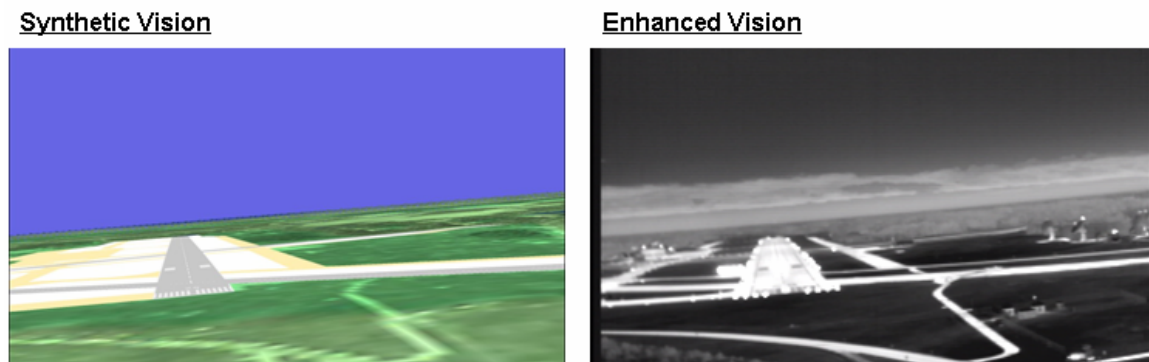

Figure 1: Synthetic Vision and Enhance Vision Comparison

Previous synthetic vision research[6] has shown that a "flight-critical" synthetic vision implementation which uses automated decision aiding functions for object detection and database alignment/navigation error detection produces superior performance to synthetic vision concepts with an EVS inset display. This result formed the motivation for developing Synthetic Vision System enabling technologies for database integrity monitoring, object detection, etc. These enabling technologies are being pursued (e.g., see References 7-8). To date, however, technology for "perfect" object detection and database/navigation error detection does _not_ exist. Further, even if these systems come to fruition, there may still be gaps, such as minimal radar cross-section objects or below-threshold detection error values, which may require independent integrity and error checks.

SVS with EVS inset displays offer one possible method to provide the pilot with information sufficient to perform navigation integrity and obstacle clearance checks. While these concepts are viable, performance and pilot workload[6] suffer in comparison to automated methods to achieve these same capabilities. Other studies have shown similar results.[9] Object detection by pilots was found to be best using a dedicated EVS display that did not include symbology. (However, the presence or absence of symbology was not tested.) From this study and others, the performance of pilots to perform navigation integrity checks and obstacle identification principally depend upon the pilot's visual acuity using the display imagery for object and image recognition[10], such as the resolution and acuity of the sensor; the characteristics of the object; the prominence of the object and surrounding external scene features; display clutter; display size; and display and object color and contrast.

While EVS might improve SVS operations, the converse warrants investigation as well. On January 9, 2004, Section 91.175 of the Federal Aviation Regulations was amended such that operators conducting straight-in instrument approach

procedures (other than Category II or Category III) may now operate below the published Decision Height and Minimum Descent Altitude when using an approved Enhanced Flight Vision System (EFVS) on the pilot's Head-Up Display. This rule change now provides "operational credit" for EVS. As such, EVS operations will become more prevalent.

An experiment was conducted to evaluate the complementary use of SVS and EVS technologies, specifically focusing on new techniques for integration and/or fusion of synthetic and enhanced vision technologies and crew resource management while operating under the newly adopted FAA rules which provide operating credit for EVS. Specifically, the objective of this experiment was to test the utility, acceptability, and usability of integrated/fused enhanced and synthetic vision systems technology concepts in two-crew commercial and business aircraft cockpit for Required Navigation Procedures (RNP)-type approaches.

2. METHODOLOGY

2.1. Subjects

Twenty-four pilots, representing seven airlines and a major cargo carrier, participated in the experiment. All participants had previous experience flying Head-Up Displays (HUDs). The subjects had an average of 1787 hours of HUD flying experience and an average of 13.8 years and 16.2 years of commercial and military flying experience, respectively. EVS experience was not required of the subjects although some pilots were familiar with imaging sensor technology from their prior military flight experience. None of the subjects was currently flying EVS in commercial and business aircraft operations.

2.2. Simulator

The experiment was conducted in the Integration Flight Deck (IFD) simulation facility (see Fig. 2) at NASA Langley Research Center (LaRC). The IFD emulates LaRC's Boeing B-757-200 aircraft and provides researchers with a full-mission simulator capability. The cab is populated with flight instrumentation and pilot controls, including the overhead subsystem panels, to replicate the B-757 aircraft. The collimated out-the-window (OTW) scene is produced by an Evans and Sutherland ESIG 4530 graphics system providing approximately 200 degrees horizontal by 40 degrees vertical field-of-view at 26 pixels per degree.
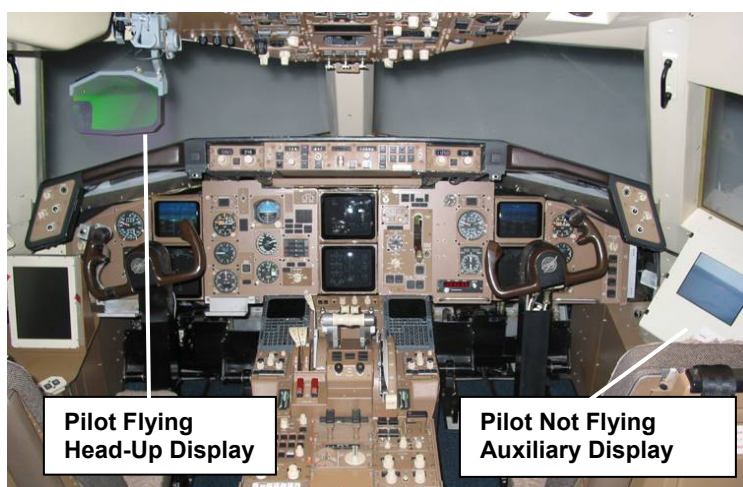

Figure 2. Integration Flight Deck Simulation Facility with HUD and AD.

The pilot participants occupied the left (as the Pilot Flying, PF) and right (as the Pilot Not Flying, PNF) seats. The left seat included an overhead HUD projection unit and the right seat included the installation of an auxiliary display (AD) under the right side window (see Fig. 2).

2.2.1. Head-Up Display

The HUD subtended approximately 32° horizontal by 24° vertical field of view. The HUD presentation was written strictly in raster format from a video source (RS-343) input. The input consisted of a video mix of symbology and computer-generated scene imagery (either EVS or SVS as described in Section 2.7.1). The symbology included "haloing" to ensure that the symbology was highlighted against the scene imagery background. Brightness and contrast controls were provided to the pilot. Also, the pilot had a declutter control, implemented as a push-button on the left hand horn of the PF yoke. The button cycled through three "declutter" states: 1) No declutter (full symbology and scene imagery); 2) "Raster" declutter (full symbology, no scene imagery); and 3) "Full declutter" (no HUD display).

2.2.2. Auxiliary Display

The PNF-Auxiliary Display (PNF-AD) was located outboard of the PNF location. The display was positioned as a compromise between optimal PNF viewing position, minimal display/instrument panel obscuration, and moderate installation complexity. The 8.4" diagonal display was full-color with 1024 x 768 pixel resolution. The display video source was a video mix of symbology and computer-generated scene imagery (either EVS or the output of a fused EVS/SVS signal as described in Section 2.7.2.).

2.2.3. Head-Down Displays

Minimal changes were made to the Primary Flight Display (PFD) and Navigation Display (ND) for the experiment so they closely resembled current B-757 equipage. The PFD was only modified to include a Flight Path Marker (FPM) and guidance cue. The PFD FPM and guidance cue were driven by algorithms identical to the HUD. Standard B-757 ship's flight director needles were disabled.

No changes to the ND were made. Of note for this experiment, the ND showed the Flight Management System programmed approach path but it did not include any Enhanced Ground Proximity Warning System "peak's mode" nor Traffic Alert and Collision Avoidance System (TCAS) information.

2.3. Synthetic Vision System

A synthetic vision database was created from a 1 arcsec (30 meter post-spacing) Digital Elevation Model (DEM) of a 53 x 53 nm area centered around the Reno-Tahoe International Airport (FAA identifier: KRNO). The airport was represented by three-dimension models of the runway, taxiways, and terminal buildings.

The DEM was draped with 1 meter/pixel satellite imagery within a 16 x 21 nm area centered around KRNO and 4 meter/pixel imagery outside this inner region.

2.4. Out-the-Window (OTW) Scene

The OTW imagery used the same source data as the SVS database but was rendered using different graphics processes and computers.

2.5. Enhanced Vision System

A physics-based Forward Looking Infra-Red (FLIR) simulation (using Evans & Sutherland EPX Sensors™) was created from the OTW visual database by applying materials properties to each component of the data. The characteristics of a short/mid-wave FLIR were simulated in a "white-hot" presentation. The time-of-day, time-of-year, and other diurnal properties were held constant.

Atmospheric properties (cloud layer, cloud height and thickness, fog, and visibility) were varied experimentally to modulate the visibility that the evaluation pilots had in the FLIR and the OTW scene presentations.

2.6. Symbology

2.6.1. HUD Symbology

The HUD stroke symbology format was based on the HGS-4000 "Primary Mode" stroke symbology set, albeit with the compass rose symbol removed (see Fig. 3). The following symbology was added: 1) a flare cue; 2) glideslope and localizer raw data indicators which included a deviation scale, angular deviation indication (i.e., glideslope and localizer deviation); and, 3) path deviation indication (i.e., "dog-bones").

The pitch-roll guidance cue ("ball") used modified pursuit guidance[11] along the desired path centerline, 5.5 seconds ahead of ownship. Horizontal and vertical position of the ball reflects the track and flight path angles to fly to the center of the desired path. The path deviation indicators showed "raw data" vertical and lateral path error as well as glideslope and localizer deviation (when available) for all the display conditions using error data scaled in "dots".
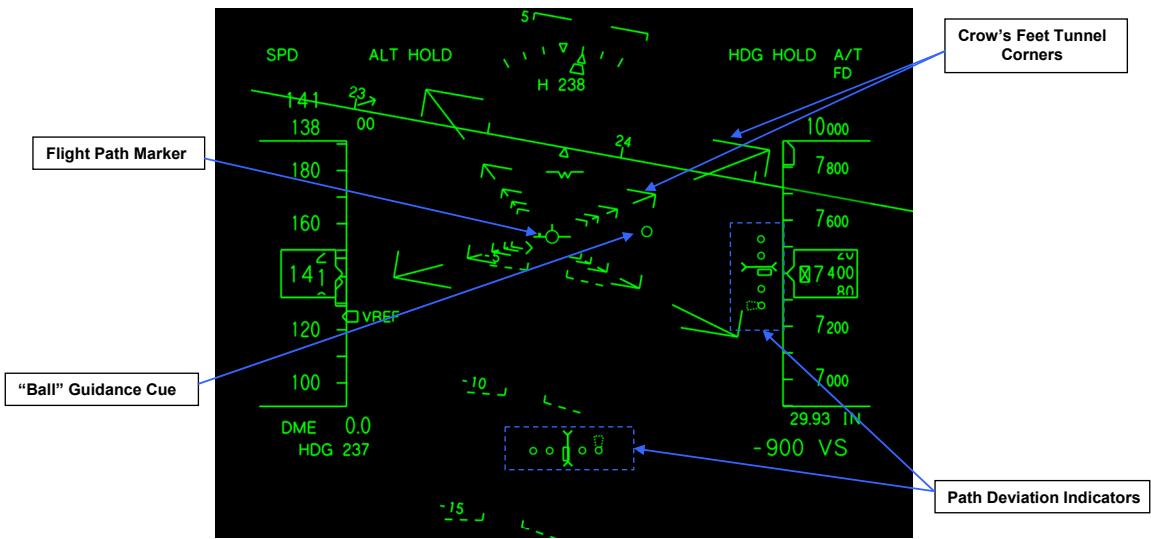

Figure 3. HUD Symbology with Advanced Pathway Guidance.

A glideslope reference line was drawn (Fig. 4) at the RNO Runway 16R Instrument Landing System (ILS) descent angle of -3.1 degrees. Also for some experimental conditions, a runway outline symbol was drawn using the threshold coordinates of the RNO 16R/34L runway based on the simulated aircraft navigation solution to conformally position the symbol. The runway outline was drawn using an 8000 ft x 200 ft runway.
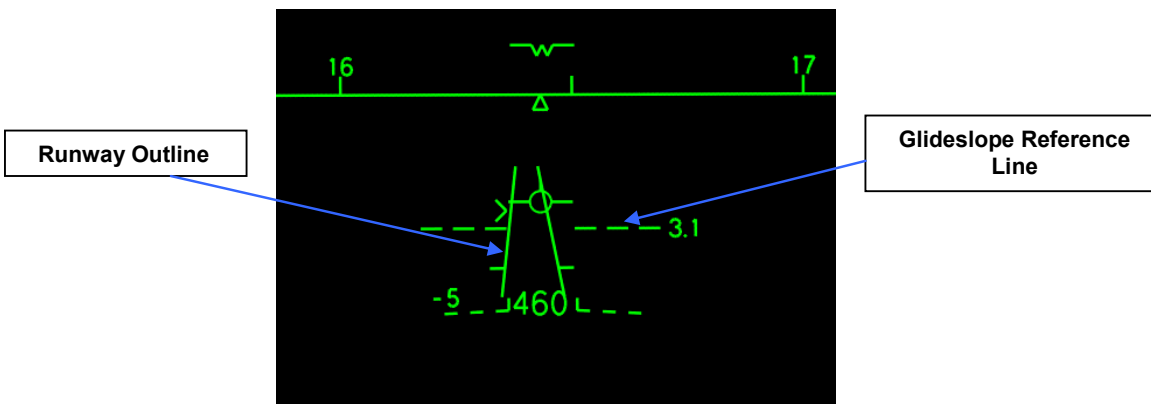

Figure 4. HUD Runway Outline Symbol and Glideslope Reference Line.

2.6.2. Tunnel

As an experiment variable, advanced pathway guidance in the form of a "minimal" tunnel was flown (see Fig 3). The minimal tunnel concept consists of a series of "crow's feet" which represented the truncated corners of nominally-connected 2-dimensional rectangles spaced at 0.2 nm increments along the desired path. The tunnel portrayed a constant 600 ft wide (±300 ft lateral) by 350 ft high (±175 ft vertical) path, 1 nm ahead of ownship position, along the desired path. One dot of vertical and lateral path error ("dogbone" deviation) corresponds to the vertical and lateral extent of the tunnel, respectively.

The minimal tunnel was selected on the concern that clutter would be critical for a HUD. Past studies[2,12,14] have shown that sufficient path information is provided by the minimal tunnel concept – at a minimal expense of display clutter – when path deviation indicators, pursuit guidance symbology and the FPM are also provided.

2.6.3. Auxiliary Display Symbology

The auxiliary display symbology (when used) was a subset of the HUD symbology to aid the PNF in monitoring the approach without obscuring too much of the raster (EVS, SVS or Fused) image. The symbology included digital readout of indicated airspeed and altitude; zero pitch attitude line (horizon line)/ heading scale with track indicator; flight path marker, pitch/roll (ball) guidance cue; path deviation indicators; ILS deviation indicators and scales; waterline; radio altitude, and event marker enunciators. (The event marker enunciators were not needed for the evaluation subjects, but were included for experimental data recording.)

2.7. Display Concepts

Four head-up display concepts and four auxiliary display concepts were evaluated by the evaluation crew (PF and PNF) while flying approaches to Reno-Tahoe International Airport. The head-down display formats were invariant.
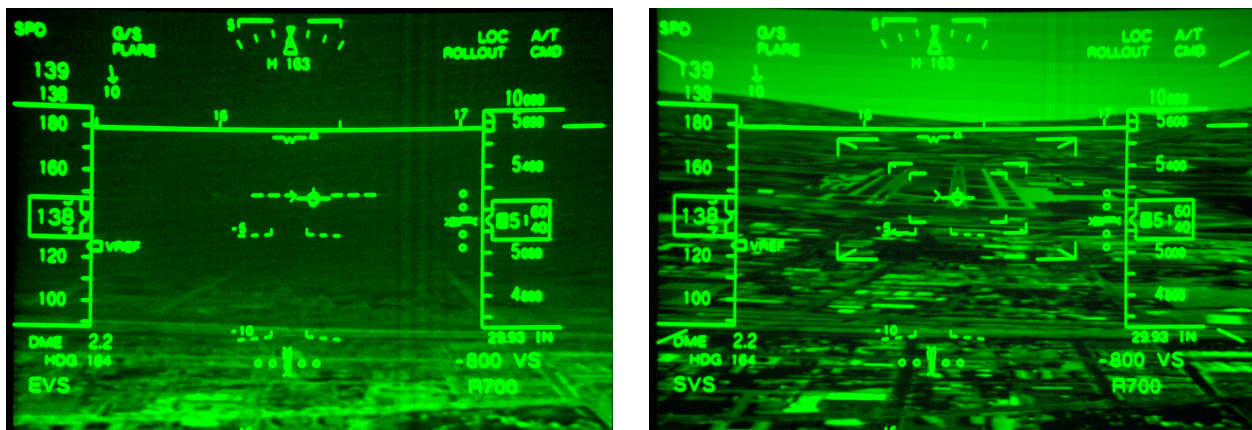
HUD Concept
EVS (FLIR) Only – Baseline

HUD Concept
Fusion – Tunnel

Figure 5. Head-Up Display (HUD) Formats

2.7.1. Head-Up Display Concepts

Four HUD display concepts were tested, differing from each other in: 1) the type of raster background presented; and, 2) in the type of symbology presented. In Figure 5, two of the concepts are shown - the FLIR-Baseline HUD (left) and the Fusion-Tunnel HUD (right).

Two raster formats were flown:

1. EVS-only (hereafter referred to as "FLIR"). The FLIR concept represented our "baseline" HUD condition.

2. A fusion of SVS/EVS imagery (hereafter referred to as "Fusion"). The Fusion raster started out as pure SVS imagery, transitioning through a fused SVS/EVS presentation beginning at 600 feet above field level (AFL), and ending with a pure FLIR imagery by 500 feet AFL. Between 600 feet and 500 feet AFL, a step function modulated the fusion from 100% SVS / 0% EVS ending at 0% SVS / 100% EVS.

Each raster concept showed FLIR below 500 ft to take advantage of the operational credit now offered by use of FLIR on the HUD. The 500 ft transition altitude was chosen for the Fusion transition altitude from a usability study prior to the test to assess an optimum altitude after which FLIR would be required.

The "Fusion" concept provides the basis to evaluate the utility, acceptability, and usability of SVS and EVS on the HUD. This concept was chosen because the "Fusion" maximizes image legibility and minimizes image confusion in that only one source is shown to the pilot and it is easily identifiable by the PF. This approach contrasts "inserts" and

other concepts for combining SVS and EVS within HUDs that have been evaluated elsewhere[4]. Pilot-controllable fusion or integration of images was not desirable as the PF already is burdened with flying; HUD brightness, contrast, and declutter control; EVS control; etc. and shouldn't be hampered with another display control task.

Two symbology sets were flown:

1) Standard HUD symbology (hereafter referred to as "Baseline")

2) Standard HUD symbology enhanced with pathway guidance and a runway outline (hereafter referred to as the "Tunnel" symbology set). The "tunnel" symbology set was tailored to transition at the same altitudes as the Fusion raster. Also, this transition altitude was based on flight test experience.[1] The tunnel was shown above the 500 ft above field level (AFL) transition altitude, the last tunnel segment was positioned at 500 ft AFL (thus, it was no longer visible below 500 ft), and, upon reaching 500 ft AFL, the runway outline was drawn until 50 ft AFL.
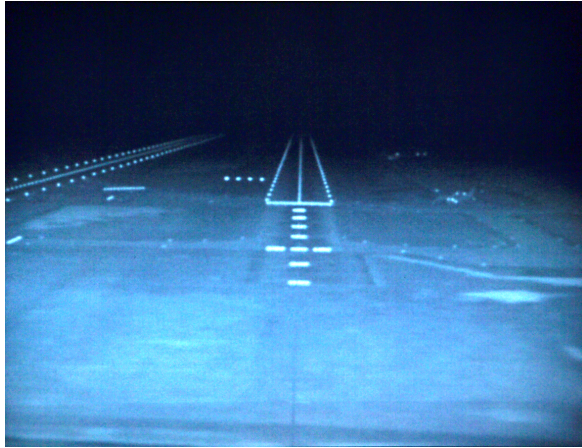
2.7.2. Auxiliary Display Concepts

Four PNF-AD display concepts were tested, differing from each other in: 1) the type of raster background presented; and, 2) the type of symbology presented. The raster was either EVS only (hereafter referred to as FLIR) or a fused SVS/EVS image (hereafter referred to as Fused). The symbology was either "On" or "Off" for the data runs. When present, the symbology was a subset of the standard HUD symbology (see Section 2.6.3). In Figure 6, two of the PNF-AD concepts are shown, the "FLIR" PNF-AD without Symbology (left) and the "Fused" PNF-AD with Symbology (right). (The terminology "Fused" was used when the pilot controlled the blending of SVS and EVS imagery, such as the case for the PNF-AD. Whereas, the term "Fusion" was used when the blending was automatically controlled, such as the case for the PF-HUD.)

The AD fused raster image was pilot-controllable and could be tuned throughout the approach to one of 10 states: FLIR only, SVS only, or 8 fusion combinations of FLIR and SVS, using an Equinox EP-3000™ fusion board. The fusion employs a feature-level extraction algorithm with two pilot control inputs. The first control biased the feature level fusion through 8 weighting values toward FLIR or toward SVS. (A value of 1 biased the extraction to 11% FLIR and 89% SVS whereas a value of 8 weighted the extraction to 89% FLIR and 11% SVS). The second control modulated the false-color coding of the fusion image through 1 of 8 values. A setting of 1 did not apply any color-coding (the display was a monochromatic fused image). A setting of 8 applied maximum green shading to the features which were assessed by the fusion algorithm to be "common" features between the two input videos and which had spatial frequency content above a threshold value (i.e., "high spatial content").

2.8. Evaluation Task

The evaluation task was selected to approximate what may be typical of the emerging NGATS concept called an "equivalent visual operation." The task was based on a published visual arrival – reflecting an efficient and preferred routing for ATC and noise-abatement – which currently requires visual meteorological conditions (VMC) for the pilot to see-and-avoid terrain, traffic, and obstacles while navigating with respect to ground references. The approach path is not too dissimilar from a Required Navigation Performance (RNP)-type arrival, requiring a curved, descending path. The evaluation task tests the ability of SVS and EVS technology to support this type of operation by providing "equivalent visual" information into the cockpit. Further, this technology succeeds in providing a visual arrival capability, the potential for operational efficiency and minimums reduction above and beyond what can be provided by RNP may be offered.

The Pilot Flying (PF) hand-flew the base and final leg portions of the Sparks Visual Arrival to RNO Runway 16R (see Fig. 7) with autothrottles engaged at an approach speed of 138 knots. The aircraft was configured for landing prior to each run (landing gear down and flaps 30 degrees). The path converged into the ILS for Runway 16R. The aircraft was configured to land and established on a stabilized course and descent by 1000 ft AFL. The Pilot Not Flying (PNF) monitored the approach from the right-hand side of the flight deck using standard instruments and the AD. Pilot participants were instructed that the run would end at main gear touchdown but that they should perform a go-around if they felt the landing was not safe.

AD Concept
EVS (FLIR) Only – No Symbology

AD Concept
Fusion – Symbology

Figure 6. Two Auxiliary Display (AD) Formats

2.9. EVS Crew Procedures

EVS crew procedures, adapted from those used currently in business aircraft EVS operations, were established. Instructions in the use of the procedures were given to each crew. An overview of these procedures is given in Table 1, including automatic call-outs. The altitude call-outs were set-up assuming a 200 ft Decision Height (DH) for the published, non-EVS approach. (A "flat-earth" model was used so differences between the barometric altitudes and radar altitudes for decision altitudes / heights were inconsequential.)

At 500 ft AFL, the "EVS Normal, System Normal" call by the PF corresponds to the point where the PF would nominally check that the EVS was set-up properly and functioning properly on the approach. In our test, this call-out cued the PF to ensure that the HUD declutter, brightness and contrast were properly set. No EVS controls were available to the flight crew in the experiment.
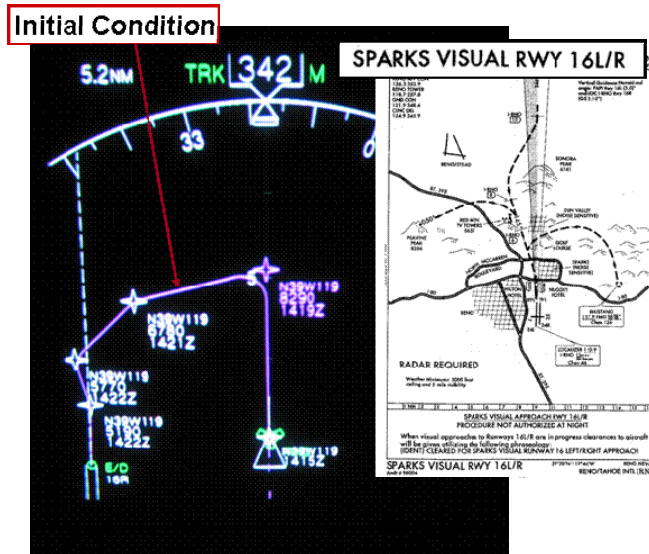

Figure 7. Sparks Visual Arrival to RNO Runway 16R.

By the published minimums of 200 ft (DH), the crew procedures dictate that the PF must have the required EVS references or the required landing visual references (using natural vision) to continue the descent. The landing

references were those published in FAR 91.175. For this test, the approach light system for RNO 16R provided the prominent EVS references. If these EVS references were visible, the PF was instructed to call "EVS Lights".

If the PF saw the lights or markings of the threshold (the predominant landing visual reference for RNO 16R), the PF called "Landing." The "landing" call was required no later than 100 ft DH.

The PNF provided monitoring, including back-up on all decision heights, and was instructed to call "go-around" if "EVS Lights" was not called at or before 200 ft DH or if "Landing" was not called by 100 ft DH.

Table 1. EVS Crew Procedures

Altitude-Based Events	AFL / Baro-Altitudes (ft)	Automatic Callouts	Pilot Flying (PF) Tasks	Pilot Not Flying (PNF) Tasks
500 feet AFL	500 / 4912	"*500*"	Response: "*Systems Normal, EVS Normal*"	Call "*500 feet*"
100 feet Above Minimums	300 / 4712	"*Approaching Minimums*"	Response: "*Check*"	Call "*100 feet Above*"
Published Minimums (200 ft AFL)	200 / 4612	"*Minimums*"	**With <u>EVS</u> <u>Visual Cues</u>,** Call "*EVS Lights*"	**When <u>Visual Cues</u> Appear** Call "*Lights*" or "*Field in Sight*"
			Without <u>EVS Visual Cues</u>, Call "*Going Around*"	**Without PF Call of 'EVS Lights'**, Call "*Go Around*"
EVS Decision Altitude (100 ft AFL)	100 / 4512		**When <u>Actual</u> <u>Visual Cues</u>,** Call "*Landing*"	**When <u>Visual Cues</u> Appear,** Call "*Lights*" or "*Field in Sight*"
			Without <u>Actual Visual Cues</u>, Call "*Going Around*"	**Without PF Call of 'Landing'**, Call "*Go Around*"

The PNF was allowed to assist the PF in picking up the required visual cues (normal or EVS). Transfer of control between the Captain and First-Officer was not permitted.

The crew procedures were new to all of the flight crews. Some procedures were counter, others consistent with their current airline Standard Operating Procedures (SOPs). In either case, the crew procedures were trained and "enforced" for the test. During the post-test debrief, questions and issues of how these procedures may or may not work within their airline operation and SOPs were discussed. Flight crews from the same airline were paired to the greatest extent possible to minimize SOP differences and influences in crew interaction.

2.10. Experiment Matrix

Nominally, forty experimental runs were completed by the evaluation crew with each pilot flying 20 approaches evaluating the HUD concepts and with each pilot monitoring 20 approaches while evaluating the AD concepts.

The wind and weather varied on each run. The nominal visibility in the EVS and OTW varied from 1 mile down to ½ mile. The required EVS visual references became visible on the HUD between 450 ft and 250 ft AFL. Four runs per flight crew were specifically designed so the EVS visual references were visible but the required runway (normal vision landing) references were not. These four runs, if properly flown using the EVS crew procedures, should conclude by a go-around initiated no lower than 100 ft AFL.

The PF was instructed to fly each approach as precisely as possible using the display information available, as the effect of the display information on the PF's ability to fly the approaches would be quantitatively and qualitatively evaluated. In addition, the PF was instructed to land as close as possible to the centerline of the runway.

A significant component of the test, in addition to the nominal runs, was met by measuring the ability of the flight crew to react and properly handle non-normal events. Four non-normal runs were flown by each crew. The non-normals were runway incursion (RI) scenarios and database integrity monitoring scenarios. The number of RI and database integrity scenarios were designed to avoid expectancy on the part of the flight crew.[15] The RI scenarios simulated an incursion with either a non-transponding baggage cart or fire truck. The database integrity monitoring scenarios purposefully introduced a lateral navigation solution error (of either 50 or 75 feet) with respect to the real runway. This

error resulted in the synthetic vision terrain, pathway and guidance cue being misaligned from the FLIR and ILS (which were defined in the flight crew briefing as being correct).

2.11. Measures

During each run, path error and pilot control inputs were recorded for analysis.

After each run, pilots completed a run questionnaire consisting of the Air Force Flight Technical Center (AFFTC) Revised Workload Estimation Scale[16], Situation Awareness Rating Technique (SART)[17], and four Likert-type (7-point) questions specific to different constructs of display clutter (see Fig. 8).

The clutter data are being used by NASA to develop improved subjective and objective measures for display clutter. (Only the overall clutter data, Rating #4 in Fig. 8, is discussed herein; the remainder will be published separately.) These metrics will be critical tools for the flight deck designer as emerging NGATS operating concepts - relying on Shared Situation Awareness concepts – dictate that massive amounts of on-board and off-board information come into the flight deck for comprehension, decision making, and action by the flight crew. Without verified and validated clutter metrics (among other things), the design process to achieve these capabilities will be hit-and-miss, at best.

After data collection was completed, pilots were administered two separate Situation Awareness – Subjective Workload Dominance (SA-SWORD)[18] and Subjective Workload Dominance (SWORD)[17] tests: one for HUD concept (Baseline-FLIR, Tunnel-FLIR, Baseline-Fusion, Tunnel-Fusion) comparisons and another for AD concept (FLIR only, FLIR with Symbology, Fused only, Fused with Symbology) comparisons. The pilots also participated in a semi-structured interview to elicit comments on the HUD/AD concepts, HUD SVS-to-EVS transition strategy for the fusion concept, AD fusion strategy, and EVS crew procedures.

For the post-run questions, separate analysis of variance (ANOVA) analyses were conducted for the HUD concepts and the AD concepts. For the HUD concepts, there were two main factors, each with two levels: 1) Raster (FLIR, Fusion) & 2) Symbology (Baseline, Tunnel). For the AD concepts, there were two main factors, each with two levels: 1) Raster (FLIR, Fused) & 2) Symbology (On, Off). When the 2nd order interaction was significant, a simple main effects analysis was conducted using $\alpha=.05$. For the post-test paired comparisons, simple ANOVAs and Student-Newman-Keuls (SNK) post-hoc tests with alpha (α) set at 0.05 were performed.

Display Clutter Ratings	Low....................................High						
	1	2	3	4	5	6	7
Rating #1 Demand on Visual Attention How much visual search time and cognitive effort was required to scan and locate task-critical display information in the display?							
Rating #2 Supply of Visual Attention Resources How much spare visual attention and mental ability was available to accomplish secondary task(s)?							
Rating #3 Understanding What was your ability to quickly and accurately understand task-critical display information?							
Rating #4 As I Performed the evaluation task, the level of display clutter on the HUD/Auxiliary Display was_____.							

Figure 8. Post-Run Display Clutter Questions.

2.12. Procedure

The subjects were given a 1-hour briefing to explain the SVS/EVS concepts on the HUD and AD, EVS crew procedures, and the expected evaluation tasks. After the briefing, a 2-hour training session in the IFD was conducted to familiarize the subjects with the aircraft handling qualities, display symbologies, crew procedures, and controls. The 'rare-event' scenarios were not discussed, although the pilot's responsibility for maintaining safe operations at all times was stressed. Data collection lasted approximately 4.5 hours followed by a 30-minute semi structured interview. The pilots were also given a take-home final questionnaire. The entire session including lunch and breaks lasted approximately 9 hours.

3. RESULTS

3.1. Path Control Performance

Root-mean-square (RMS) calculations of lateral and vertical path error were used as the measures for flight path control performance. Separate ANOVA analyses were performed on RMS path error (lateral and vertical) for two segments of the run: approach and final. The approach segment began at the task starting point and ended at 500 feet AFL. The final approach segment was between 500 feet and 100 feet AFL. Two runs were excluded from these analyses due to simulation problems (e.g., lost path on navigation display during those runs). The non-normal runs with a lateral navigation error (50 feet or 75 feet) were not included in the final segment analyses.

3.1.1. Approach Lateral Path Error

An ANOVA revealed that HUD concept ($F_{(3,443)}=9.73$, $p<.01$) was statistically significant for RMS lateral path error during the approach segment of the flight. Post-hoc tests revealed two unique subsets: 1) FLIR-Baseline (mean=42 feet) & Fusion-Baseline (mean=37 feet) and 2) FLIR-Tunnel (mean=30 feet) & Fusion-Tunnel (mean=28 feet).

3.1.2. Approach Vertical Path Error

An ANOVA revealed that HUD concept ($F_{(3,443)}=6.69$, $p<.01$) was statistically significant for RMS vertical path error during the approach segment of the flight. Post-hoc tests revealed that the FLIR-Baseline (mean=12 feet) concept had significantly higher RMS vertical path error than the other three concepts: Fusion-Baseline=10 feet; FLIR-Tunnel=9 feet; & Fusion-Tunnel=9 feet.

3.1.3. Vertical and Lateral Path Error on Final

There were no significant differences ($p>.05$) for HUD concept for the RMS lateral path error (mean=7 feet) or RMS vertical path error (mean=7 feet) during the final segment of the flight.

3.1.4. Path Control Results Discussion

The path control results show that the tunnel concepts have lower RMS lateral path error than the baseline, non-tunnel HUD concepts on the approach segment. The primary difference between the configurations was presence or absence of the tunnel and the turn anticipation cues that it provides. However, while statistically significant, the operational significance of the differences appears weak. An analysis of the path error data for Flight Technical Error (FTE) as a component of Required Navigation Performance has not yet been conducted to determine the operational significance of the effect. The approach vertical path error also showed statistical significance, but the very small differences imply little operational significance.

The pilot performance results are supported by past research.[1, 12, 13, 19] Minimal performance differences were expected since each display concept utilized the same pursuit guidance control laws and symbology (i.e., the flight path marker, integrated cue guidance symbol, and path deviation indicators).

Subjectively, the EPs also felt that the SVS background, when present, also improved flight path control performance because the database imagery in the background provided stronger roll reference visual cues. This limited a tendency to overcontrol in roll in the no-motion B-757 simulator, particularly when flying the baseline symbology set (i.e., compensatory guidance symbol only).

3.2. Mental Workload

Mental workload was assessed after each experimental run, using the AFFTC workload estimate tool, and post-test, using SWORD.

3.2.1. AFFTC Workload Estimate – PF HUD

The main factors of HUD raster ($F(1,366)=4.47$, $p=.035$) and symbology ($F(1,366)=25.06$, $p<.01$) were significant for workload. However, the raster by symbology interaction was not significant ($p>.05$). Post-hoc tests (SNK using $\alpha=.05$) showed two unique subsets for workload ratings with the 4 HUD combinations: 1) Fusion-Tunnel (mean = 3.1) and 2) FLIR-Tunnel (mean = 3.4); Fusion-Baseline (mean = 3.5) & FLIR-Baseline (mean = 3.6).

Pilots ranked the Fusion-Tunnel HUD concept as having significantly less workload than the other 3 HUD concepts tested. On the AFFTC Workload Scale, a value of "3" indicates "Moderate Activity – Easily Managed; Considerable Spare Time" and a value of "4" indicates "Busy – Challenging but Manageable; Adequate Time Available".

3.2.2. AFFTC Workload Estimate – PNF-AD

For the PNF-AD concepts, there were no significant ($p>.05$) differences between raster type (FLIR, Fused) and symbology (Off, On) or the interaction between these two factors for post-run workload. A mean pilot rating of 2.6 was given for the AD concepts by the pilots. On the AFFTC Workload Scale, a value of "2" indicates "Light Activity; Minimum Demands" and a value of "3" indicates "Moderate Activity – Easily Managed; Considerable Spare Time."

This result indicates that the presence or absence of symbology and the presence or absence of fusion controls for the PNF does not have a measurable effect on pilot monitoring workload.

3.2.3. SWORD

Pilots were administered the pair-comparison SWORD scale that enabled ratings of mental workload across the four display concepts (raster*symbology) for both the PF and PNF displays. The definition of mental workload used was "The amount of cognitive resources available to perform a task and the difficulty of that task."

The post-test SWORD data indicate that there were no significant ($p<.05$) differences among the HUD concepts for the PF ratings of mental workload.

The SWORD data show that the AD concept was highly significant ($F(3, 69)=15.02$, $p<.001$). Post-hoc tests (SNK using $\alpha=.05$) showed three unique subsets for the mental workload ratings with the 4 PNF-AD concepts: 1) FLIR-Symbology & Fused-Symbology (lowest workload); 2) Fused-No Symbology; and 3) FLIR-No Symbology (highest workload).

3.2.4. Workload Discussion

The workload data do not show *substantial* differences associated with PF-HUD concepts. Pilots ranked the Fusion-Tunnel HUD concept as having significantly less workload than the other 3 HUD concepts tested, post-run. Operationally, the workload data suggest that the Fusion-Tunnel concept reduced PF workload to the extent that the average workload is "easily managed", whereas the other concepts elicited workload ratings tending toward "challenging but manageable" workload levels. However, the post-test SWORD data indicate that there were no significant ($p<.05$) differences among the HUD concepts for the SWORD ratings of mental workload. Since the definitions of workload and the types of tests differ, the data suggests weak differences, if any, in the workload associated with the PF-HUD concepts. Pilot commentary suggested that the workload when flying the tunnel symbology concepts was easier (less scan between the HUD and ND, easier to anticipate the turns), but the differences were not of a magnitude to warrant concern.

Similarly, the post-run AFFTC workload ratings for the PNF-AD concepts showed no statistically significant differences, but post-test, pilots ranked the two AD concepts with symbology as requiring significantly less mental workload in their SWORD ratings. (There were no appreciable differences between the FLIR Symbology and Fused Symbology AD concepts for mental workload ratings.) Pilot commentary typically noted the advantage of symbology in reducing the visual scan and cognitive task of integrating the different display information. However, the differences, again, were not of a magnitude to warrant concern.

3.3. Situation Awareness

Situation awareness was assessed after each experimental run, using the post-run SART, and post-test, using SA-SWORD measures. Situation awareness was assessed in both PF and PNF roles. SART is a multi-dimensional rating technique using the constructs of: 1) demand on attentional resources; 2) supply of attentional resources; and, 3) understanding. From these components, the SART rating is "understanding" reduced by the difference of "demand" minus "supply" (i.e., SART = {(understanding) – (demand – supply)}).

3.3.1.   SART – PF HUD

An ANOVA revealed that both HUD raster type (F(1,366)=3.23, p<.01) and symbology type (F(1,366)=38.10, p<.01) and the interaction between these factors (F(1,33)=4.22, p=.04)) were significant for PF-HUD SART ratings. Pilots rated their SA significantly higher when the HUD symbology included pathway/tunnel guidance and the Fusion imagery. A simple main effects analysis revealed that the effects of symbology type on SA ratings was stronger within the Fusion raster type than within the FLIR raster type.

3.3.2.   SART – PNF-AD

An ANOVA on SART ratings for PNF display found no significant results for the main effects (raster, symbology) or the interaction (p > .05).

3.3.3.   SA-SWORD

Similar to the SWORD described above, the SA-SWORD is a paired comparison technique that provides relative situation awareness ratings across the four display concepts for both the PF and PNF displays. For these comparisons, SA was defined as "The pilot's awareness and understanding of all factors that will contribute to the safe flying of their aircraft under normal and non-normal conditions." The SA-SWORD measure differs from the post-run SART measure construct as it is a pair-comparison test, it was administered post-test, and the underlying definition and construct for the SA ratings are different.

An ANOVA revealed that the HUD concept was highly significant (F(3, 69)=43.61, p<.001) for the SA-SWORD ratings. Post-hoc tests (SNK using α=.05) showed three unique subsets for situation awareness ratings with the 4 HUD concepts: 1) Fusion-Tunnel; 2) Fusion-Baseline & FLIR-Tunnel; and 3) FLIR-Baseline. Pilots ranked the Fusion-Tunnel HUD concept as having significantly higher SA than the other 3 HUD concepts tested

The PNF-AD concepts were highly significant (F(3, 69)=37.78, p<.001) for the SA-SWORD ratings. Post-hoc tests (SNK using α=.05) showed three overlapping subsets for situation awareness ratings with the 4 AD concepts: 1) Fused-Symbology; 2) FLIR-Symbology & Fused-No Symbology; and 3) Fused-No Symbology & FLIR-No Symbology. Pilots ranked the Fused-Symbology AD concept as having significantly higher SA than the other 3 AD concepts tested. During the post-run SART ratings, no statistically significant differences were found.

3.3.4.   Situation Awareness Discussion

Pilot commentary noted that the presence of the tunnel gave the pilots a much better understanding and appreciation of the curving, descending visual arrival path. Without the tunnel guidance, pilots commented that they had to use the head-down ND more frequently for path (turn) guidance. (Head-tracking data was unfortunately not recorded for the PF.) Also, the SVS component in the Fusion HUD concept provides significant terrain information unavailable in any other cockpit displays. These SA components emerged in the SA results, both post-test and post-run. Superior SA was rated for the Fusion-Tunnel HUD concept whereas the FLIR-Baseline was rated inferior for SA. Interestingly, by the SA measures, the Fusion-Baseline & FLIR-Tunnel concepts were not significantly different. This result would imply that the pilots felt the SVS contribution to SA was equivalent to the tunnel contribution to SA.

No statistically significant differences were noted in SA (post-run) using SART for the PNF-AD concepts, but post-test, the benefits of Fused imagery and symbology on the PNF-AD emerged. The post-test and post-run SA differences might be attributed to the fact that SA-SWORD asks for a *general* appraisal whereas the SART asks for ratings from what was experienced for *that* pilot on *that* run. SA can be high - it was high in all conditions, including the baseline, as they were all highly skilled pilots - and the task really wasn't extremely demanding of the PNF. But, when asked to compare the PNF display concepts post-test to each other, SA differences emerged. Post-test, the pilots commented that they felt SA was impacted by several issues. SA was significantly improved with Fused imagery on the PNF-AD by providing a way to better monitor the EVS and navigation system performance and improve their understanding of their

flight path with respect to terrain. Symbology on the PNF-AD provided two key SA benefits. First, the FPM and guidance cue (with FLIR and/or SVS imagery) provided visual evidence that the PF was flying to the proper point on the ground (i.e., flying to the intended runway, touchdown point) and the raw data displays on the PNF-AD symbology was the only location of path error for the PNF (i.e., the "dog bones" were not shown on the PNF's PFD.) Without the dogbones, the PNF had to use the ND to monitor approach-tracking performance. These PNF-AD attributes may not have been critical to the experiment on each run (i.e., minimal impact on post-run SART) but they can contribute significantly to SA for a PNF in general in this type of operation.

3.4. Pilot Display Preferences

Separate post-test paired comparisons for pilot display preferences were made on the HUD and AD concepts after data collection was completed. HUD concept was highly significant ($F(3, 69)=73.17$, $p<.001$) for the Pilot-Preferred Display ratings. Post-hoc tests (SNK using $\alpha=.05$) showed three unique subsets for the pilot-preferred display ratings with the 4 HUD concepts: 1) Fusion-Tunnel; 2) Fusion-Baseline & FLIR-Tunnel; and 3) FLIR-Baseline. Pilots ranked the Fusion-Tunnel HUD concept as being preferred significantly more than the other 3 HUD concepts tested.

AD concept was highly significant ($F(3, 69)=23.74$, $p<.001$) for the Pilot-Preferred Display ratings. Post-hoc tests (SNK using $\alpha=.05$) showed three overlapping subsets for the pilot-preferred display ratings with the 4 AD concepts: 1) Fused-Symbology; 2) FLIR-Symbology & Fused-No Symbology; and 3) Fused-No Symbology & FLIR-No Symbology. Pilots ranked the Fused-Symbology AD concept as being preferred significantly more than the other 3 AD concepts tested.

3.5. Subjective Assessments of Display Clutter

After an experimental run, each pilot gave ratings for the 4 Likert-type questions on display clutter (Figure 8) for the display (HUD or AD) concept they had just flown; only the "overall" clutter data are discussed here.

For the overall display clutter ratings, there were no significant differences for raster type, symbology type, or the raster by symbology interaction for the PF-HUD concepts. The average rating for the HUD concepts was 3.3 which corresponds to a moderate amount of display clutter for all concepts.

Symbology type ($F(1,363)=28.89$, $p<.01$) was highly significant (but not operationally) for the PNF-AD post-run overall clutter rating with the baseline symbology concept (1.8) having a lower rating (less overall clutter) than the baseline symbology plus pathway guidance concept (2.2). These ratings for the AD concepts correspond to a moderately low level of display clutter.

The pilots noted that the symbology on the PNF-AD was beneficial to SA, but contributed to clutter. The pilots – as always – want symbology *and* a completely clear FLIR or Fusion raster on the AD to promote better readability and understanding of the imagery. The proposed solution was to include symbology on the PNF-AD and also, include a PNF-AD "declutter" button, analogous to the PF-HUD, so the symbology could be toggled on and off as needed.

3.6. Fusion Controls

By providing a plethora of controls to the PNF for the Fused AD concept, the experiment data provided first-order determination if: a) a "fusion" concept was viable in the commercial cockpit; b) allowing the PNF to control their presentation was viable or desirable; and, c) SVS and EVS was necessary for the PNF.

In Figure 9, the percentage of time that a fusion control settings was used by the PNFs during the experiment is plotted by altitude range. The altitude ranges correspond to the approach segment ("APP": run start to 500 ft AFL), and 100 ft increments from 500 ft AFL to touchdown (or go-around). A fusion setting of 1 corresponds to SVS-only on the PNF-AD, a fusion setting of 10 corresponds to EVS-only , and a setting of 2 through 9 corresponds to the feature-level extraction algorithm bias as described in Section 2.7.2.

The data indicate a consistent trend. At altitude (on Approach down to approximately 500 ft AFL), the PNFs tended to use the feature-level fusion of EVS and SVS. The most prevalent settings were heavily weighted toward EVS (i.e., settings of 8 and 9 in Figure 9). On the approach segment, the EVS did not have any information content because of simulated clouds on the approach. With feature-level extraction, the fusion image shows the SVS database image without significant alteration or contrast reduction. The PNFs often used an intensity (false color-coding) value of 8 so when color appeared on the PNF-AD, this cued the pilot that the EVS was starting to show useful information. The color signaled that they could effectively begin using an "EVS-Only" setting. The data indicate that fusion control was used - albeit not to its full-range - and the PNF gathered significant information that assisted in their monitoring

function. The EPs quickly learned how the fusion worked and what the most effective means to employ the controls were.
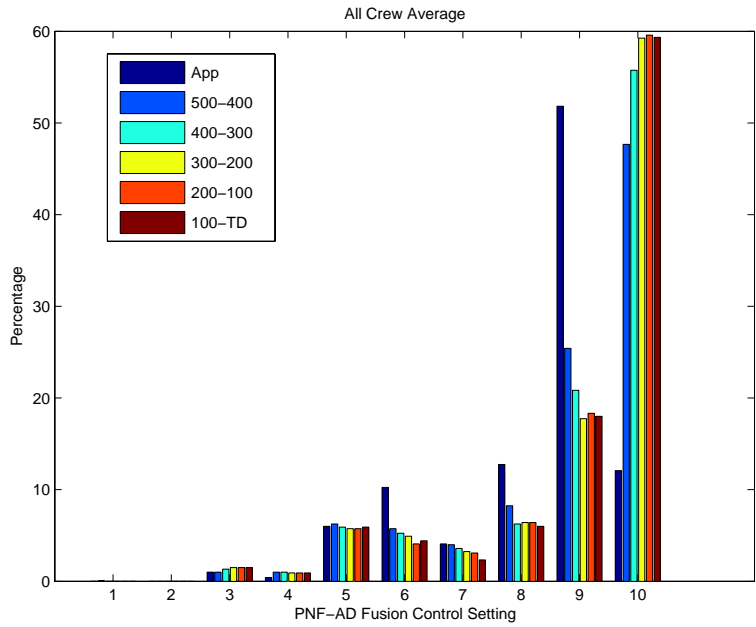

Figure 9. All Fusion Control Settings by Altitude.

These data are further collapsed to highlight the trends in Figure 10. The percentage of time that the PNF used any fusion settings (i.e., SVS-Only or feature-level fusion values of 1-8, see Section 2.7.2) or "EVS-Only". On the approach, fusion was used more than 85% of the time, but in the final approach segment, EVS-only was used 60% of the time and Fusion reduced to 40%. The 60-40 distribution in EVS-only and "Fusion" settings suggests that the PNF used both information sources cooperatively and effectively.
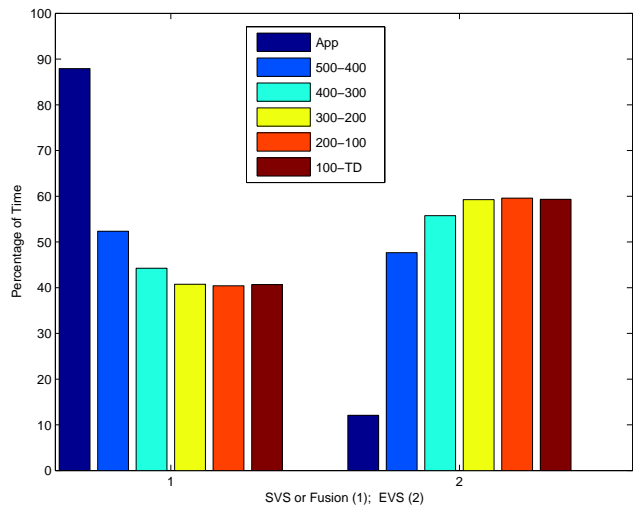

Figure 10. SVS/EVS Fusion and EVS-Only Setting by Altitude.

3.7. Non-Normals

Non-normals were injected into the test unbeknownst to the evaluation subjects. The non-normals were two runway incursions and four lateral offsets for each flight crew.

3.7.1. <u>Runway Incursions</u>

The runway incursions were made by a baggage cart and a fire truck. Both items were positioned in the same location, approximately 850 ft from the RNO Runway 16R landing threshold and just slightly offset from the centerline. The weather on the runway incursions was held constant at 2400 ft RVR (OTW) with the lowest cloud layer at 500 ft AFL. The FLIR visibility was very good in this condition – approximately 4 times the OTW RVR.

The baggage cart runway incursion was always performed before the fire truck incursion. The baggage cart was much more difficult to see due to its low visual contrast against the runway and small size. This ordering tested for "just noticeable differences" for runway incursion detection.

For the 12 flight crews, only one crew member (PNF) saw the baggage cart visually and initiated a go-around. The other 11 crews had a runway incursion with the baggage cart. All crews saw the incurring fire truck visually (i.e., not with the HUD EVS or PNF-AD), except for one PNF that saw it on the AD. Typically the fire truck incursion was detected first by the PNF looking out the window, except on three occasions where the PF saw the incursion OTW. Upon seeing the incursion, all crews, except one, initiated a go-around. The one crew that didn't initiate a go-around flew over the fire truck and landed long.

The incurring vehicles were visible in the PNF-AD and HUD, yet the data suggests that EVS on the HUD and PNF-AD were not useful for RI detection. In the HUD, the incurring vehicles were largely occluded by symbology on the HUD (FPM and guidance cue) and the small size and relatively low resolution of the HUD made vehicle detection extremely difficult for the PF.

In contrast, the vehicles were much more apparent in the PNF-AD. Again, the vehicle size and contrast to the surroundings made detection on the PNF-AD moderately difficult above 200 ft AFL, particularly if the PNF only used cursory looks at the PNF-AD. Below 200 ft AFL, the vehicles were much more obvious in the image, but the PNF noted that they were head-out the vast majority of the time. (Head tracking measurements were made and will be analyzed to quantify this statement.) Based on observation and pilot comments, the PNF was concentrating on picking-up the required visual references for landing, not runway incursion detection. Thus, the use of the PNF-AD for incursion detection was not practical in this scenario. Current flight crews are not familiar with using head-down displays on short final to check for incursions. This was not part of the pre-experiment flight crew instructional briefing.

3.7.2. <u>Navigation Error</u>

The lateral navigation offsets could be detected by either the PF or the PNF. The errors were noticeable from one of several principle ways (depending upon the display configuration):

- By a disagreement between the "dogbones" path error and the localizer deviation symbology (PF-HUD and PNF-AD with symbology).
- By a non-zero localizer deviation on the PFD when the PF is flying the final approach on the path centerline.
- By differences between the SVS and the EVS registration using the PNF-AD Fusion controls.
- By differences between the runway outline and the EVS imagery on the PF-HUD.
- By differences in the pitch/roll guidance symbol and the EVS imagery (PF-HUD and PNF-AD)

A complete analysis of the navigation error data has not yet been completed. The preliminary data does show that approximately 70% of the time the database offsets were verbally noted by the flight crew. They were predominately noted by the PF when they noticed that the pitch/roll guidance symbol was leading them to the left or right of the runway (seen through the EVS HUD presentation). Statistical distribution of the data has not yet been completed.

Only one person (flying as PNF) noted the navigation error through the dogbone and ILS raw data presentation.

None of the pilots executed a go-around with this anomaly. Each performed a lateral correction and landed near the runway centerline.

4. CONCLUSIONS

An experiment was conducted to evaluate the complementary use of SVS and EVS technologies, specifically focusing on new techniques for integration and/or fusion of synthetic and enhanced vision technologies and crew resource management while operating under the newly adopted FAA rules which provide operating credit for EVS. From these data, significant improvements in SA can be provided by the integration and/or fusion of synthetic and enhanced vision technologies for the PF and PNF.

The data showed that the tunnel concept promotes lower RMS lateral path error during the approach segment. However, while statistically significant, the operational significance of the differences appears weak. Further, qualitative data suggests that SVS added to background of the PF-HUD improved path control performance but quantitative evidence of this improvement was not conclusive. Analyses are on-going to further investigate these findings and their potential operational significance.

Workload for the PF and PNF was not substantially different when flying with the tested concepts. Thus, increasing the "informational complexity" of the HUD by adding SVS and tunnel data, and increasing the number of controls and symbology on a PNF-AD did not affect PF or PNF workload.

In contrast, SA for the PF and PNF was improved by the addition of tunnel and SVS on the HUD and by adding fusion control and symbology on the PNF-AD.

The ability of the flight crew to handle a substantial navigational solution error was not impacted by the display concepts. In all display concepts, the navigation error was detected or ignored. The pilots landed safely. Further analyses are on-going to tease out statistical correlations.

The ability of the flight crew to handle a runway incursion was neither impacted nor aided by the display concepts tested. Although the increase in near-domain symbology information (runway outline) did not degrade pilot response to the fire truck runway incursion event, there was also not an observed enhancement in incursion detection as hypothesized for the FLIR. Moreover, only one pilot detected the baggage cart despite the FLIR imagery. Analyses are on-going to further evaluate specific effects between the display concepts to determine the etiology of these findings.

Numerous suggested improvements have been identified and are being worked. For instance, the PNFs strongly suggested that a declutter capability on the PNF-AD should be developed. Symbology on the PNF-AD was strongly preferred and rated highly, but the presence of symbology degraded the readability of the raster, particularly of the runway and touchdown point.

Further analysis of the runway incursion data is on-going. The results must be decomposed into components that span the breadth of the problem, including human perception, sensor design and detection theory, crew procedures, and crew interface issues. The display concepts and scenarios tested in this experiment – typical of current and future PF HUD and PNF-AD displays - did not show adequate incursion detection functionality. All but one of the runway incursion scenarios were detected without the use of the cockpit displays.

Further analysis of the data is on-going. Several items, in particular, have not yet been addressed at the time of this publication, including the quantification of the PNF visual scan patterns, the operational significance of path performance data, the development of subjective and objective display clutter metrics, and an analysis of EVS crew procedures for commercial airline operations.

5. REFERENCES

1. Kramer, L.J., Arthur, J.J., Bailey, R.E., and Prinzel, L.J. III, *Flight Testing An Integrated Synthetic Vision System,* In Proceedings of SPIE, Enhanced and Synthetic Vision 2005, Editor: Jacques G. Verly, Volume 5802, Paper 1, April 2005.
2. Arthur, J.J., Prinzel, L.J., Kramer, L.J., Bailey, R.E., & Parrish, R.V. (2003). *CFIT Prevention Using Synthetic Vision.* In Proceedings of SPIE, Enhanced and Synthetic Vision 2003, Editor: Jacques G. Verly, Volume 5018, Paper 16, April 2003.
3. Schiefele, J. et al *Human Factors Flight Trial Analysis for 3D SVS: Part II,* In Proceedings of SPIE, Enhanced and Synthetic Vision 2005, Editor: Jacques G. Verly, Volume 5802, Paper 18, April 2005.
4. Schnell, T., Theunissen, E., and Rademaker, R. *Human Factors Test & Evaluation of an Integrated Synthetic Vision and Sensor-Based Flight Display System for Commercial and Military Applications*, paper presented at the NATO

Research and Technology Organization, Human Factors and Medicine panel workshop entitled "Toward Recommended Methods for Testing and Evaluation of EV and E/SV-Based Visionic Devices", Williamsburg, VA, USA, 26-27 April 2005.

5. Arthur, J.J., Kramer, L.J., and Bailey, R.E., and Prinzel, L.J. III, *Flight Test Comparison Between Enhanced Vision (FLIR) and Synthetic Vision Systems,* In Proceedings of SPIE, Enhanced and Synthetic Vision 2005, Editor: Jacques G. Verly, Volume 5802, Paper 3, April 2005.

6. Parrish, R.V, Busquets, A.M., Williams, S.P., & Nold, D.E. (2003). *Evaluation of Alternate Concepts for Synthetic Vision Flight Displays with Weather-Penetrating Sensor Image Inserts During Simulated Landing Approaches.* NASA TP-2003-212643, October 2003.

7. Harrah, S.D., et al. *The NASA Approach To Realize A Sensor Enhanced-Synthetic Vision System (SE-SVS).* paper presented at 21st IEEE / AIAA Digital Avionics Conference, October 2002.

8. Uijt De Haag, M., et al. *Flight Test Results of a Synthetic Vision Elevation Database Integrity Monitor.* Paper presented at SPIE's 15th Annual International Symposium on Aerospace / Defense Sensing, Simulation and Controls, *AeroSense*, Orlando, FL, April 16-20, 2001

9. McKay, D., Guirguis, M., Zhang, R. and Newman, R.L. *Evaluation of EVS for Approach Hazard Detection.* paper presented at the RTO SET Workshop on "Enhanced and Synthetic Vision System", held in Ottawa, Canada, September 2002, published as NATO RTO-MP-107.

10. Boff, K.R. and Lincoln, J.E., (Editors). *Engineering Data Compendium: Human Perception and Performance.* H.G. Anderson Medical Research Laboratory, Ohio.

11. Merrick, V.K. and Jeske, J.A., *Flightpath Synthesis and HUD Scaling for V/STOL Terminal Area Operations.* NASA Technical Memorandum 110348, April 1995.

12. Kramer, L.J., Prinzel, L.J., Bailey, R.E., & Arthur, J.J. (2003). *Synthetic Vision Enhances Situation Awareness And RNP Capabilities For Terrain-Challenged Approaches.* Proceedings of the American Institute of Aeronautics and Astronautics Third Aviation Technology, Integration, and Operations Technical Forum, AIAA 2003-6814, 1-11.

13. Glaab, L.J., Kramer, L.J., Arthur, T., & Barry, J.S. (2003). *Flight Test Comparison Of Synthetic Vision Display Concepts At Dallas/Fort Worth International Airport.* NASA Langley Research Center: NASA Technical Paper TP-2003-212177.

14. Kramer, L.J., Prinzel, L.J., Arthur, J.J., & Bailey, R.E. (2004). *Pathway Design Effects On Synthetic Vision Head-Up Displays.* SPIE. In Proceedings of SPIE, Enhanced and Synthetic Vision 2004, Editor: Jacques G. Verly, Volume 5424-08.

15. Foyle, D.C and Hooey, B.L., *Improving Evaluation and System Design Through The Use of Off-Nominal Testing: A Methodology for Scenario Development.* Proceedings of the 12th International Symposium on Aviation Psychology, pp. 397-402. Dayton, OH …

16. Ames, Lawrence L. & George, Edward J. (1993). *Revision and verification of a seven-point workload estimation scale.* Air Force Flight Technical Center: AFFTC-TIM-93-01.

17. AIAA (1993). *Guide to Human Performance Measurement.* Washington D.C.: American Institute of Aeronautics and Astronautics.

18. Vidulich, M.A. & Hughes, E.R. (1991). *Testing A Subjective Metric Of Situation Awareness.* Proceedings of the Human Factors & Ergonomics Society, 35, 1307-1311.

19. Kramer, L.J., Prinzel III, L.J., Bailey, R.E., Arthur III, J.J., & Parrish, R.V. (2004). *Flight Test Evaluation of Synthetic Vision Concepts at a Terrain-Challenged Airport.* NASA Langley Research Center: NASA Technical Paper TP-2004-212997.