Fusion of Synthetic and Enhanced Vision for All-Weather Commercial Aviation Operations

Randall E. Bailey, Lynda J. Kramer, and Lawrence Prinzel III
NASA Langley Research Center
Hampton, VA 23185
USA

E-mail: randall.e.bailey@nasa.gov

ABSTRACT

NASA is developing revolutionary crew-vehicle interface technologies that strive to proactively overcome aircraft safety barriers that would otherwise constrain the full realization of the next-generation air transportation system. A piloted simulation experiment was conducted to evaluate the complementary use of Synthetic and Enhanced Vision technologies. Specific focus was placed on new techniques for integration and/or fusion of Enhanced and Synthetic Vision and its impact within a two-crew flight deck during low-visibility approach and landing operations. Overall, the experimental data showed that significant improvements in situation awareness, 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. During non-normal operations, the ability of the crew to handle substantial navigational errors and runway incursions were not adversely impacted by the display concepts although the addition of Enhanced Vision did not, unto itself, provide an improvement in runway incursion detection.

1.0 INTRODUCTION

The United States air transportation system is undergoing a transformation to accommodate a projected 3-fold increase in air operations by 2025.[1] Technological and systemic changes are being developed to significantly increase the capacity, safety, efficiency, and security for this Next Generation Air Transportation System (NGATS). One of the key capabilities envisioned to achieve these goals is the concept of Equivalent Visual Operations (EVO), whereby Visual Flight Rules (VFR) operational tempos and also, perhaps, operating procedures (such as separation assurance) are maintained independent of the actual weather conditions. One methodology by which the goal EVO might be attainable is to create a virtual visual flight environment for the flight crew, independent of the actual outside weather and visibility conditions, through application of Enhanced Vision (EV) and Synthetic Vision (SV) technologies.

An experiment was conducted to evaluate the complementary use of SV and EV technologies, specifically evaluating the utility, acceptability, and usability of integrated/fused enhanced and synthetic vision technologies and its effect on two-crew operations. This work begins the development of an all-weather commercial aviation operations capability, approaching that which might create an EVO capability.

2.0 BACKGROUND

The Integrated Intelligent Flight Deck Technologies project, under the National Aeronautics and Space

Administration (NASA) Aviation Safety Program, comprises a multi-disciplinary research effort to develop flight deck technologies that mitigate operator-, automation-, and environment-induced hazards. Towards this objective, crew/vehicle interface technologies are being developed that reduce the propensity for, and minimize the risks associated with, pilot error to improve aircraft safety for current and future aircraft; thus proactively overcoming aircraft safety barriers that would otherwise constrain the full realization of the NGATS. Part of this research effort involves the use of SV and EV systems and other interface modalities as enabling technologies to meet these challenges, such as creating a visual flight environment for the crew, independent of the actual outside weather and visibility conditions, through application of EV/SV and associated technologies.

Synthetic Vision (SV) is a computer-generated image of the external scene topography, generated from aircraft attitude, high-precision navigation, and data of the terrain, obstacles, cultural features, and other required flight information. SV provides significant improvements in terrain awareness and reductions in the potential for Controlled-Flight-Into-Terrain incidents/accidents compared to current cockpit technologies.[2-5]

Enhanced Vision (EV) (or Enhanced Flight Vision System, EFVS) 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 (MMWR). In 2004, Section 91.175 of the US Federal Aviation Regulations (FAR) was amended such that operators conducting straight-in instrument approach procedures (in other than Category II or Category III operations) may now operate below the published Decision Height (DH) or Minimum Descent Altitude (MDA) when using an approved EFVS shown on the pilot's Head-Up Display (HUD). This rule change provides "operational credit" for EV equipage. No such credit exists for SV.

2.1 Integrated and Fused Synthetic and Enhanced Vision Systems Concepts

The complementary capabilities of SV and EV have been well-recognized[6] with the premise that[7] "the strengths of enhanced system can compensate for the deficiencies in the synthetic system and that the strengths of synthetic system can compensate for the deficiencies in the enhanced vision system." While these goals are obvious, optimal methods and capabilities are not on hand.

2.2 Previous Research

Several studies[8-10] have shown that the optimal combination of SV and EV technology likely uses the direct display of SV to the flight crew *without* direct display of EV, but instead, using EV "behind the glass" for navigation error detection, database integrity monitoring, and real-time obstacle/object detection. Image processing performs these functions automatically without intervention by the flight crew. This arrangement provides a highly usable display presentation (i.e., SV) that is impervious to the actual weather and visibility conditions, yet if un-charted obstacles, database errors or navigation errors are detected by the EV running in the background, the situation is annunciated, and almost "perfect" decision-making by the pilot occurs.[8-10] In fact, the study conclusion from Parrish[8] – "SVS concepts should *not* be implemented without incorporating image-processing decision aides for the pilot" – launched a 5 year effort at NASA and elsewhere developing SV *Systems* (SVS) enabling technologies for database integrity monitoring and EV object detection. [11-15]

While degrees of success in developing these decision aids have been met, technology for "perfect" object detection and database/navigation error detection does not yet exist. Further, there may always be gaps which may still warrant flight deck procedures and human interventions for integrity and error checks.

Research suggests that the optimal configuration of displays, controls, and formats depends upon the intended function of the system, the phase flight, and the role of the pilots[16,17] involving a balance between:

- Understanding – how the design promotes the perception of the SV and EV information, the comprehension of its meaning, and the projection of their status in the near future.[18]
- Workload – how the design minimizes the physical (e.g., use of controls, visual scanning) and mental (e.g., amount of cognitive effort necessary to generate understanding) workload for effective use and understanding of the information.

A key interconnection between the factors of understanding and workload is the concept of display clutter – i.e., an excessive number and/or variety of color and symbols, that obscures essential information, or presents distracting, disorganized and unnecessary information delaying visual detection.

Clutter warrants special consideration for EV/SV as the designer is intentionally increasing the volume of data presented to the flight crew. For the HUD, in particular, FAA certification policy[16] formalizes this conundrum in that "clutter should be minimized" yet, "essential or critical (HUD) data must always be displayed" such as that necessary for EFVS operational credit. Thus, task-critical or essential EV/SV imagery must be displayed but the information must not excessively interfere with pilots' ability to see the actual runway environment through the display or impede their perception, understanding, and use of the information. Declutter control in present EFVS implementations allow the pilots to selectively add or remove symbology or raster (i.e., SV or EV) imagery, but declutter controls also introduce pilot workload during task- or time-critical situations and raise the possibility that critical information might be inappropriately removed.

A common method of combining SV/EV information is by use of an integrated single display format with simultaneous presentation of unadulterated SV and EV content. By visual proximity, this method minimizes the visual scan and cognitive effort in integrating the disparate information. Typically, "inset" displays of EV information have been used to take advantage of the "unlimited" field-of-regard in SV information to complement a limited field-of-regard EV sensor. Integration in this manner allows the user to retain comparison of the separate sources for improved understanding. However, careful design of the integration process must be made to prevent an unintended *loss* of understanding. Pilot-control of control of the inset "opaqueness" by averaging or "alpha-blending" has *not* been found to be optimal.[10] Image degradation of the blended image was the primary factor.[17,19] A pure pixel-averaging technique may create a situation where a poor quality (low content) sensor image obscures good synthetic data without adding any valued to the image; conversely, a good quality sensor image is obscured by uncorrelated synthetic data. More sophisticated blending methods are available to minimize this problem,[17] but pilot control introduces the costs of workload for control, a higher potential for miss-setting of the controls, and the clutter primarily negates its utility.[10]

Display clutter can be mitigated by *not* displaying both EV and SV on the same display simultaneously using:

- Spatial separation – by locating SV and EV information on different displays. This process forces the pilot to look across displays and mentally perform the information integration. This methodology has been demonstrated under NASA flight test, wherein SV information was presented on a head-down primary flight display, and EV information was presented on the HUD. The pilots transitioned to the HUD at a specific height above the landing runway. This methodology was found to be acceptable although not necessarily without improvements being desired. Sufficient information must be on the HUD to restrict the need for the Pilot Flying (PF) to go "head-down" to acquire task critical information; otherwise, HUD usage can create "a loss of situation awareness" and additional workload from visual scanning. Designs utilizing visual momentum assist in the integration.[19]
- Temporal separation – by displaying *either* SV or EV information on the same display, using

automatic or manual selection for which source is displayed. An automatic transition has been tested and it felt "natural" to the pilot.[20] Without retaining some elements of SV or EV display, however, "visual momentum" and the complementary benefits of both EV and SV may not be lost.

Using a single display, image fusion can reduce clutter. Fusion, in this case, creates an image wherein the SV and EV sources are consolidated to the extent that the sources are indistinguishable. The drawback is the loss of understanding from source comparisons and contrasts. The specific fusion methodology significantly affects the utility of the resultant display.[17] Feature-level fusion methods generally enabled higher image contrast while retaining the source information; however, it may also retain some undesirable noise content of the sensor inputs. Registration of the two image sources is of paramount importance for fusion.

Another method of EV/SV fusion displays SV-derived information symbolically, using symbols, icons, and/or perspective shapes for conformal overlay with EV imagery. For instance, SV-derived symbology, in the form of runway outlines, with simulated EV imagery provided visual momentum yet didn't create any perceptual confusion even when flying with significant navigational errors inserted into the SV-derived guidance.[21-22]

2.2.1 Pilot-Flying Role

EV/SV integration on head-down displays permits the use of color to differentiate and declutter the respective data attributes.[23] In addition to color, head-down displays do not require conformality but its limitations must be recognized.[24] The head-down to head-up transition for the landing task is critical, however, and the literature shows a clear, consistent preference for flight information to be displayed head-up during approach and landing operations for the pilot-flying (PF).[6] Further, the HUD provides many other advantages due to its conformal display of symbology and imagery.[25-27] However, its display characteristics – monochromatic, variable lighting background, and limited luminance capabilities – introduce constraints.[23]

The presence of EV/SV imagery (i.e., terrain) for the PF has *not* been found to be a significant factor in flight path performance when compared to the overwhelming influence of flight guidance symbology[6,28] unless, however, the guidance information is obscured or hindered by the presence of SV and/or EV imagery.[22] In this latter case, performance degrades in the presence of EV/SV information. More typically, terrain imagery has been regarded as supplemental information that complements guidance information or the EV/SV imagery provides required visual references allowing certain tasks or operations, such as EFVS.[29]

The complementary use of EV/SV information for the PF to perform navigation integrity checks and obstacle/object detection has been evaluated in simulation and, in general, performance is relatively mediocre.[8-10] The results appear to be influenced by the EV/SV integration and fusion methods employed but conclusive cause-and-effect is not available.

In an evaluation of an integrated SV/EV Head-Down Display (HDD) concepts[8] (using an EV inset), four pilots (~10% of the non-normal trials) flew to touchdown on the synthetic runway image, essentially ignoring a lateral navigation error in the EV (actual) inset image. Similarly, when presented with runway incursions, 3 pilots missed the incursions (~12.5% of the non-normal trials) when using this same display concept.

EV imagery may provide an element of traffic detection above and beyond unaided natural vision, but several factors must be considered to realize this capability. First, imaging sensors are not a panacea. Detection and recognition is dependent upon the sensor resolution and its sensitivity to the environment and "target", the size and contrast of the obstacle/object against the image background, and the display resolution.[6,30-31] Second, if the PF display uses flight vector information, the flight path marker symbol and guidance symbol, if provided, may obscure an object.[16] Finally, the pilot-flying's attention is primarily focused on the landing task, and the PF may not have sufficient attentional resources for "unexpected" scenarios. This task has been rated as "unacceptable" and resulted in unacceptable workload.[9]

2.2.2 Pilot-Monitoring Role

While the pilot-flying is primarily focused on flying the aircraft, the pilot-not-flying (PNF) or pilot-monitoring in a two-crew operation *does* have the primary responsible for monitoring and verifying flight path performance, cross-checking guidance and raw data indicators, and identifying the visual runway environment and clearing the landing area.[31] Unfortunately, like the PF data, the use of EV/SV information by the PNF to perform the navigation integrity, obstacle detection, and runway incursion detection is relatively mediocre. The results, again, appear to be influenced by the EV/SV integration and display methods employed.

During low-visibility auto-landings[10] where most included an "anomaly" consisting of SV database misalignment, runway incursions, and uncharted obstacle, piloted evaluations showed that:

- Using an EV-only HUD, 57% of the obstacles were undetected and 50% of the runs with the database misalignment went undetected. 38% of the runway incursions went undetected. Despite this relatively poor performance, the pilots subjectively preferred the Head-up implementation.

- A Primary Flight Display (PFD) concept, showing SV with a pilot-controllable EV inset image, exhibited the worst performance for anomaly detection. The PFD inset concept missed 100% of the uncharted obstacles - all leading to hazardous situations and a 33% rate of missed detections for the database misalignments. Detection of runway incursions showed a missed detection rate of 33%. The poor performance and low-acceptance of the head-down EVS insert concept was attributed to the clutter of the image, the small EV image size, and confusion between the SV/EV images.

- With EV-only shown on a HDD, all uncharted obstacles were successfully detected (but 20% of runs continued into hazardous conditions because the evaluation pilots (EPs) didn't recognize the severity of the threat) yet 33% of the runs missed the misalignment. A missed detection rate of 22% for runway incursions was found. This concept was well liked for its large image size and minimal display clutter, but disliked because of the workload in transitioning from head-down to head-up flight.

Direct research into the effects of EV/SV usage on crew resource management and crew coordination, in a two-place cockpit, has not been studied extensively. Crew coordination in an EVS approach operation was the primary focus of one study,[32] but sensor-unique (in this case, a millimeter wave radar) issues of range/azimuth versus elevation/azimuth imagery presentations predominated the work. Anecdotal evidence from a flight demonstration program suggests that "a centrally-located display of EVS imagery" may facilitate CRM and training.[33]

2.3 Experiment Objectives

An experiment was conducted to evaluate the complementary use of SV and EV technologies, specifically focusing on new techniques for integration and/or fusion of synthetic and enhanced vision technologies and CRM. The objective of this experiment was to test the utility, acceptability, and usability of integrated/fused EV and SV technology concepts in a two-crew commercial or business aircraft cockpit.

The experiment leverages off of the new operating rules available under Section 91.175(l) of the US FARs for conducting instrument approach using an approved EFVS. Under these rules, EV is presented head-up to the pilot-flying (Captain) whose primary responsibility is flight path control, flight path maintenance, recognition and identification of the EV "visual references" under 91.175(l), and recognition and identification of the "visual references" under 91.175(l)(4), without reliance on the enhanced flight vision system. In this operation, the PF remains in control of the aircraft through the approach and landing.

The experiment assumes a "best practices" approach from the available literature to bound the experimental scope:

- While EV might improve SV operations, the converse warrants investigation as well. Accordingly, these complementary roles are investigated.

- Previous research shows that navigation integrity and obstacle clearance checks by the PF carries a workload and performance burden. These should be considered secondary tasking. The HUD formats are designed to emphasize the PF role for aircraft control, and the recognition/identification of the EV "visual references." HUD clutter is minimized without additional PF controls.

- Object detection by the PNF has previously been found to be best[10] using a dedicated EV display. The tested EV display did not include symbology – thus, minimizing clutter and aiding detection – but the test participants recommended that symbology be added to aide performance monitoring.

3.0 METHODOLOGY

3.1 Experiment Method

A fixed-based simulation experiment was conducted to evaluate the effect of adding synthetic vision information to an enhanced vision HUD for the PF during low-visibility approach and landing operations. In addition, the experiment evaluated the effect of adding synthetic vision information and symbology to the PNF's display of the EV sensor data. A two-crew experiment was conducted to assess the interaction of these display concepts on crew resource management.

Twenty-four pilots, representing seven airlines and a major cargo carrier, participated in the experiment. All participants had previous experience flying with 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. EV experience was not required although some pilots were familiar with imaging sensor technology from prior military flight experience. None of the subjects were currently flying EV in operations.

3.2 Simulator

The experiment was conducted in the Integration Flight Deck (IFD) simulation facility (see Fig. 1) at NASA Langley Research Center (LaRC). The IFD emulates LaRC's Boeing B-757-200 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 resolution.

The participants occupied the left (as PF) and right (as PNF) seats. The left seat included an overhead HUD projection unit and the right seat included an auxiliary display (AD) under the right side window (see Fig. 1).

3.3 Head-Up Display

The HUD subtended approximately 32° horizontal by 24° vertical field of view. The HUD presentation was written strictly in a raster format from a video source (RS-343) input. The input consisted of a video mix of symbology and computer-generated scene imagery (either EV or SV as described in Section 3.10.1). The symbology included "haloing" to ensure high-contrast symbology 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).

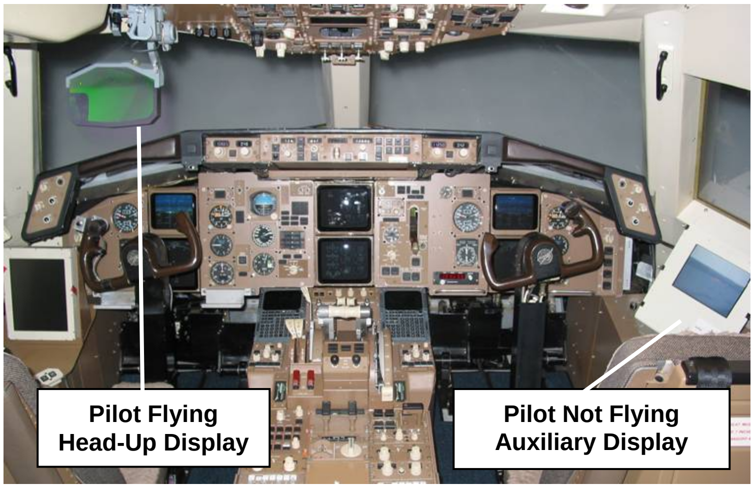


Figure 1. Integration Flight Deck Simulation Facility with HUD and AD.

3.4 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 "haloed" symbology and computer-generated scene imagery (either EV or the output of a fused EV/SV signal as described in Section 3.10.2).

3.5 Head-Down Displays

Minimal changes were made to the Primary Flight Display (PFD) and Navigation Display (ND). 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. The ND showed the Flight Management System approach but did not include Enhanced Ground Proximity Warning System or Traffic Alert and Collision Avoidance System information.

3.6 Synthetic Vision Database

An SV database was created from a 1 arc-sec Digital Elevation Model (DEM) of a 53 by 53 nm square area centered around the Reno-Tahoe International Airport. The airport was represented by three-dimensional models of the runway, taxiways, and terminal buildings. The DEM was draped with 1 meter/pixel satellite photographic imagery within a 16 x 21 nm area centered around the airport and 4 meter/pixel imagery outside.

3.7 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.

3.8 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 database. 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 in Section 4.3. The EV imagery was provided in 640 horizontal by 480 vertical pixel resolution.

3.9 Symbology

3.9.1 HUD Symbology

The HUD format shown in Figure 2. Vertical and horizontal path deviation was always shown by linear path deviation indicators (i.e., "dog-bones") simultaneously with the standard instrument landing system (ILS) (angular) glideslope and localizer "raw data" indicators. One "dot" of linear path deviation equaled 87.5 ft vertical and 150 ft lateral path error.

The pitch-roll guidance cue ("ball") used modified pursuit guidance[34] along the desired path centerline, 5.5 seconds ahead of ownship. Horizontal and vertical position of the ball corresponded to the track and flight path angles to fly to the center of the desired path.

Depending upon the experimental condition, the following symbology elements were also used, as described in Section 3.10.1 A glideslope reference line was drawn (Fig. 3) at the RNO Runway 16R ILS descent angle of -3.1 degrees. Also, a runway outline symbol was drawn to conformally position the symbol based on the threshold coordinates of Runway 16R/34L based on the simulated aircraft navigation position. The symbol portrayed an 8000 ft long by 200 ft wide runway, consistent with certified Head-up Guidance Systems..

3.9.2 HUD Tunnel

As an experiment variable, advanced pathway guidance in the form of a "minimal" tunnel was flown (see Fig 2). 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 used to minimize HUD clutter. Past studies[2,28] have shown that sufficient path information is provided by the minimal tunnel concept – at a minimal cost of display clutter – when path deviation indicators, guidance symbology and the FPM are also provided.

3.9.3 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 image. The symbology included digital readout of indicated airspeed and altitude; zero pitch attitude line (horizon line); 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.) Alternative symbology sets to clear the center of the display were tried, but the pre-test usability results were not encouraging enough to move forward with the concepts.

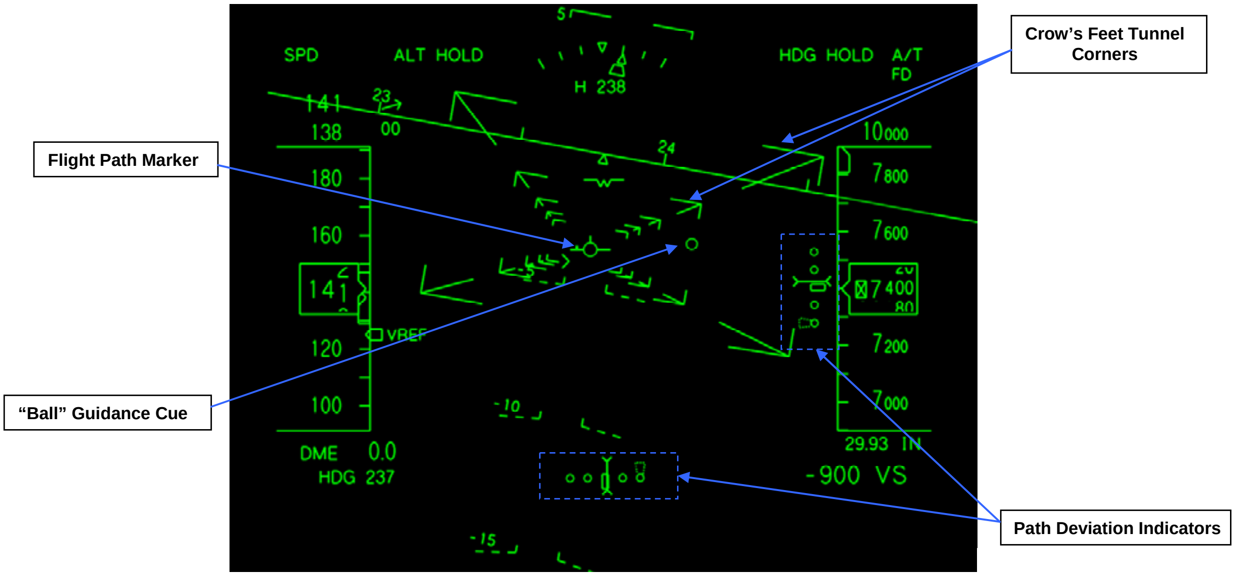


Figure 2. HUD Symbology with Advanced Pathway Guidance.

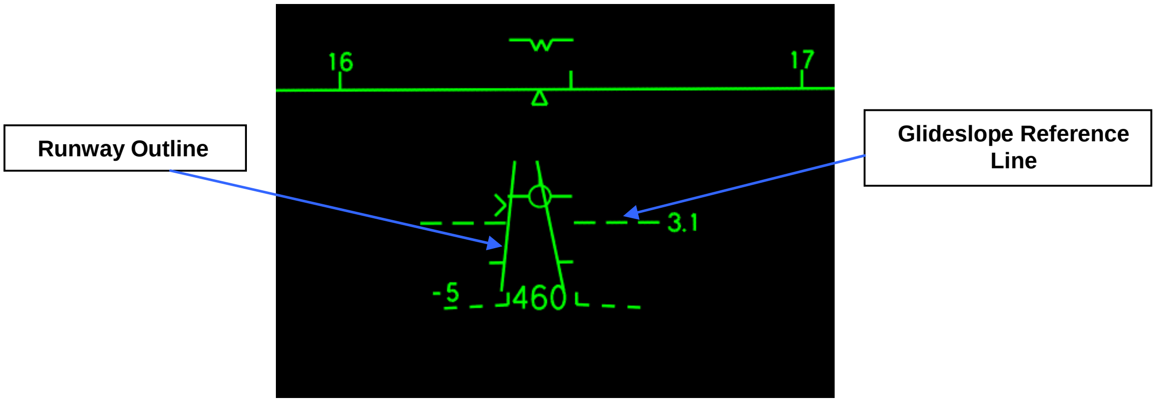


Figure 3. HUD Runway Outline Symbol and Glideslope Reference Line.

3.10 Display Concepts

Four HUD concepts and four AD concepts were evaluated by the evaluation crew (PF and PNF) while flying approaches to Reno-Tahoe Airport, Runway 16R. The head-down display formats were invariant.

3.10.1 Head-Up Display Concepts

The four HUD concepts were differed from each other with respect to: 1) the type of raster background presented; and, 2) in the type of symbology presented.

Two raster (background) formats were flown:

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

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

Each raster concept showed FLIR alone below 500 ft to enable the operational credit now offered by EFVS. The 500 ft transition altitude was chosen from a usability study prior to the test and flight experience[2] as the altitude: 1) after which FLIR would be required, 2) with sufficient time to become assimilated to the FLIR, and, 3) at or just after the recommended minimum stabilized approach altitude to allow full utilization of SV.

The "fusion" concept provides the basis to evaluate the utility, acceptability, and usability of SV and EV on the HUD. This methodology was picked as a "best practice" from the literature review as it maximizes image legibility and minimizes image confusion and clutter for the pilot. In this methodology, SV and EV are shown in the flight regimes where they are most advantageous to the PF, and pilot-controllable fusion methods are not implemented as the PF is already over-burdened with controlling the HUD brightness, contrast, declutter, and EVS control.

Two symbology sets were flown:

1) Standard HUD symbology (hereafter referred to as "Baseline"). The glideslope reference line was always drawn.

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. 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 glideslope reference line was drawn and a runway outline was projected until reaching 50 ft AFL.

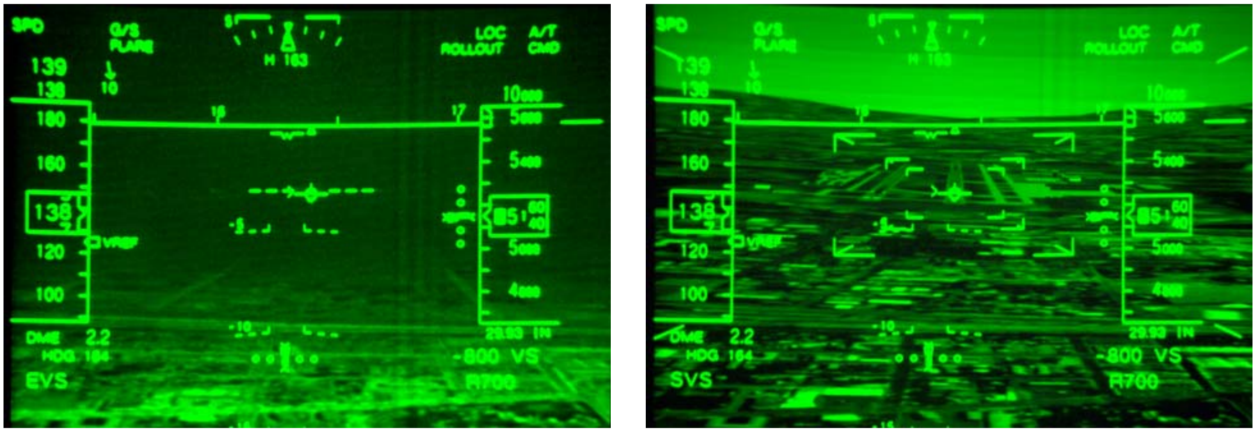
The glideslope reference line is part of both the Baseline and Tunnel symbology sets as it is identified as a required symbology element under FAR §91.175 EFVS operations.

In Figure 4, two of the concepts are shown - the FLIR-Baseline HUD and the Fusion-Tunnel HUD. In the FLIR-baseline concept, a minimum of symbology is used and the FLIR does not necessarily provide terrain and runway cues, depending for instance, upon the atmospheric conditions. Conversely, the Fusion-Tunnel concept uses tunnel guidance for distinct path demarcation and SV for clear terrain and runway references, above 500 ft AFL. Below 500 ft with the Tunnel symbol set, the runway outline provides an element of SV visual momentum within the EV raster background image. The tunnel is decluttered at 500 ft AFL to minimize clutter and is replaced by the glideslope reference line to ensure approach path angle awareness. Below 500 ft, the only additional symbology in the Tunnel symbology set, over that of the Baseline set, is the runway outline symbol. Below 500 ft, the FLIR-Baseline and Fusion-Baseline configurations are identical.

3.10.2 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 symbology was either “On” or “Off.” When present, the AD symbology was a subset of the standard HUD symbology (see Section 3.9).

The raster was either: 1) EV only (hereafter referred to as “FLIR”); or, 2) a fused SV/EV image (hereafter referred to as “Fused”).



HUD Concept
EVS (FLIR) Only – Baseline

HUD Concept
Fusion – Tunnel

Figure 4. Head-Up Display (HUD) Formats

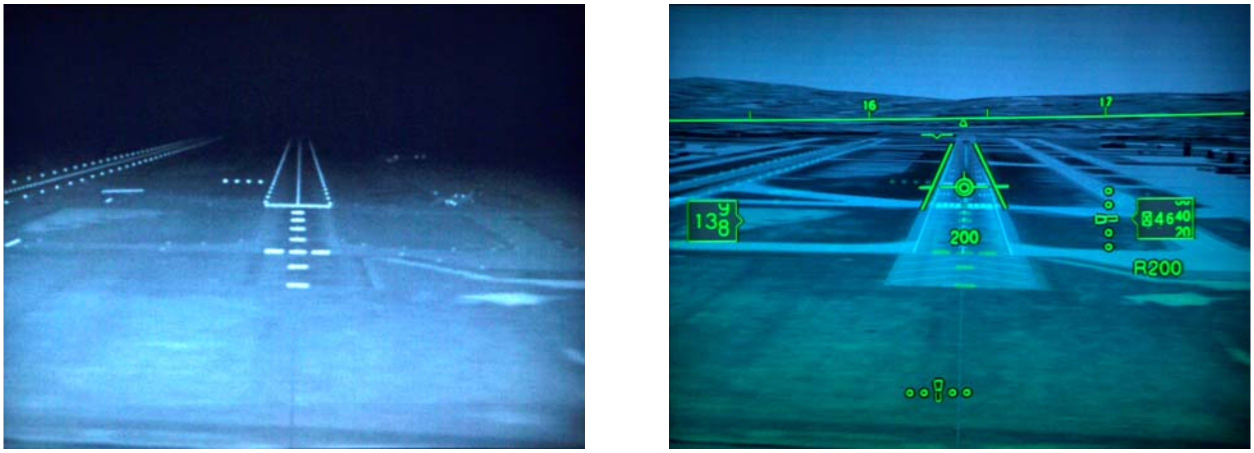
The AD fused raster imagery was pilot-controllable and could be tuned throughout the approach to one of 10 states: FLIR only, SVS only, or 8 fusion combinations of FLIR and SVS, using an Equinox EP-3000™ fusion board. The fusion board employs a feature-level extraction algorithm with two pilot control inputs. The first control biased the feature level fusion through 8 weighting values weighting EV or SV. (A value of 1 biased the extraction to 11% FLIR and 89% SV whereas a value of 8 weighted the extraction to 89% FLIR and 11% SV). 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 maximal 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.

In Figure 5, 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 SV and EV 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.)

4.0 EXPERIMENT

4.1 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 air traffic control 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 selected evaluation task tests the ability of SV and EV technology to support this type of operation by providing "equivalent visual" information into the cockpit. Further, if this technology succeeds in providing a visual arrival capability, the potential for operational efficiency above and beyond what can be provided by RNP may be offered.



AD Concept
EVS (FLIR) Only – No Symbology

AD Concept
Fusion – Symbology

Figure 5. Two Auxiliary Display (AD) Formats

The Pilot Flying (PF) hand-flew the base and final leg portions of the Sparks Visual Arrival to Runway 16R 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) for this curving, descending approach. The path converged into the ILS for Runway 16R. When properly executed, the aircraft was configured to land and established on a stabilized approach by 1000 ft AFL. The PNF monitored the approach using standard instruments and the AD. The crew was 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 or if the required visibility per §91.175 was not met.

4.2 EVS Crew Procedures

EFVS crew procedures, adapted from those used currently in business aircraft EFVS operations, were used. Instructions and training in the use of the procedures were given to each crew. These procedures are given in Table I, including automatic cockpit aural call-outs. The automatic call-outs were set up assuming a 200 ft DH for the published, non-EVS approach.

By the published minimums of 200 ft DH, the crew procedures dictate that the PF must have the required EFVS 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 Runway 16R provided the prominent EVS references. If these EFVS references were visible, the PF was instructed to call "EVS Lights".

Unaided by the EFVS, if the PF saw the lights or markings of the threshold (the predominant landing visual reference for Runway 16R), the PF called "Landing." The "landing" call was required no lower than 100 ft AFL.

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 AFL or if "Landing" was not called by 100 ft AFL. 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 to, while other procedures were consistent with, their current airline Standard Operating Procedures (SOPs). In either case, the crew procedures were trained and "enforced." 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.

Table I: 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 EVS Visual Cues,** Call "*EVS Lights*"	**When Visual Cues Appear** Call "*Lights*" or "*Field in Sight*"
			Without EVS Visual Cues, Call "*Going Around*"	**Without PF Call of 'EVS Lights'**, Call "*Go Around*"
EVS Decision Altitude (100 ft AFL)	100 / 4512		**When Actual Visual Cues,** Call "*Landing*"	**When Visual Cues Appear,** Call "*Lights*" or "*Field in Sight*"
			Without Actual Visual Cues, Call "*Going Around*"	**Without PF Call of 'Landing'**, Call "*Go Around*"

4.3 Experiment Matrix

Nominally, forty experimental runs were completed by each crew with each pilot flying 20 approaches evaluating the HUD concepts and 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 EFVS 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. The PF was also instructed to land as close as possible to the runway centerline.

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.[34] The RI scenarios simulated an incursion with either a non-transponding baggage cart or fire truck. The database integrity 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 SV terrain, pathway, runway outline, if shown, and guidance cue being misaligned from the FLIR, OTW, and ILS (which were all defined as being correct).

4.4 Measures

During each run, path error, pilot control inputs, PNF head-direction, and PNF-AD 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[36], Situation Awareness Rating Technique (SART)[37], and four Likert-type (7-point) questions specific to different constructs of display clutter.

After data collection was completed, pilots were administered two separate Situation Awareness – Subjective Workload Dominance (SA-SWORD)[38] and Subjective Workload Dominance (SWORD)[37] 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.

5.0 RESULTS

5.1 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 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 and 100 ft AFL. (The non-normal runs with a lateral navigation error were not included in the final approach segment analyses.)

As expected, there were no significant differences ($p>0.05$) 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. Once on final, the only difference between the configurations was the presence or absence of the runway outline. This symbology element should be, and was found to be, inconsequential to flight performance.

On the approach segment, an ANOVA ($F(3,443)=9.73$, $p<0.01$) and post-hoc tests revealed that lateral flight performance with the FLIR-Baseline (mean = 42 feet) and Fusion-Baseline (mean = 37 feet) was significantly worse than the FLIR-Tunnel (mean = 30 feet) and Fusion-Tunnel (mean = 28 feet) HUD concepts. For vertical performance, an ANOVA ($F(3,443)=6.69$, $p<0.01$) 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; and Fusion-Tunnel = 9 feet.

While these results show statistical significance, the operational significance of these differences is questionable. To tease out the possible operational differences between configurations, the lateral and vertical tracking error data on approach was analyzed in terms of Flight Technical Error (FTE) as part of Required Navigation Performance (RNP).

Lateral and vertical path error was cast into a histogram analysis for the nominal Sparks 16R approach runs for the four HUD concepts (FLIR-Tunnel, FLIR-Baseline, Fused-Tunnel, and Fused-Baseline). The horizontal and vertical path steering error components of the RNP calculation include both FTE and display error. For this analysis, it was assumed that display error was negligible, so FTE was the only component of path steering error. It was also assumed that the other two components (path definition error and position estimation error) of the horizontal RNP calculation would be equivalent across the display concepts evaluated. Similarly, it was also assumed that the other three components (altimetry system error, vertical path definition error, and horizontal coupling error) of the vertical RNP performance calculation would be equivalent across the display concepts evaluated.

With these assumptions, all HUD concepts yielded a horizontal FTE navigational accuracy of 0.02 nmi at least 95% of the time. No statistically significant differences were found between display concepts using this FTE analysis. For vertical FTE, it was assumed that the pilot was flying a specified vertical profile so that the required vertical navigation performance accuracy was 300 feet.[39] With these assumptions, all HUD concepts yielded a vertical FTE navigational accuracy of 300 feet at least 99.7% of the time. In fact, the Fused - Tunnel, Fused-Baseline, and FLIR-Tunnel concepts yielded a vertical FTE navigational accuracy of 80 feet at least 99.7% of the time. The FLIR-Baseline HUD concept yielded a vertical FTE navigational accuracy of 120 feet at least 99.7% of the time.

Overall, the path control performance results are consistent with past research[2,40] that showed HUD guidance concepts, using flight path-centered symbology, can enable manual RNP operations with lateral FTE of 0.05 nmi. 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 felt that the tunnel provided good turn anticipation cues and the SVS background, when present, also improved flight path control performance because the database imagery in the background provided stronger roll reference visual cues. These features also limited a tendency to overcontrol in roll in the fixed-base B-757 simulator, particularly when compared to flying the baseline symbology set (i.e., compensatory guidance symbol only). These differences are subtly apparent in the data. For instance, it could be hypothesized that the pilots with a Tunnel and/or SV information in the raster background were better able to attend to the dual task of vertical and lateral path tracking. In any case, the pilots were able to easily meet operational performance requirements, such as FTE for RNP, flying manually.

5.2 Mental Workload

5.2.1. PF HUD

In the post-run data, the main factors of HUD raster (F(1,366)=4.47, p=0.035) and symbology (F(1,366)=25.06, p<0.01) were significant for workload. However, the raster by symbology interaction was not significant (p>0.05). Post-hoc tests showed the pilots rated the Fusion-Tunnel HUD concept as having significantly less workload than the other 3 HUD concepts tested: 1) Fusion-Tunnel (mean = 3.1) and 2) FLIR-Tunnel (mean = 3.4); Fusion-Baseline (mean = 3.5) and FLIR-Baseline (mean = 3.6). 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.”

In the post-test data, SWORD data indicate that there were no significant (p<0.05) differences among the HUD concepts for the SWORD ratings of mental workload.

Overall, the data suggests weak differences, if any, in the workload associated with the PF-HUD concepts as the only significant differences (i.e., the Fusion-Tunnel concept reduced PF workload) reduced the average workload toward “easily managed”, whereas the other concepts elicited workload ratings tending toward “challenging but manageable” workload levels. Pilot commentary suggested that the workload when flying the tunnel symbology concepts was easier (less scan between the HUD and ND for path awareness; easier to anticipate the turns), but the differences were not of a magnitude to warrant concern.

5.2.2 PNF-AD

In the post-run data, there were no significant ($p>0.05$) 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, meaning that the workload was rated between "Light Activity; Minimum Demands" and "Moderate Activity – Easily Managed; Considerable Spare Time."

In the post-test data, SWORD data show that the AD concept was highly significant ($F(3, 69)=15.02$, $p<0.001$). Post-hoc tests (SNK using $\alpha=0.05$) showed that the lowest workload was associated with the FLIR-Symbology and Fused-Symbology (no discrimination between); followed by the Fused-No Symbology; and, 3) FLIR-No Symbology (highest workload).

Pilot commentary typically noted the advantage of symbology in reducing the visual scan and cognitive task of integrating the different display information. Further, the presence of SV and EV also provided workload reduction, indicating that the physical workload increase due to manipulating fusion controls was more than offset by the reduction in mental/cognitive workload. The location of the PNF-AD was often noted as being sub-optimal. It was too far away from the forward field of view / instrument panel.

5.3 Situation Awareness

5.3.1 PF HUD

From the post-run SART data, an ANOVA revealed that both HUD raster type ($F(1,366)=3.23$, $p<0.01$) and symbology type ($F(1,366)=38.10$, $p<0.01$) and the interaction between these factors ($F(1,33)=4.22$, $p=0.04$)) 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.

From the post-test SA-SWORD data, an ANOVA revealed that the HUD concept was also highly significant ($F(3, 69)=43.61$, $p<0.001$). Post-hoc tests showed three unique subsets for situation awareness ratings with the HUD concepts: 1) Fusion-Tunnel (highest SA); 2) Fusion-Baseline & FLIR-Tunnel; and 3) FLIR-Baseline (lowest SA).

The SA data and pilot commentary consistently reflect 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. Also, the SVS component in the Fusion HUD concept provides significant terrain information unavailable in any other cockpit displays. Interestingly, by the SA measures, the Fusion-Baseline & FLIR-Tunnel concepts were not significantly different. This result would imply that the pilots felt the SV contribution to SA was essentially equivalent to the tunnel contribution.

5.3.2. PNF-AD

An ANOVA on the post-run SART ratings for PNF display found no significant results for the main effects (raster, symbology) or the interaction ($\underline{p}>0.05$). However, the post-test SA-SWORD data showed that the PNF-AD concepts were highly significant ($F(3, 69)=37.78$, $p<0.001$). Post-hoc tests showed three overlapping subsets for SA-SWORD with the 4 AD concepts: 1) Fused-Symbology (highest SA); 2) FLIR-

Symbology & Fused-No Symbology; and 3) Fused-No Symbology & FLIR-No Symbology (lowest SA).

The benefits of Fused imagery and symbology on the PNF-AD emerged in the post-test SA data, but no statistically significant differences were noted in SA post-run using SART. 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 in general for a PNF in this type of operation.

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.

5.4 Head-Tracking Analysis

A head-tracker was used to quantify what display the PNF was using and when. An eye tracker would have been preferable for this task, but its installation was impractical. The head-tracker was unable to differentiate whether the PNF was looking OTW or at the instrument panel HDDs because very small head-movements, if any, were used by the PNFs to look between the HDDs and OTW. But the head-tracker could very effectively distinguish when the PNF was looking at the AD or at the OTW/HDDs.

In Figure 6, the percentage of time that the PNF looked at either the AD, the OTW/HDD, or elsewhere is shown, broken down by the aircraft altitude (in feet) Above Field Elevation (AFL). Simple ANOVAs were highly significant ($p<0.001$) within each altitude bins and post-hoc tests showed 3 unique subsets across the head locations (Forward, AD, or Other). The data clearly indicates that the pilot was using the AD only about 14% of the time (collapsed across PNF-AD concept) above 1000 ft but increased their usage (up to 30% of the time around 400 to 500 ft AFL). Below 400 ft, usage quickly reduces to less than 10% of the time below 200 ft AFL. The PNF was mostly looking forward (HDD/OTW) below 200 ft (~90%) and pilot commentary suggested that the PNFs was looking OTW for the runway visual references.

In Table II, the percentage of time the PNF spent looking at the AD is shown depending upon the AD display concept. Simple ANOVAs were conducted across the different auxiliary display concepts (FLIR, with and without symbology and Fused, with and without symbology). Every altitude segment, except the Alt < 100 feet segment, showed significant differences for the auxiliary display concept.

The data shows that on approach, above 500 ft AFL, pilots focused on display concepts with symbology more so than those with terrain information (SV or EV). (Given the simulated weather conditions, the FLIR did not provide any terrain information above 1000 ft AFL.) For instance, the FLIR-Symbology configuration induced more gaze on the part of the PNFs than the Fused-No Symbology configuration (terrain information but no symbology). This result is likely due to the PNF-AD symbology which included the linear path error – this information was not readily available on any other PNF display.

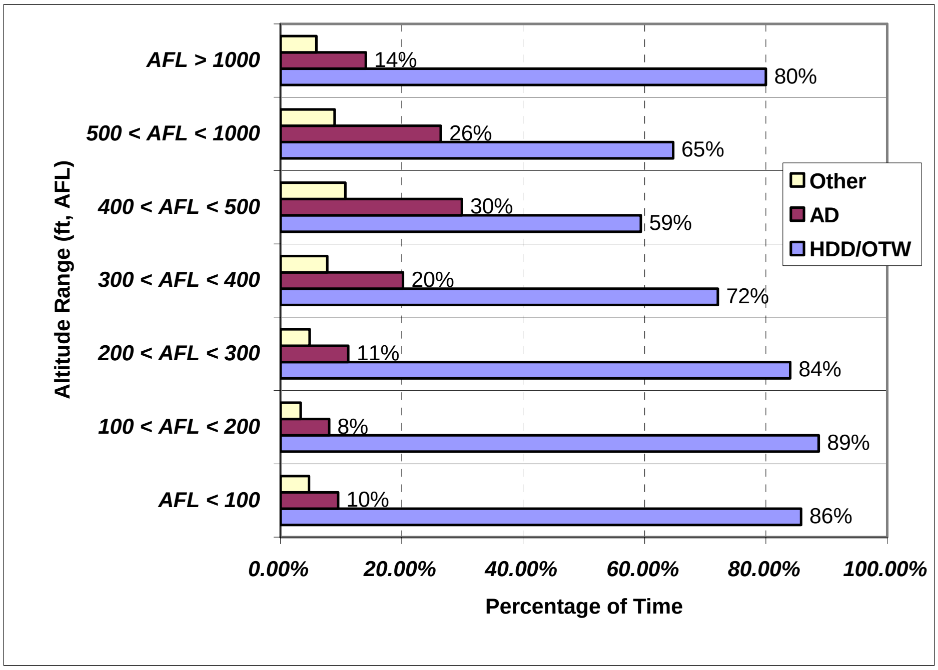


Figure 6: Percentage of Time Where PNFs Directed Gaze

Between 500 and 200 ft AFL, the usage of all the display concepts increased. The presence of symbology primarily and SV secondarily induced attention to the AD. While statistical significance is shown in Table II below 200 ft, the differences in display concept are operationally insignificant. (10% percentage represent approximately 1 second of time looking at the PNF between 100 and 200 ft AFL.)

Table II. Percentage of Time PNF Directing Gaze, By Display Configuration

AD Concept	Altitude Range (ft, AFL) AFL > 1000	500 < AFL < 1000	400 < AFL < 500	300 < AFL < 400	200 < AFL < 300	100 < AFL < 200	AFL < 100
Significance?	4 Unique Groups	4 Unique Groups	2 Unique Groups	2 Unique Groups	2 Overlapping Groups	2 Overlapping Groups	None
FLIR - No Symbology	4%	18%	24%*	15%*	9%*	6%*	9%
Fused - No Symbology	13%	22%	24%*	15%*	10%*	7%*†	10%
FLIR-Symbology	16%	30%	36%†	25%†	12%*†	9%†	11%
Fused - Symbology	23%	35%	36%†	27%†	14%†	9%†	9%

* Member of Group 1 † Member of Group 2

By providing a plethora of controls to the PNF for the Fused AD concept, the experiment data provided firstorder 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. No “guidance” was provided to the pilots for optimal usage. Instead, the system operation was explained and they were given the training runs to “play” with the controls. After training, the pilots were allowed to use whatever control settings that felt necessary and appropriate. The EPs quickly learned how the fusion worked and what the most effective means to employ the controls were.

The percentage of time that the PNF used any fusion settings (i.e., SVS-Only or feature-level fusion values, “SV”) or “EV-Only” when the “Fused” AD concept was flown is plotted in Figure 7. The data shows that, on the approach, fusion or SV information was displayed more than 85% of the time. Below 500 ft, in the final approach segment to landing, EV-only was used 60% of the time and Fusion (SV) reduced to 40%. The 60-40 distribution in EVS-only and “Fusion” settings, respectively, suggests that the PNF used both information sources cooperatively and effectively. The common PNF strategy was cycle between EV-only and the highest level of SV/EV fusion. They could do this quickly because they were adjacent switch positions.

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) as a “data available” cue. That is, with color-coding enabled, when color appeared on the PNF-AD, this cued the pilot that the EVS was starting to show useful information. The color signalled that they could effectively begin using an “EVS-Only” setting.

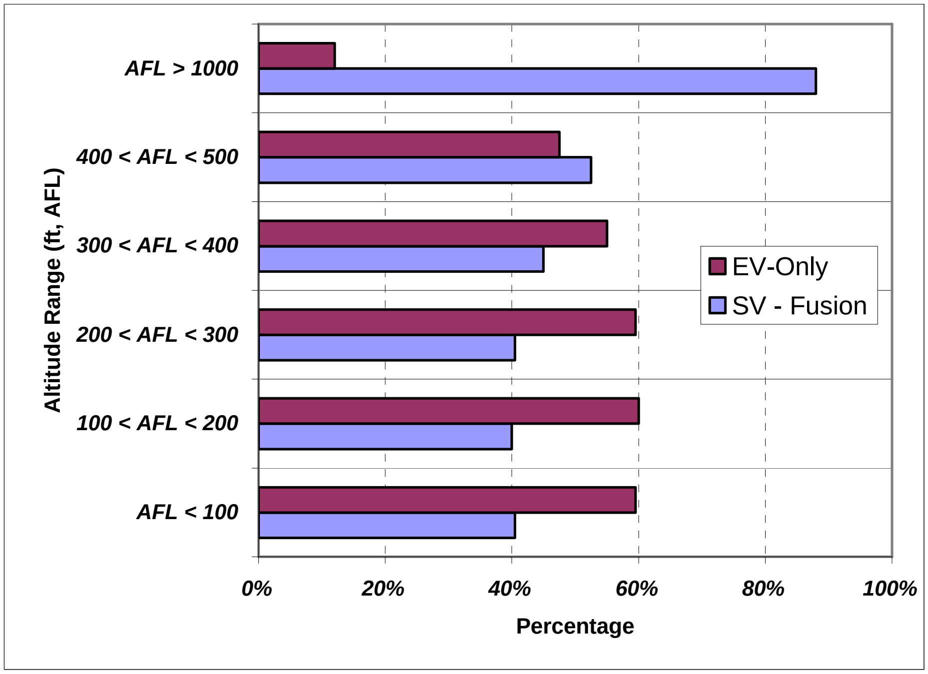


Figure 7. SVS/EVS Fusion and EVS-Only Setting by Altitude.

5.6 Subjective Assessments of Display Clutter

After an experimental run, each pilot gave ratings for the 4 Likert-type questions on display clutter 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 ($F(1,363)=28.89$, $p<0.01$) 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.

5.7 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.

5.7.1 Runway Incursions

The runway incursions were represented by a baggage cart and a fire truck. Both vehicles were positioned in the same location, approximately 850 ft from the RNO Runway 16R landing threshold and just slightly offset from the centerline. They were both positioned perpendicular to the runway (i.e., they were facing toward the runway edgelines.) 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 small size. This ordering tested for "just noticeable differences" for runway incursion detection.

To put the runway incursion (RI) results into perspective, the EVS and OTW (simulated) visual scene resolutions were used to compute the altitude AFL that the RIs might be reasonably expected to occur. While target detection is critically dependent upon visibility, lighting, target contrast, color, gray scales, etc., resolution is only used in this example. It is assumed that 10 pixels (scan lines) are required for a human observer to recognize a target/object in this example.[29] The baggage cart consisted of a tug and a cart. The tug was approximately 7.5 ft tall and 10 ft long, tied to a cart 6.5 ft tall and 10 ft long. The firetruck was 31.6 ft long and 13 ft tall. An operating rotating beacon was depicted atop the firetruck.

In Table III, the height of the aircraft when 10 lines (pixels) draw the incurring vehicles is shown. The analysis assumes a 3 degree glideslope and the limiting resolution of the EVS shown on the HUD and PNF-AD was the simulated FLIR (640x480 resolution). The HUD and PNF-AD EVS resolution was 20 pixels per degree. The OTW resolution provides 26 pixels per degree. The vertical and horizontal resolutions were identical.

Table III: Altitudes Above Field Level for Theoretical Detection of Runway Incursions

Object	Dimension / Direction	AFL for EVS "Detection"	AFL for OTW "Detection"
Baggage Cart	Height	42 ft	55 ft
	Width	120 ft	156 ft
Fire Truck	Height	78 ft	101 ft
	Width	190 ft	246 ft

The theoretical detection ranges show that the Fire Truck should be detectable at almost twice the distance as the Baggage Cart, particularly its vertical extent (i.e., height). The OTW provided better detection capability because of its higher resolution. Finally, none of the vehicles were "detectable" in this analysis above 200 ft AFL using the EVS. (It should be noted however, that the Fire Truck *was* detectable above 200 ft AFL on the EVS, if the observer was cued to its existence and studied the display.)

The experimental results showed that, 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. From the analysis of Table III, the baggage cart should be "detectable" between 50 and 150 ft AFL but this leaves only 5 to 15 seconds before landing – not very much time to spot a small object, parked slightly short of the intended touchdown zone.

Eleven crews saw the Fire Truck OTW (7 by the PNF, 3 by the PF, and 1 simultaneously by the PF and PNF) and the one remaining crew saw it on the PNF-AD. Upon seeing the incursions, all crews initiated a go-around (all lower than 50 feet AGL).

The incurring vehicles were visible in the PNF-AD and HUD, yet the experimental 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, because of the lack of symbology (in some cases) and the improved gray scale contrast. But, as the analytical data shows, the vehicle size and EVS resolution made detection 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. The head tracking measurements shown above quantify this pilot comment. The PNF was head-out the vast majority of the time, ranging from 86% to 100% of the total time below 200 ft. AFL. Based on these data and the pilot comments, the use of the PNF-AD for incursion detection was not practical. The presence of symbology on the PNF-AD could also obscure the vehicles.

The display concepts tested in this experiment – typical of current and future PF HUD and PNF-AD displays – showed poor incursion detection functionality. Only one of the runway incursion scenarios was detected through use of the cockpit displays. Therefore, requirements for display and sensor technology for runway incursion detection should be developed which span the breadth of the problem, including human perception, sensor design and detection theory, crew procedures, and crew interface issues. Current flight crews are not familiar with using head-down displays on short final to check for incursions. The displays are not necessarily optimized for this role. This role was not intentionally included in the pre-experiment crew briefing.

5.7.2 Navigation Error

The navigation errors were either a 50 foot or 75 foot lateral offset (see Figure 8 for PNF-AD with Symbology example). The offsets could be detected by either the PF or the PNF using numerous display indications.

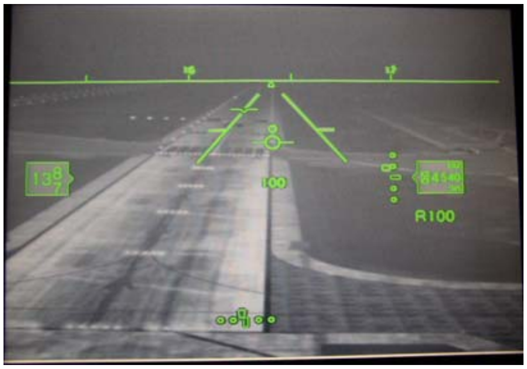


Figure 8. 75 foot localizer offset

The majority of flight crews verbally noted the presence of the 50 foot offset (15/24) and 75 foot offset (22/24) during the approach. None of the pilots executed a go-around with this anomaly. Each performed a lateral correction and landed near the runway centerline. Video analysis showed that navigation errors were predominately noted by the PF (~85%) when they noticed that the pitch/roll guidance symbol was leading them to the left or right of the runway. One person (flying as the PNF) noted the non-zero localizer deviation on the PFD presentation while tracking the path centerline.

The flight crews were not instructed on the course of action to take when confronted with a navigation error, and the pilots had relatively little training and experience with the system. Despite this, the study showed that lateral navigation errors were verbally acknowledged a significant percentage of time and, even when unrecognized (i.e., not explicitly verbalized), all flight crews landed safely and accurately on the runway. These results suggest that dissociations between raw data, sensor, and/or database presentation should be easily recognized and managed by experienced pilots. Pilot training to recognize these discrepancies could further improve operations in the event of this anomaly.

5.7.3. Illegal Landings

Although not technically a “non-normal,” each flight crew was confronted with four trials where weather conditions obscured the required visual cues to complete the landing from 100 ft HAT as defined by FAR §91.175. Of the 48 “illegal landing” rare event trials, only during six of these trials did pilots continue and land the aircraft. On each of the six illegal landings, the pilot flying had excellent visibility of the runway using the FLIR on the HUD. However, the §91.175 rule requires visual acquisition of the runway references without use of the EFVS.

No statistically significant effects of the HUD or PNF-AD concept was observed. The operational procedures necessary to follow the §91.175 regulation were found to be awkward for the PF, requiring the PF to declutter the HUD or look-around the HUD combiner. The radio altitude shown on the HUD could be used for judging HAT by the PF. The PNF had to go head-down to read the altitude on the PFD or PNF-AD. The experiment did not use a "100 ft" AFL call-out. The aural call-outs were set to Cat. I decision heights.

The few occurrences of "below minimums" landings suggest that the current regulations can be operationally viable. However, an aural call-out at 100 ft AFL may help overcome the lack of awareness to the HAT. Nonetheless, there still exists an awkwardness in the transition from EV/HUD-to-visual runway references. The PFs typically commented that the EFVS provided suitable visual references to complete the landing.

5.8 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<0.001) and post-hoc tests showed three unique subsets for the 4 HUD concepts: 1) Fusion-Tunnel (most preferred); 2) Fusion-Baseline & FLIR-Tunnel; and 3) FLIR-Baseline (least preferred).

The AD preference rating data was also highly significant (F(3, 69)=23.74, p<0.001) with post-hoc tests showing three overlapping subsets for the 4 AD concepts: 1) Fused-Symbology (most preferred); 2) FLIR-Symbology & Fused-No Symbology; and 3) Fused-No Symbology & FLIR-No Symbology (least preferred). The rankings indicate a strong preference for symbology on the AD. Post-test pilot comments noted that the PNF-AD should allow pilot-controllable symbology declutter for object detection/recognition.

6.0 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.

These data show that significant improvements in SA can be provided by the integration and/or fusion of synthetic and enhanced vision technologies for the PF and PNF. 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 significantly 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. 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. Sensor and display design must be tailored to this function and corresponding crew procedures and interfaces developed to support RI detection.

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.

7.0 REFERENCES

[1] Joint Planning and Development Office. *Next-Generation Air Transportation System Integrated Plan, Version 1*. Dec 2004.

[2] 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.

[3] Arthur, J.J., Prinzel, L.J., Kramer, L.J., Bailey, R.E., and Parrish, R.V. *CFIT Prevention Using Synthetic Vision*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2003, Editor: Jacques G. Verly, Volume 5018, Paper 16, April 2003.

[4] 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.

[5] 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.

[6] 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.

[7] Craig, G., Jennings, S., Link, K., and Kruk, R.: *Flight Test of a Helmet-Mounted, Enhanced and Synthetic Vision System for Rotorcraft Operations.* paper presented at the RTO SET Workshop on "Enhanced and Synthetic Vision Systems" held in Ottawa, CA, Sept. 2002, and published in RTO-MP-107.

[8] Parrish, R.V, Busquets, A.M., Williams, S.P., & Nold, D.E.: *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.

[9] Theunissen, E., et al: *Integration of Imaging Sensor Data into Synthetic Vision Display*. The 23rd Digital Avionics Systems Conference, 2004. DASC 04. Volume 1, 24-28 Oct. 2004. pp. 4.A.1 - 41-10.

[10] 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.

[11] 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.

[12] 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

[13] D. R. Jones, C. C. Quach, and S. D. Young: *Runway Incursion Prevention System – Demonstration and Testing at the Dallas/Fort Worth International Airport*. Proceedings of the AIAA/IEEE 20th Digital Avionics Systems Conference, 2001.

[14] Young, S.D.: *On the Development of In-flight Autonomous Integrity Monitoring of Stored Geo-spatial Data Using Forward-looking Remote Sensing Technology*. PhD Dissertation, Ohio University, March 2005.

[15] Ganoe, G., and Young, S.D.: *Utilization of GPS Surface Reflected Signals To Provide Aircraft Altitude Verification For SVS*. SPIE. In Proceedings of SPIE, Enhanced and Synthetic Vision 2005 Editor: Jacques G. Verly, Volume 5802-12.

[16] Federal Aviation Administration*: HUD Certification Criteria - Working Paper*. PS-ANM100-2001-00085 **(**Draft), Dated January 13, 2005.

[17] Bender, E.J., Reese, C.E., and van der Wal, G.S.: *Comparison of Additive Image Fusion vs. Feature-Level Image Fusion Techniques for Enhanced Night Driving*. Low-Light-Level and Real-Time Imaging Systems, Components, and Applications, C. Bruce Johnson, Divyendu Sinha, Phillip A. Laplante, Editors, Proceedings of SPIE, Vol. 4796 (2003), pp. 140-151.

[18] Endsley, M. R.: *Situation Awareness in Aviation Systems*. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), Human factors in Aviation Systems (pp. 257-276). Hillsdale, N.J.: Lawrence Erlbaum. 1999.

[19] Simard, P., Link, N.K., and Kruk, R.V.: *Feature Detection Performance with Fused Synthetic and Sensor Images*. Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, pp. 1108-1112, 1999.

[20] Andre, A.D., Wickens, C.D., and Moorman, L.: *Display Formatting Techniques for Improving Situation Awareness in the Aircraft Cockpit*. International Journal of Aviation Psychology, Vol. 1, No. 3, 1991, pp. 205-218.

[21] Greig, I., Shaw, G., and Lumsden, B.: *Augmented Vision Systems for the All Weather Autonomous Landing Guidance of Fixed Wing Aircraft: An Investigation of the Key Parameters by Piloted Flight Simulation*. From Proceedings of the 10th CEAS Conference on Free Flight: 1997. pp, 32-1 to 32-15

[22] Mann, T.L.: *Autonomous Landing Guidance Concept - The Effects of Video and Symbology Dynamics on Pilot Performance*. Paper presented at the 6th SAE Aerospace Behavioural Engineering Technology Conference, Paper 872390, 1988, pp. 131-138.

[23] Schnell, T.; Ellis, K.; and Etherington, T.: *Flight Simulator Evaluation Of An Integrated Synthetic And Enhanced Vision System For Terrain Avoidance.* The 24th Digital Avionics Systems Conference, DASC 2005. Volume 1, 30 Oct.-3 Nov. 2005, pp: 4.E.4-1 - 4.E.4-18.

[24] Bailey, R.E., Parrish, R.V., Kramer, L.J., Harrah, S.D., and Arthur, J.J. (Trey) III*: Technical Challenges In the Development of a NASA Synthetic Vision System Concept.* 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.

[25] Snow, M.P., & French, G.A.: *Human Factors In Head-Up Synthetic Vision Displays*. In proceedings of the 2001 Aerospace Congress (ACE2001CD-01-2652). Warrendale, PA: Society of Automotive

Engineers (2001)

[26] Ververs, P. M., & Wickens, C. D.: *Head-Up Displays: Effects Of Clutter, Display Intensity, And Display Location On Pilot Performance.* International Journal of Aviation Psychology, 8(4), 377-403, 1998.

[27] Snow, M.P., Reising, J.M., & Liggett, K.K.: *Flying Complex Approaches Using A Head-Up Display: Effects Of Visibility And Display Type*. Air Force Research Laboratory, Human Effectiveness Directorate Report No. AFRL-HE-WP-TR-2000-0011. 2000.

[28] Kramer, L.J., Prinzel, L.J., Arthur, J.J., & Bailey, R.E. *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.

[29] Larimer, J., Pavel, M., Ahumada, A. and Sweet, B.: *Engineering a Visual System for Seeing Through Fog.* AeroTech '92, SAE Paper No. 921130, pp. 428-432. 1992.

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

[31] Federal Aviation Administration Advisory Circular*: Standard Operating Procedures for Flight Deck Crewmembers*, AC 120-71A, Dated February 27, 2003.

[32] Lorenz, B. and Korn, B.: *Crew Coordination Issues Of EVS-Approaches.* In Enhanced and Synthetic Vision 2004, edited by Jacques G. Verly, Proceedings of SPIE Vol. 5424., pp. 89-97.

[33] Tuttle, T.T., Imrich, T., Etherington, T.J., and Theunissen, E.*: Terminal Area Operations with Enhanced and Synthetic Vision: Experience in the Boeing Technology Demonstrator.* in "Enhanced and Synthetic Vision 2003", Edts Jacques G. Verly, Proc. of SPIE Vol. 5081, pp. 136-145.

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

[35] 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.

[36] Ames, Lawrence L. & George, Edward J.: *Revision and Verification of a Seven-Point Workload Estimation Scale.* Air Force Flight Technical Center: AFFTC-TIM-93-01, 1993.

[37] American Institute of Aeronautics and Astronautics: *Guide to Human Performance Measurement*. Washington D.C.: American Institute of Aeronautics and Astronautics. 1993.

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

[39] RTCA: *Minimum Aviation System Performance Standards: Required Navigation Performance for Area Navigation.* RTCA DO-236A. Washington, D.C.: RTCA, Incorporated, 2000.

[40] Kramer, L.J., Prinzel, L.J., Bailey, R.E., & Arthur, J.J.: *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, 2003.