SYNTHETIC VISION SYSTEMS IN GA COCKPIT- EVALUATION OF BASIC MANEUVERS PERFORMED BY LOW TIME GA PILOTS DURING TRANSITION FROM VMC TO IMC

Takallu, M.A., PhD, CFII, Lockheed Martin, Hampton, Virginia
Wong, D.T. and M.D. Uenking, NASA Langley Research Center, Hampton Virginia

Abstract
An experimental investigation was conducted to study the effectiveness of modern flight displays in general aviation cockpits for mitigating Low Visibility Loss of Control and the Controlled Flight Into Terrain accidents. A total of 18 General Aviation (GA) pilots with private pilot, single engine land rating, with no additional instrument training beyond private pilot license requirements, were recruited to evaluate three different display concepts in a fixed-based flight simulator at the NASA Langley Research Center’s General Aviation Work Station. Evaluation pilots were asked to continue flight from Visual Meteorological Conditions (VMC) into Instrument Meteorological Conditions (IMC) while performing a series of 4 basic precision maneuvers.

During the experiment, relevant pilot/vehicle performance variables, pilot control inputs and physiological data were recorded. Human factors questionnaires and interviews were administered after each scenario.

Qualitative and quantitative data have been analyzed and the results are presented here. Pilot performance deviations from the established target values (errors) were computed and compared with the FAA Practical Test Standards. Results of the quantitative data indicate that evaluation pilots committed substantially fewer errors when using the Synthetic Vision Systems (SVS) displays than when they were using conventional instruments. Results of the qualitative data indicate that evaluation pilots perceived themselves to have a much higher level of situation awareness while using the SVS display concept.

Introduction
Limited visibility is the single most critical factor affecting both the safety and capacity of worldwide aviation operations. The Synthetic Vision Systems (SVS) Project of the National Aeronautics and Space Administration’s (NASA) Aviation Safety Program (AvSP) is striving to eliminate poor visibility as a causal factor in aircraft accidents as well as enhance operational capabilities of all aircraft.

The Synthetic Vision Systems-General Aviation (SVS-GA) element of NASA’s Aviation Safety Program is developing technology to eliminate low visibility induced General Aviation (GA) accidents through the application of synthetic and enhanced vision techniques. SVS displays present computer generated 3-dimensional imagery of the surrounding terrain on the primary flight display (PFD) to greatly enhance pilot’s Situation Awareness (SA) and reduce pilot’s Spatial Disorientation (SD), thus reducing or eliminating Controlled Flight into Terrain (CFIT), as well as Low-Visibility Loss of Control (LVLOC) accidents. SVS-conducted research is facilitating development of intuitive display concepts that provide the pilot with an unobstructed view of the outside terrain, regardless of weather conditions and time of day. Both accident types involve limited visibility conditions as a causal factor.

During the course of instrument training, pilots are trained to increasingly rely on the visual orientation cues provided by the cockpit instrumentation and to progressively manage their vestibular sense of orientation. Currently, pilots still require substantial training to obtain instrument-rating privileges, and even at that level of training, LVLOC accidents continue to occur at an unacceptable rate. Common errors associated with
instrument flight involve the improper scanning and interpretation of the flight instruments. These errors are known as fixation, omission, and emphasis [1]. However, the root cause of these errors is the non-intuitive presentation of the aircraft attitude and position with respect to the outside world. The objective of the SVS Project is to develop cockpit display systems with intuitive visual cues that replicate the safety and operational benefits of flight operations in clear day Visual Meteorological Conditions (VMC).

As an initial investigation, the SVS-GA team conducted a series of studies that focused on determining the associated benefits of SVS displays towards reducing LVLOC and CFIT accidents for GA pilots. While previous studies have been conducted regarding the understanding of SA and SD, leading to some novel concepts, only relatively minor progress has been made towards measuring the effectiveness of these new flight displays [2-6]. The study discussed in this paper was conducted to measure the effectiveness of these new technologies in the GA cockpit for addressing LVLOC. To obtain a quantitative measure of the new displays, new and state of the art analysis tools also needed to be developed and implemented. A brief description of some of these analysis tools will be also presented.

**Description of Experiment**

In a recent work, reference [7], the description of the experimental set-up was detailed. In this paper, the focus of the report is on the analyses of the data gathered during the experiment and only a brief description of the experiment is presented in the following paragraphs.

**Experimental Theory**

It was theorized that a Visual Flight Rules (VFR) pilot in normal flight acts as a feedback controller who makes control inputs based on visual and vestibular information to correct for deviations from the desired aircraft state. When non-instrument rated pilots process information solely from the aircraft instruments, the additional processing time required increases feedback lag time, creating an unstable situation, which leads to Low-Visibility induced Aircraft Upset (LVAU). LVAU without recovery might lead to a LVLOC accident. The presence of computer-generated terrain on the primary flight display should enable low-time GA pilots to maintain superior mental models of the outside world while operating in IMC, enhancing spatial orientation and situation awareness, and thus eliminating LVLOC and CFIT accidents.

**Experimental Setup**

The experiment was conducted in the GA Work Station (GAWS) at NASA Langley Research Center (LaRC). The hardware at GAWS was based on Precision Flight Control’s PC-based Aviation Training Device (PCATD) Model PI-142 instrument procedure trainer. Initiative Computing’s (IC) Elite Electronic IFR Training Environment Version 6.2 was used as the flight simulation software. The out the window imagery was provided by IC’s GenView software. The SVS research display software was developed in-house under a contract with Raytheon Incorporated. Three Pentium-3 class computers with high-end graphics cards hosted all the software. Aircraft position and orientation information from the Elite computer were relayed to the front visual and SVS computers via Ethernet interfaces. A 15” LCD was used to display the SVS imagery for the evaluation pilots. The research terrain database, which was created for the Asheville, North Carolina area using 3-Arcsec digital elevation model (DEM) data, was employed to generate the SVS imagery. Texturing applied to the terrain database was colored, based on the absolute altitude ranging from dark green for altitudes from 0 to approximately 800 ft mean sea level (MSL) to white for altitudes greater than 9,000 ft MSL. The GAWS control room area was isolated from the pilot station. Only the evaluation pilot was located in the GAWS cockpit simulation area during the experiment.
**Flight Displays**

Three distinct display concepts were employed for the study, and each pilot flew all 3 displays:

Display 1-the baseline display, referred to as Attitude Indicator (AI), replicated conventional instrumentation common in today’s GA aircraft. Illustrated in Figure 1, the AI display included the basic 6 gauges (airspeed, attitude, altitude, turn coordinator, directional gyro, and vertical speed indicator) along with a tachometer gauge that are typical of current GA aircraft. All gauges for the AI display concept were 3” in diameter.

Display 2- The Electronic Attitude Indicator (EAI), Figure 2, featured an enlarged attitude indicator that was more representative of a “glass display”; it also included some advanced flight symbology, such as a velocity vector. Enlarging the attitude indicator to approximately 4” by 4” provided the ability to evaluate the effect of attitude indicator size as well. In order to limit the number of displays evaluated within this study, some advanced symbology was added to the EAI display. Basic symbology included on the EAI display was a horizon line, a pitch grid, a roll scale with sideslip wedge and a digital heading. Advanced symbology included a velocity vector with sideslip flag and acceleration caret. The systems, such as Air Data, Attitude, and Heading Reference Systems (ADAHRS), required for generating the symbology are considered to be essential part of GA glass cockpit systems. The symbology was presented on top of a blue-sky brown-ground background.

Display 3- The SVS display, Figure 3, was identical to the EAI except that computer-generated terrain imagery replaced the conventional blue-sky/brown-ground background of the EAI. In order to keep the symbology identical to the EAI display concept, a fixed 50-degree Field of View (FOV) was employed for the SVS imagery. The conformal (non-reduced) FOV is about 10.3 degrees. Thus, for this SVS display concept, a reduction or minification factor of approximately 4.8 was created.

**Training Scenarios**

Each evaluation pilot was trained on use of all three displays and the experimental equipment. The training scenarios were similar to the actual experimental flight scenarios. Then the pilots were instructed to perform the standardization scenarios within the limits of Federal Aviation Administration (FAA) Practical Test Standards (PTS) for private pilot certificates. One of the evaluation pilots was not able to perform to the above standards during the training and during the VMC portion of the test scenarios. Consequently, the data gathered from this pilot was disqualified.

**Test Scenarios**

Each scenario started with an out of window visibility of more than 20 statute miles (SM), which was gradually reduced to 3 SM within a 3 minutes period. Even though 3 miles visibility is considered marginally VFR, for the purposes of this experiment, 3 miles out of window visibility was considered to be IMC. The duration of each scenario was a total of 5 minutes. Pilots were briefed to use out the window pilotage as much as possible and upon entering IMC to execute one of four specified scenarios:

Scenario 1: Straight and Level - fly straight and level while maintaining airspeed, altitude and heading.

Scenario 2: U-Turn - make a 180° turn with a 20° bank while maintaining altitude and airspeed.
Scenario 3: Descent - descend 1,000 feet while maintaining heading and airspeed.

Scenario 4: Climb - climb 1,000 feet at 80 kt while maintaining heading.

All four scenarios were initiated at 2,500 Above Ground Level (AGL) with a speed of 100 kt and a heading of 20 degrees. The evaluation pilots were asked to maintain airspeed within +/-10 kt, altitude within +/-100 ft, and heading within +/-10 degrees of the assigned values. A LVAU condition was considered to be when either pitch angle was greater than +25/-10 degrees or bank angle was greater than +/-45 degrees. Massive loss of situation awareness was defined as altitude errors greater than 1,000 feet and heading errors greater than 45 degrees. Experimental scenario presentation sequences were grouped into 12 possible combinations. The experimental combination was selected at random for each evaluation pilot and not re-used. Evaluation pilots would perform the 4 evaluation maneuvers for each display configuration before proceeding to the next display configuration. None of the scenarios were intended to excite CFIT conditions.

**Measurement Variables**

Quantitative measures evaluated for this study included, pilot/vehicle performance variables such as heading, altitude, airspeed, and bank angle along with pilot control inputs such as longitudinal, lateral, directional, and throttle control inputs. Physiological measurements included heart rate, skin temperature on the left (flying) hand, as well as muscle flexure on the left forearm. Qualitative measures included the NASA TLX and a Stress-Arousal Checklist (SACL) [8] after each run and responses to post-test questionnaires.

**Analyses Tools**

Pilot performance and control inputs for all flight conditions (62 variables) were recorded by a custom version of Elite Simulation Solutions software at 60 HZ. Out of 62 variables recorded, 10 time dependent variables of interest were down-selected for further analysis within a computer program developed in-house called the SVS Analysis Tool (SVSAT). With the use of the SVSAT, each scenario was broken up into 5 different segments of a maneuver (except scenario 1 which was divided into 2 segments only). Since every scenario was initiated with a straight and level flight in VMC and transitioned into IMC, the actual maneuver (i.e. the U-turn) was labeled as segment 3 and separated from the starting level flight portion (segment 1), the transition to the turning maneuver (segment 2), the transition to level flight (segment 4) and the final level flight portion (segment 5) of the scenario.

![Figure 4: Segmentation of a Scenario for Different Maneuvers](image)

**Physiological Data**

The physiological data collection apparatus was the MP100TM system developed by BIOPAC Systems Inc. Physiological state data such as heart rate, skin temperature, and muscle response were recorded using the above system to determine stress and workload. Data was transmitted from the BIOPAC system to a Pentium-2 class PC for analysis.
Other Data Collections

Audio and Video recording of pilot activity and post experiment briefings and exit interviews were conducted. Subject matter experts (instructor pilots) were present to observe pilot performance and record the Pilot Performance Observation (PPO) forms. Between each test scenario pilots were asked to complete NASA TLX and Stress and Arousal Checklist (SACL) forms. After each scenario the run log from the PPO forms, TLX and SACL forms and Elite recordings were cataloged and archived.

Results and Discussions

Results of the analysis are grouped into different categories and are listed in the following sections, based on the type of data obtained. The duration of each scenario was a total of 5 minutes.

Quantitative results:

![Graphs showing Time History of Pertinent Variables](image_url)

Figure 5: Time History of Pertinent Variables
A sample time history of the pertinent performance variables is shown in Figure 5 for one of the
evaluation pilots executing scenario 3 (descent maneuver). Deviations of pilot/vehicle performance variables, \(v_j\),
from the established target or reference values, \(r_j\), can be defined as the performance errors, \(e_j\):

\[ e_j = v_j - r_j \]

The reference values for the above equation were described above, in the test scenarios section.

The first column of the plots shows the time history of pilot/vehicle performance errors, \(e_j\), such as
errors in aircraft heading, roll angle, altitude above mean sea level, and indicated airspeed for all display types.
The first plot in the second column shows the value of the out of window (flight) visibility in Nautical miles
(Nm). As can be seen, the visibility was lowered from VMC to IMC in 3 minutes. The second and third plots in
the second column show pilots’ roll and pitch control inputs during this task. In this example, the values of the
pilots’ performance errors and control inputs show less amplitude and smoother changes in both frequency and
amplitude when using the SVS display as compared to when either the AI or EAI display was used.

Statistical analysis of pilot performance and control inputs:

Each one of seventeen pilots flew three iterations of each of the four scenarios, once with each display
type. Each run was partitioned into two or five flight segments, depending upon the scenario. Four separate
analyses, using data from one flight segment of each scenario, were done repeatedly. Statistical analyses were
conducted on several measures (e.g., Root Mean Square (RMS), standard deviation (StD), time ratio statistics)
computed from the pilot performance errors and control-input activities. In statistical analysis of the measures
described above, the pilot performance errors and control-input activities were computed for each segment of
each scenario and the magnitude of statistical significance of each measure was carefully verified. Across
various displays, scenarios, and segments the computed differences were statistically significant in many cases.
An example is shown in Figures 6-9. In this case, there was a weak but statistically significant result in the
standard deviation of heading error for scenario 1. It is showing that the measure when using display 3 was
smaller than with displays 1 and 2 (F(1,17)=3.3, p=0.004). In scenarios 3 and 4 there are significant differences
in the RMS, StD and time ratio measures for the heading error. The post hoc analyses show that error while
using display 1 is significantly different (larger) from displays 2 and 3. The details of the statistical analyses of
all variables for all scenarios will be reported in a later NASA report.

Figures 6-9 show the statistically significant results of the primary segment (segment 2 for scenario 1
and segment 3 for the other 3 scenarios) for all 4 scenarios and all 3 displays. In each figure, the values of
computed errors within the maneuver are plotted for the particular scenario. The first column of the plots shows
mean value (over 17 pilots) of RMS of each performance variable (aircraft altitude for scenarios 2 and heading
for other scenarios). The second column shows the mean value (over 17 pilots) of RMS of control inputs
(lateral input for scenario 2 and lateral input for other scenarios). The third column shows the counts of the
number of times the parameters (aircraft altitude for scenarios 2 and heading for other scenarios) exceeded the
PTS limits. The results indicate that the RMS of pilot performance errors and control inputs and the time with
error ratio measures were lower when using the SVS display than when using the EAI and AI displays.

Scanning and Error Correlations:

During the course of instrument training, pilots are trained to increasingly rely on the visual orientation
cues provided by the cockpit instrumentation and to progressively manage their vestibular sense of orientation.
As stated previously, common scanning and interpretation errors associated with instrument flight include
fixation, omission, and emphasis. However, the root cause of these errors is the non-intuitive presentation of the
aircraft attitude and position with respect to the outside world. In statistical analyses of the errors described
above, measures of the pilot performance errors and control-inputs were computed for each variable and the
magnitude of statistical significance of each measure was established over the number of samples and pilot
population. The statistical analyses showed the level of the error for each measure in isolation from the others.
However, an untrained pilot flying with reference to instruments typically deviates from altitude while
concentrating on maintaining heading or vice versa.
Figure 6: Scenario 1, Straight/Level

Figure 7: Scenario 2, Level U-Turn

Figure 8: Scenario 3, Straight Descent and Level-off

Figure 9: Scenario 4, Straight Climb and Level-off
Assume that the deviation from the assigned value for a pilot task (i.e. heading, bank, airspeed and altitude for each scenario and display type) is the numerical error of a system process. One can then display the interdependence of these errors as a measure of cross-correlation between errors of primary variables. Such a pattern of errors (a phase plot) is illustrated in Figure 10 for scenario 3 (descent maneuver). In this figure, the error patterns of all 3 displays used by the same pilot are summarized in a plot of the normalized value of indicated airspeed error versus normalized value of heading error. The errors, $e_j$, were normalized by the tolerance values, $t_j$, which were described above, in the test scenarios section, and computed as:

$$e_j = \frac{e_j}{t_j}$$

In the above figure, each separate curve represents an evaluation pilot’s performance when using each display during this task. In this manner, the size of enclosed area by each curve can be interpreted as the measure of the pilot’s inability to process the changes in aircraft altitude and airspeed simultaneously (in a parallel manner). By visual inspection of the curves in the above example, Figure 10, it can then be concluded that this pilot performed best (processed multiple information sources and acted upon them in a parallel manner) when using the SVS display as compared to using EAI or AI displays. Computations of these areas are rather complex and computationally costly.

![Figure 10: Scanning Error Pattern during Scenario 3, for Evaluation Pilot 11](image)

A simpler method, than computing the error areas, is to compute the Norm of a vector space described by $j$-number of performance errors. A norm is defined [8] as “real-valued function that provides a measure of the size or length of multi-component mathematical entities such as vectors and matrices.” A p-Norm defined as

$$||e|| = \left( \sum |e|_p \right)^{1/p}$$

The Euclidean Norm of a vector space is the $p=2$ and L2 Norm for a two variable system becomes:

$$L2 \text{ Norm} = \left( e_1^2 + e_2^2 \right)^{1/2}$$

In the above example, $e_1$ and $e_2$ are the normalized errors of altitude and airspeed, respectively.
Figure 11: L2 Norm behavior for Scenario 3, for Pilot 11

Figure 11 illustrates an example plot of the L2 Norms for 3 different display conditions over the duration of the scenario 3 (5 minutes) for the same evaluation pilot as in figure 10. Since the L2 Norm combines the errors of all sources for this scenario, it can uniquely reveal any combined system performance effect as illustrated in Figure 11. The lower L2 Norm values are particularly apparent three to four minutes into the flight as the pilot entered into the IMC portion of the task. Any errors due to pilot induced inputs are further amplified during this period. It can be seen that the evaluation pilot performed better when using the SVS display than when using the EAI display or when using the conventional instruments (AI).

Figure 12: Total Scanning Error over the Length of the Scenario for each Pilot

The areas under the curves in Figure 11 can be seen as the total scanning error of the pilot during the entire 5-minute flight. Therefore, a simple way to compare the effectiveness of the three display types for all pilots will be to compare their corresponding total scanning errors. Figure 12 shows a plot of the time averaged scanning errors of each subject pilot for the three display types tested. The total scanning errors, or the areas under the L2 norm curves, were computed using the trapezoidal rule numerical integration method. It turns out that 8 out of the 16 pilots performed better using the SVS display. For those pilots whose results indicate that they did not perform as well using SVS display, the magnitudes of the corresponding scanning errors, using the other two displays, were very similar to the errors when using the SVS display. It is assumed that these
differences arose because these pilots were most likely trained in using the turn rate and airspeed as primary instrument (primary/support method) instead of the attitude indicator (control and performance method) [1].

Physiological Data
The physiological data was analyzed as part of this experiment to determine if the various display concepts produced measurable physiological differences that would either support or refute hypotheses. No statistically significant findings were found in the physiological data. Since all test subjects were asked to fly a specific flight condition when they experienced planned IMC conditions, the test subjects were able to anticipate what the required course of action needed to be and thus experienced no surprises. The only physiological measure that approached statistically significant levels was the skin temperature measurements. In the event of increased workload, periphery skin temperature decreases. Likewise, in the event of decreased workload, periphery skin temperature increases. In the two most workload-intensive maneuvers, descent (scenario 3) and climb (scenario 4), the skin temperature measurements showed higher skin temperature in IMC for the SVS display versus the AI. This would indicate lower workload for the test subjects during IMC portion of the task when using the SVS display. The general linear model for repeated measures and the one-way Analysis of Variance (ANOVA) were employed to analyze the data using the SPSS® software (Figure 13). The skin temperatures for the descent maneuver between the baseline and SVS displays were near significance (F(1,17)=1.911, p=.185).

Figure 13: Average Skin Temperature for Descent Maneuver

Qualitative Results:
The recorded questionnaires and the transcripts of exit interviews with subject pilots were analyzed and are shown in the following sections.

TLX and SACL Data
The evaluation pilots filled out TLX and SACL questionnaires after each task. Statistical analysis of the NASA TLX shows excellent values of significance in the workloads between all three displays (F(2,204)=5.680, p=.004). Figure 14 indicates a decrease in workload when using the SVS displays compared to the other two displays. The Tukey Post Hoc test showed an acceptable significance (p=0.043) between the baseline display and the EAI display and an even more significance (p=0.003) between the baseline display and the SVS display. There was no significance (p=0.675) shown between the EAI display and the SVS display. The SACL calculates stress and arousal, and although there is a trend for decreased stress levels for the SVS display as shown in Figure 15, after further analysis using the SPSS® software, it was shown that this difference is not
significant (F(2,204)=0.792, p=0.454). The arousal values showed a fairly constant level of arousal across all three displays.

Figure 14: TLX Workload

Figure 15: SACL Mean Stress

Pilot acceptance of SVS

During the exit interviews with subject pilots the following questions were posed: (1) “Of the three display concepts evaluated (baseline AI, EADI, and SVS), which provided the best IMC performance.” (2) “Of the three display concepts evaluated (baseline AI, EADI, and SVS), which provided the best situation awareness.” (3) “Of the three display concepts evaluated (baseline AI, EADI, and SVS), which provided the best spatial awareness.” The transcripts of evaluation pilots’ comments are summarized in Figure 16. It is shown that the overwhelming majority of evaluation pilots preferred the SVS concept to the other 2 displays.

Figure 16: Results of Pilot Preferences during Exit Interviews
Concluding Remarks

Common pilot errors associated with instrument flight involve the improper scanning and interpretation of the cockpit instruments. These errors are due to pilot’s inability to develop and maintain an appropriate mental model of the conditions in which the aircraft is operating. The root cause of these errors is the non-intuitive presentation of the flight-critical data to the pilot. A new primary flight display known as the Synthetic Vision Systems display presents critical flight information and 3-dimensional navigation information in one display. The subject study was conducted to measure the effectiveness of these new technologies in the general aviation cockpit. To obtain quantitative measures of the new displays, state of the art analysis tools were developed and implemented. Results of both quantitative and qualitative data obtained have been presented. The reduced level of errors committed by the evaluation pilots when using the SVS display as compared to the other displays were interpreted as improved pilot situation awareness and a high level of pilots’ perception of his spatial orientation. Enhancements of pilot SA, inferred through the reduction of pilot performance errors combined with qualitative data, can be extended to imply that SVS displays will facilitate reductions of the rate of CFIT and LVLOC accidents. The above results were obtained for a generic Synthetic Vision Systems display in a fixed-base low-fidelity simulation facility and should be seen as an introductory experiment of this type.

Many issues, such as proper terrain portrayal, display symbology, cockpit integration/work load, and FAA certification, remain at the core of a successful synthetic vision systems display. Subsequent simulation and actual flight experiments will be conducted to enhance and extend the current simulation results.

Acknowledgements:

The authors would like to thank the following people for their contribution to the content of this paper and/or for many valuable discussions: Mr. Lou Glaab (SVS-GA element lead), Dr. Ray Comstock (Crew/Vehicle Integration Branch), Dr. Harold Atkins (Computational Modeling and Simulation Branch) of NASA LaRC, Dr. Richard Newman of Embry Riddle Aeronautical University), and Mr. Mel Bailey and Mr. Jake Berry of Lockheed Martin Corporation.

References