

Feature Detection Performance with Fused Synthetic and Sensor Images

Philippe Simard
McGill University
Montreal, Quebec

Norah K. Link and Ronald V. Kruk
CAE Electronics Ltd.
St-Laurent, Quebec

The operational (airborne) Enhanced/Synthetic Vision System will employ a helmet-mounted display with a background synthetic image encompassing a fused inset sensor image. In the present study, three subjects viewed an emulation of a descending flight to a crash site displayed on an SVGA monitor. Independent variables were: 3 fusion algorithms; 3 visibility conditions; 2 sensor conditions; and 9 sensor/synthetic image misregistration conditions. The task was to detect specified terrain features, objects and image anomalies as they became visible in 16 successive fused image snapshots along the flight path. The fusion of synthetic images with corresponding sensor images supported consistent subject performance with the simpler algorithms (averaging and differencing). Performance with the more complex opponent process algorithm was less consistent and more image anomalies were generated. Reductions in synthetic scene resolution did not degrade performance, but elevation source data errors interfered with scene interpretation. These results will be discussed within the context of operational requirements.

INTRODUCTION

The present study is part of a Canadian Forces Search And Rescue (CF SAR) Technology Demonstrator project to provide an Enhanced and Synthetic Vision System (ESVS) to SAR helicopter crews in poor visibility conditions. The ESVS system includes an on-board Data Base and Computer Image Generator to generate a synthetic image of local terrain, registered through the aircraft navigation system, and a near visual wavelength (Infra Red) sensor to provide a correlated out-the-window image. Both images are presented on a Helmet-Mounted-Display with the IR image fused as an inset in the center of the synthetic image field of view.

The IR sensor responds to objects in degraded visual conditions, particularly at night, but the sensor image is sub-optimal in the following ways: it typically has a small field of view when resolution matched; it is subject to degradation due to weather effects (especially with respect to resultant low spatial frequency); and it can be noisy. Synthetic images can be generated with both large field of view and high resolution and they have inherent high spatial frequency characteristics. However, they will suffer real-world correlation problems due to the resolution of the polygonal representation of terrain and cultural features and due to the resolution and accuracy of available source data. The ESVS fuses these two sources of information with the goal of providing accurate and relevant visual information to the pilot at all times.

Supporting research for ESVS has examined issues of pilot performance against parameters such as field-of-view, design eye, system (temporal) delays, and navigation data stability (CMC, 1996; Kruk et al, 1999). The current study was developed to assess image fusion algorithms for ESVS. Three fusion algorithms of varied complexity were applied to fuse

emulated IR sensor images with synthetic images in a variety of weather and synthetic data error conditions.

METHOD

Subjects

Three experienced psychophysical observers with vision corrected to Snellen 20/20 served as experimental subjects.

Apparatus/ Image Generation

Twenty Pentium II - 266 MHz PC's, running 24 hours per day for 20 days were used to generate the 240 source images (both sensor and synthetic) and 2592 fused images used in the experiment, as well as over 3000 additional source images used to generate dynamic sequences for further evaluation.

Sensor and Synthetic Image Configuration

Figure 11 shows the image configuration. This configuration was designed to register the pixels of the sensor and synthetic images automatically. It also provided an opportunity to assess boundary conditions between the inset fused image and the background synthetic image. A full description can be found in Kruk et al, 1998.

Infrared object simulation. The IR images were generated as a post-process to image generation using standard sensor controls and responses as developed at CAE Electronics Ltd. This includes random noise generation, misalignment of the sensor elements, image filtering (noise removal) and brightness/contrast and black hot/white hot controls. Objects in the database were color tuned to mimic their thermal signature for a fixed time of day and time of year (around 5 p.m.

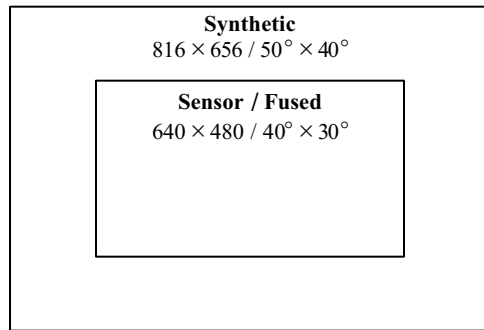


Figure 1 - Image configuration

on an early fall day, ambient temperature approximately 10°C). Finally, scene fading due to atmospheric effects was simulated.

Fusion algorithms. Three fusion methods were investigated and modified to fit the ESVS requirements:

- **Pixel Averaging Method.** The pixel averaging method assigns a weighted average of the luminance of each corresponding pixel from the sensor frame and the synthetic frame to the fused image. The weights were modified smoothly across the fused region to take into account the quality of the sensor image within 36 sub-windows.
- **TNO Method.** This method, which combines differencing and averaging, was developed at the Netherlands TNO Human Factors Institute to fuse low-intensity visible CCD camera images with infrared sensor images (Toet, 1996). We adapted and tuned the algorithm to account for sensor image quality (as for the averaging method) and synthetic scene content.
- **MIT Method.** The MIT method was adapted from a method developed at the Massachusetts Institute of Technology Lincoln Laboratory (Waxman et al, 1995, 1996a&b, 1997). Its purpose is similar to the TNO method: fusion of low-intensity visible CCD camera images with infrared sensor images. The method is based on opponent processing in the form of feed-forward center surround shunting neural networks.

Condition 1	Control: Identical database and viewpoint to sensor image (40 m terrain elevation posts).
Condition 2	Synthetic viewpoint inaccuracies introduced on flight path ($\pm 15\text{m}$).
Condition 3	Missing objects, object position offsets.
Condition 4	Global database offset (extreme - 1° , 5 m).
Condition 5	Decreased terrain resolution (mid - 160 m).
Condition 6	Local terrain elevation errors (extreme - 45m).
Condition 7	Global database offset (mid - 0.5° , 3 m).
Condition 8	Decreased terrain resolution (extreme - 320 m).
Condition 9	Local terrain elevation errors (mid - 15 m).

Table 1 - Synthetic scene registration conditions

An autogain function (inhibited in the sensor simulation) was applied following the fusion process for all algorithms.

Procedure

Flight path. A flight path was modeled to simulate a typical SAR approach up a box canyon into a crash site in hilly, forested terrain. It consisted of an approach descending from 500 ft to 30 ft over rising terrain (see Figure 22). The crash site was located on a hill-side just below a saddle ridge.

Matrix of test conditions. Six sensor conditions were developed with black hot and white hot images each at three visibility ranges (3 nm, 1.5 nm and 0.5 nm). Nine different synthetic scene registration conditions were defined to study separately the different effects that misregistration between the two image sources would have on fusion algorithms. The conditions are listed in Table 11. Together with the three fusion methods, this led to 162 different sequences of fused images (6 sensor conditions \times 9 registration conditions \times 3 fusion methods).

Task. The subjects were instructed to perform a target detection task which consisted of assessing the visibility of given features and in verifying if there were conflicts between

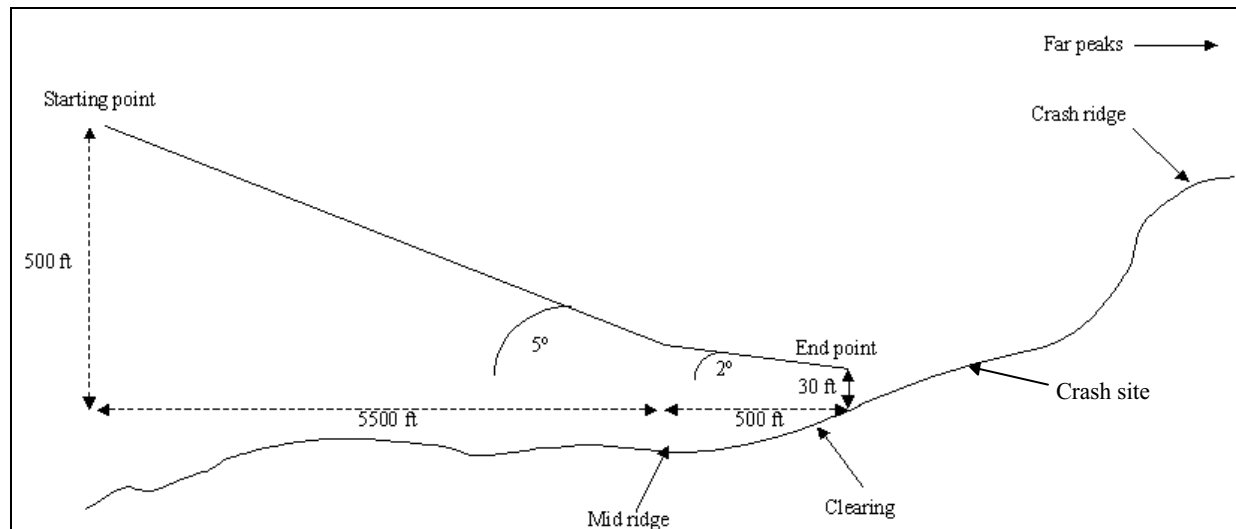


Figure 2 - Flight path and terrain profiles

objects or terrain features due to registration problems. The features (located in Figure 22) were: the far peaks (required for general route planning); a mid ridge rising to the left of the flight path (terrain obstacle on approach); the clearing and clearing obstacles; the ridge behind the crash site (also a terrain obstacle and required for route planning); and the crash site itself (visible only in the sensor image). Subjects were required to view the 16 images along the flight path and to note at which image the features first became visible and which images contained terrain or object conflicts.

Statistics. Standard t-tests to compare populations were applied to the results to assess pair-wise differences in feature-detection performance between the sensor baseline and fused sequences, as well as between the three fusion algorithms, for individual features in each visibility condition.

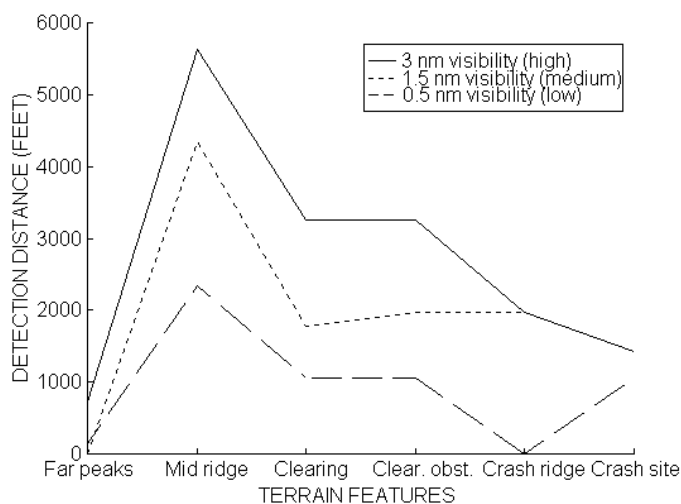


Figure 3 - Sensor baseline results (white hot mode)

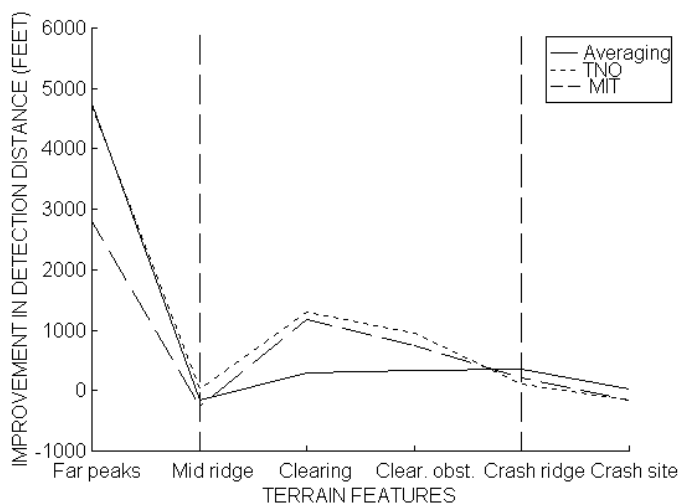


Figure 4 - Fused differentials, 3.0 nm visibility (high)

RESULTS

Baseline

The sensor sequences were evaluated first to create a baseline. The results are displayed graphically in Figure 33. A distance is associated with each feature and corresponds to the distance from the observer to the endpoint of the flight path when the feature was first detected. The average of the detection distances recorded by the three subjects in the white hot polarity is shown for each visibility condition. Note that although a curve connects the observations, it does not imply continuous results. Rather, it facilitates the comparison of the visibility conditions (and also of algorithms by clearly identifying crossover points in performance).

The far peaks were generally not visible in the sensor images, due to the fact the ceiling was low and it therefore obscured long range features. In addition, the curves are progressively lower from the high visibility condition to the low visibility condition, consistent with the fact that we see

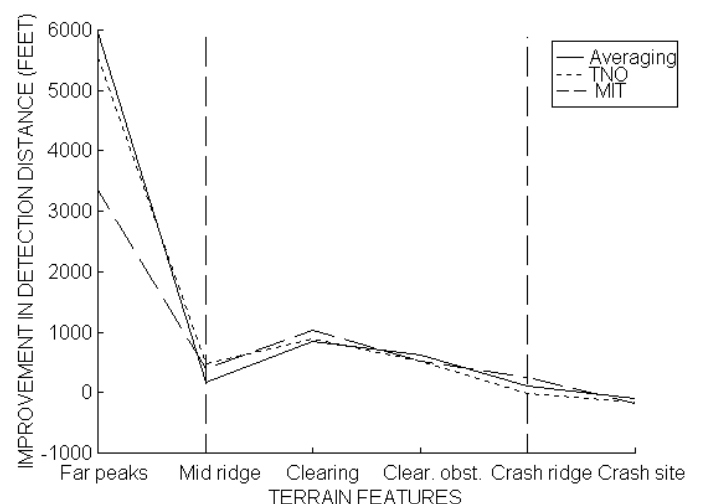


Figure 5 - Fused differentials, 1.5 nm visibility (medium)

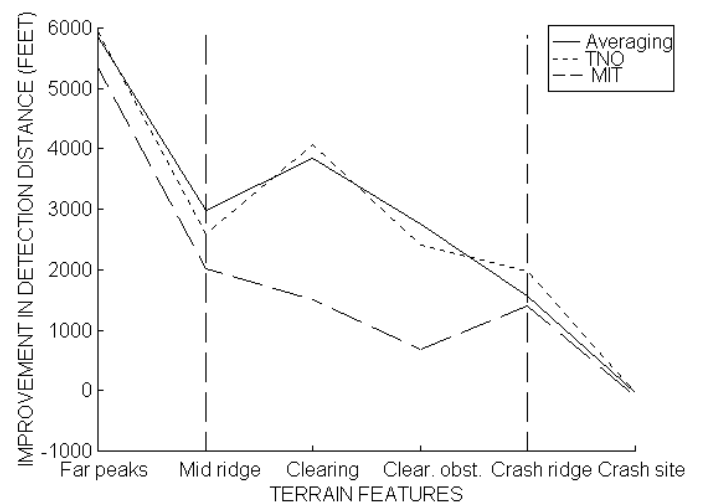


Figure 6 - Fused differentials, 0.5 nm visibility (low)

objects from greater distances in better visibility conditions. The black hot polarity results were nearly identical to white hot (i.e. sensor polarity had little or no impact on the detection distances).

Algorithm Performance by Visibility Condition

Results were pooled for each visibility condition to obtain a larger statistical sample. White hot and black hot sensor observations were combined, as were the database registration conditions

Graphs. The results for the high, medium and low visibility conditions are shown in Figure 44, Figure 55 and Figure 66. In these graphs, the distance associated with each feature corresponds to the *IMPROVEMENT* in detection distance of the fused images over the sensor alone (negative observations represent reduced performance). The graphs are segmented into three regions separated by vertical dashed lines. The left-hand region contains the far peaks. Because these were obscured by cloud in all sensor baseline trails, this region represents the extent to which features in the synthetic image were obscured by pure sensor noise. Conversely, the right-hand region contains the crash site, which was not present in the synthetic databases. This region shows to what extent each fusion algorithm allowed the synthetic image to obscure sensor data. The center region includes features present and detectable in both sensor and synthetic images.

Summary. All three algorithms provided a significant improvement over the sensor baseline in the detection distance of the far peaks (>99% confidence) and no significant difference in the detection distance of the crash site. The region of interest is therefore the central region in the result graphs, as follows.

The averaging algorithm provided no significant improvement over the sensor baseline in the high visibility condition. However, its performance improved as visibility decreased, and the improved detection distances were significant (>99% confidence) at low visibility.

The TNO algorithm produced significantly improved detection distances (>95% confidence) for the clearing and clearing obstacles at all visibilities, for the mid ridge at medium and low visibility, and for the crash ridge at low visibility. The MIT algorithm had similar results, with the exception of the mid ridge which was only improved in the low visibility condition.

A pair-wise comparison of the algorithms indicated significant differences between MIT and averaging and between TNO and averaging at high visibility. There were no differences at medium visibility, and averaging and TNO both performed significantly better than MIT at low visibility.

Database Condition Effects

The number of object (OC) and terrain (TC) conflicts were pooled for sensor conditions and were compiled by registration condition and by fusion algorithm. The results are tabulated in Table 22. Registration condition 6 (extreme local terrain elevation errors) stands out with a very high number of

Regis. Cond.	Averaging		TNO		MIT		Total
	TC	OC	TC	OC	TC	OC	
1	2	1	1	2	4	4	14
2	2	15	1	30	1	60	109
3	4	27	5	38	2	72	148
4	1	5	2	8	10	39	65
5	22	30	17	52	23	45	189
6	84	95	82	104	111	108	584
7	1	6	1	10	1	46	65
8	9	27	20	43	11	60	170
9	1	1	2	15	9	24	52
Total	126	207	131	302	172	458	

Table 2 - Number of object and terrain conflicts

terrain and object conflicts across algorithms. Conditions 5 and 8 (moderately and extremely coarse horizontal terrain resolution) generated increased but equivalent reported error performance.

Among algorithms, the number of reported object and terrain conflicts are highest for the MIT method. This could be caused by the general property of the MIT algorithm to keep more of the synthetic image even when the sensor image quality is good. The TNO algorithm produced results between the averaging method and the MIT algorithm.

DISCUSSION

Algorithm Performance

The results indicate that fusion improves the useable content of independent synthetic and sensor images. All three algorithms examined in this study provided observers the capability to detect important features from greater distances than with the sensor alone.

Pair-wise comparison of the algorithms showed the TNO algorithm was superior over a broad range of visibility conditions. Both the TNO and MIT algorithms use a high-contrast third feature – local image differences – in computing the fused image. This likely accounts for their superior performance over simple averaging in the high visibility condition. The MIT algorithm, however, did not perform as well in the low visibility condition. This method uses a fixed normalized filter to process the images, and the particular receptive field size chosen was somewhat sensitive to the sensor noise produced by sensor mismatch and by atmospheric conditions. While the normalization is an advantage for fusing two sensor images of varying image quality but correlated content, in our case the (potentially uncorrelated) synthetic image always had very sharp, high-contrast edges present compared to the sensor. This method was therefore more difficult to tune to respond to varying sensor conditions, and it resulted in significantly more conflict reports.

Registration Condition Effects

Database offset. Experimental results as well as individual observations indicated that typical database offsets due to navigation system inaccuracies or typical misalignment of coordinate systems would not have a severe impact on the system. Such errors appeared to be both detectable and tolerable to the observers.

Horizontal terrain resolution. The different terrain resolutions were a source of particular interest as low-cost image generation technologies will not yet support the very high terrain resolutions it was originally thought might be necessary at the required 60 Hz update rate. The conflict generation performance of the moderately coarse and extremely coarse horizontal resolution conditions was equivalent and subjects' comments indicated that the lowest resolution terrain had sufficient detail to permit accurate identification of key terrain features to support route planning.

Local terrain elevation errors. Results show that medium and large errors in local terrain elevation (present in source data) may result in performance problems. Initial subject reports were that the mid ridge separated into two ridges (one behind the other), and because the databases were otherwise identical it was difficult to determine which ridge was the "real" one. However, terrain objects that were vertically displaced made it more obvious as the ridge was approached, and in general, the terrain errors were then detected by the subjects as being an offset in elevation. This effect may also be less severe when the synthetic image has a lower resolution content (e.g. polygonal forest canopies) and when obvious synthetic textures are applied.

Summary

The results of the present study indicate that fusion of accurate synthetic image content with sensor - sourced images could significantly enhance pilot performance in terrain and obstacle avoidance in poor visibility operational conditions. Among the array of fusion algorithms currently available, the simpler ones seem to perform the best, albeit with considerable tuning and optimization for the conditions and task. There are distinct tradeoffs between performance enhancements in some areas, e.g. superior performance of TNO and MIT in good visibility (see Figure 44), and inferior performance of those algorithms with respect to generation of anomalies. In the current study, the TNO algorithm provided the best combination of flexibility imparted by more complex processing, as well as robust performance across a variety of conditions.

Present and Future Work

At time of writing, the pixel averaging and TNO algorithms are being implemented in flight hardware for the ESVS 2000 technology demonstrator and will be test flown in the NRC Canada Bell 205A flying testbed. Navigation system and database errors will be evaluated in pilot-in-the-loop studies

using the University of Toronto Institute for Aerospace Studies Full Flight Simulator later this (1999) year. A number of candidate active sensor systems are under consideration for database error correction and registration.

ACKNOWLEDGEMENTS

The ESVS program is supported by the Canadian Department of Defense - Chief, Research and Development, and the Search and Rescue Directorate. This study was conducted under contract # 03SD.W8477-7-AC27.

REFERENCES

- Canadian Marconi Company (CMC) and CAE Electronics Ltd., "Enhanced/ Synthetic Vision System Scoping Study", CMC Doc 1000-1102, 1996.
- Kruk, R.V., Link, N., and Simard, P., "Synthetic Vision Implementation Project Final Report", CAE Electronics Ltd. CD 342734-01-8-300, 1998.
- Kruk, R.V., Link, N.K., MacKay, W.J., Jennings, S. "Enhanced and Synthetic Vision System for Helicopter Search and Rescue Mission Support." *Proc. American Helicopter Society 55th Annual Forum*, Montreal, Quebec (Canada), May 1999.
- Toet, A. & Walraven, J., "New false colour mapping for image fusion", *Optical Engineering*, 35(3), 650-658, 1996.
- Waxman, A.M. et al, "Color night vision: fusion of intensified visible and thermal IR imagery", *Proc. SPIE Conference on Synthetic Vision for Vehicle Guidance and Control*, vol. SPIE-2463 (pp. 58-68), 1995.
- Waxman, A.M. et al, "Electronic imaging aids for night driving: low-light CCD, thermal IR, and color fused visible/IR", *Proc. SPIE Conference on Transportation Sensors and Controls*, SPIE-2902, 1996a.
- Waxman, A.M. et al, "Progress on color night vision: visible/IR fusion, perception & search, and low-light CCD imaging", *Proc. SPIE Conference on Enhanced and Synthetic Vision*, vol. SPIE-2736 (pp. 96-107), 1996b.
- Waxman, A.M. et al, "Color night vision: opponent processing in the fusion of visible and IR imagery", *Neural Networks*, 10(1), 1-6, 1997.