Real-time Enhancement, Registration, and Fusion for a Multi-Sensor Enhanced Vision System

Glenn D. Hines[a], Zia-ur Rahman[b], Daniel J. Jobson[a],Glenn A. Woodell[a]

[a]NASA Langley Research Center, Hampton, VA 23681;
[b]College of William & Mary, Department of Applied Science, Williamsburg, VA 23187

ABSTRACT

Over the last few years NASA Langley Research Center (LaRC) has been developing an Enhanced Vision System (EVS) to aid pilots while flying in poor visibility conditions. The EVS captures imagery using two infrared video cameras. The cameras are placed in an enclosure that is mounted and flown forward-looking underneath the NASA LaRC ARIES 757 aircraft. The data streams from the cameras are processed in real-time and displayed on monitors on-board the aircraft. With proper processing the camera system can provide better-than-human-observed imagery particularly during poor visibility conditions. However, to obtain this goal requires several different stages of processing including enhancement, registration, and fusion, and specialized processing hardware for real-time performance. We are using a real-time implementation of the Retinex algorithm for image enhancement, affine transformations for registration, and weighted sums to perform fusion. All of the algorithms are executed on a single TI DM642 digital signal processor (DSP) clocked at 720 MHz. The image processing components were added to the EVS system, tested, and demonstrated during flight tests in August and September of 2005. In this paper we briefly discuss the EVS image processing hardware and algorithms. We then discuss implementation issues and show examples of the results obtained during flight tests.

Keywords: enhanced vision system, image enhancement, retinex, digital signal processing, sensor fusion

1. INTRODUCTION

NASA Langley Research Center (LaRC) has been involved in numerous aviation safety improvement programs. Many of these programs are tested on the NASA LaRC ARIES 757 (NASA 757) aircraft. During August and September of 2005, several new aviation safety technologies were demonstrated on the NASA 757 as part of the Follow-On Radar, Enhanced and Synthetic Vision Systems Integration Technology Evaluation (FORESITE) program. Day and night flights were conducted and most flight paths were between LaRC and the NASA Wallops Flight Facility. One of the new technologies demonstrated was an Enhanced Vision System (EVS) that will be used to aid pilots while flying in poor visibility conditions.

The EVS consists of three video cameras that are placed in an enclosure mounted forward-looking underneath the NASA 757 aircraft, and video distribution and processing components that are located in the cabin of the plane. Proper processing of the video signals provides better-than-human-observed imagery particularly during poor visibility conditions. However, to achieve this level of image improvement requires several stages of processing and specialized hardware. In April 2005 we discussed the conceptual design of the image processing system.[1] We then developed and implemented the processing components on a digital signal processor (DSP), and in August and September of 2005 we tested the system during NASA 757 flight tests. In this paper we briefly review the EVS, image processing algorithms, and processing hardware. We then discuss several minor issues and constraints that surfaced during implementation, and show examples of the results obtained during the flight tests.

2. EVS IMAGE PROCESSING

The EVS captures imagery using three video cameras and, in real-time, processes this information to improve the quality of the data. Two of the three EVS video cameras are currently targeted for processing: a long-wave infrared (LWIR) camera that senses radiation in the 7.5–14 μm band and has a field of view (FOV) of $34° \times 25°$, and a short-wave infrared (SWIR) camera that senses radiation in the 0.9–1.68 μm band and has a FOV of
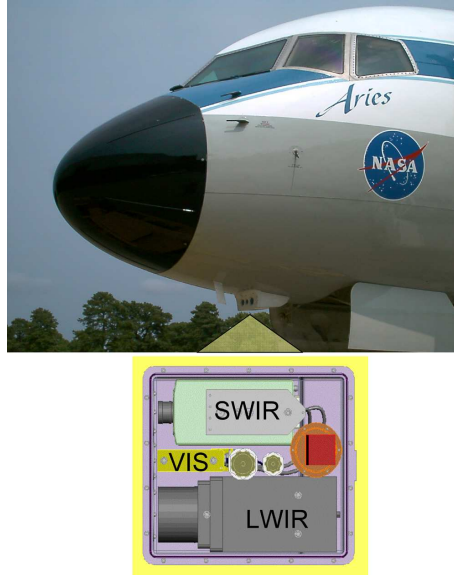
Figure 1. EVS cameras mounted under the NASA 757. LWIR and SWIR cameras are used in the current application.

Figure 2. DM642 EVM, signal splitter boards, and power supply in flight box.

Figure 3. Flight box in flight pallet on NASA 757.

$39° \times 29°.^2$ The third video camera uses a visible-band sensor. It is currently only used as a reference for what would have been seen by a human viewing the same scene as the other cameras. The cameras are mounted to the same baseplate in an enclosure that is flown forward-looking underneath the NASA 757. Figure 1 shows the enclosure installed on the NASA 757.

The EVS image processing architecture is shown in the top of Figure 4. The analog NTSC RS-170 outputs of the cameras are routed through a video distribution unit to the DSP board. The DSP board is placed in a pallet on-board the NASA 757 approximately 120 feet away from the cameras. Digital camera outputs are also distributed to the pallet using fiber optic cables, but we were unable to develop a digital input interface to our current DSP board in time for flight tests. Figure 2 shows the DSP board in our flight box and Figure 3 shows the box placed in a pallet on-board the NASA 757.

The image processing functions performed on the video data streams are shown in the bottom of Figure 4. The data streams (channels) from the cameras must be resized and enhanced, registered, and fused into a single image stream in real-time — 15-30 frames per second (fps). Both camera video streams are enhanced using the Retinex to improve the information content of the imagery. The Retinex is a general-purpose image enhancement

algorithm that simultaneously provides dynamic range compression, color constancy, and color and lightness rendition.[5, 6] It is an ideal enhancement solution in the context of the EVS because of its superb performance in improving low-contrast images. For this application the single-scale, monochrome version of the Retinex is used since it provides good enhancement of single-band infrared imagery while minimizing computational requirements.[7] The single-scale Retinex is given by

$$R_i(x_1, x_2) \quad = \quad \log(I_i(x_1, x_2)) - \log(I_i(x_1, x_2) * F(x_1, x_2)), \qquad i = 1, \ldots, S$$

where I_i and R_i are the ith spectral band of the input and output image, respectively. For a monochrome image $S = 1$. The log is the natural logarithm function and "$*$" represents convolution. F is a Gaussian surround function defined by

$$F(x_1, x_2) \quad = \quad \kappa \exp[-(x_1^2 + x_2^2)/\sigma^2]$$

where σ controls the spatial extent of the surround, and $\kappa = 1/\left(\sum_{x_1} \sum_{x_2} F(x_1, x_2)\right)$ is a normalization factor. Gain, α, and offset, β, values are applied to convert the Retinex output into the user display domain, so the final form of the single-scale Retinex is

$$R_i(x_1, x_2) \quad = \quad \alpha\big(\log(I_i(x_1, x_2)) - \log(I_i(x_1, x_2) * F(x_1, x_2))\big) - \beta, \qquad i = 1, \ldots, S$$

Values for α, β, and σ are application dependent and determined empirically. Since the cameras have different imaging sensors, a different set of Retinex parameters is applied to each camera output. Our real-time version of the Retinex current operates on images of size 256×256, so the input data streams must be cropped, subsampled or padded to these dimensions.

Registration is required to remove the FOV differences in the cameras and to correct bore-sighting inaccuracies. The SWIR data is used as the baseline since it has the smallest FOV. The LWIR data is registered to the SWIR data by applying an affine transform to the LWIR imagery.[3] A general representation of an affine transform is $[y_1, y_2, 1] = [x_1, x_2, 1]T$ where

$$T \quad = \quad \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix},$$

x_1 and x_2 reference the input coordinate system, y_1 and y_2 reference the output coordinate system, and a_{ij} are transform coefficients.[4] The mapping functions are given as

$$y_1 \quad = \quad a_{11}x_1 + a_{21}x_2 + a_{31} \text{ and}$$
$$y_2 \quad = \quad a_{12}x_1 + a_{22}x_2 + a_{32}.$$

Prior to flight, a set of control points are selected based on corresponding features from sample images acquired at the same time from the cameras. The control points are analyzed using multiple linear regression to approximate the transform coefficients a_{ij}, which is then applied to the LWIR image. The transformed LWIR image is then resampled using bilinear interpolation to align the LWIR image to the same grid as the SWIR image. The same transform coefficients are used on all LWIR video frames during flight since both the FOV and the camera alignment should not change.

The enhanced and registered images are fused by effectively performing a weighted sum of the two processed outputs since a different Retinex is applied to each channel. Pixels are summed on an inter-frame basis. Other methods, such as interleaving frames or fields causes severe flicker. Lastly, the fused data stream is output as a standard composite NTSC signal into a display. The fused video stream contains more information than either individual camera output and also provides the additional benefit of producing a single output to observe. Our sequence of tasks is as follows:

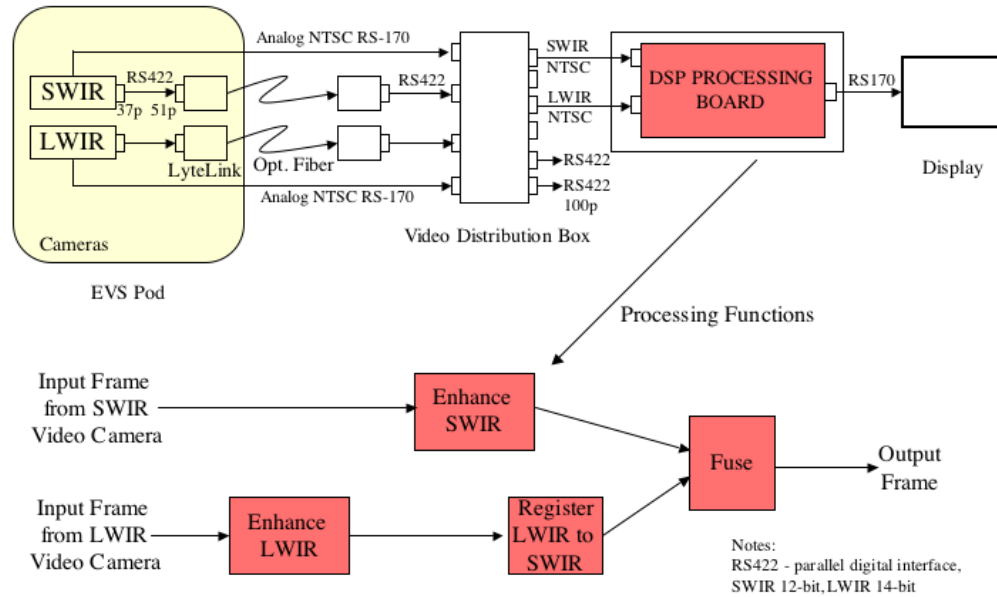- resize the LWIR input image to 256×256 pixels,

Figure 4. Image processing architecture and tasks of the EVS. The NTSC analog signals are used in our current implementation. The SWIR data is used as the baseline for registration since it has the smallest field of view.

- Retinex the LWIR image,

- resize the SWIR input image to 256×256 pixels,

- Retinex the SWIR image,

- register the enhanced LWIR image to the enhanced SWIR image,

- interpolate the LWIR image to the SWIR grid,

- fuse and output the final processed image.

Higher quality imagery is achieved by enhancing the LWIR image before performing registration, instead of registering first, since registration may eliminate part of the original image when it is transformed.

We used the TI 32-bit fixed point DM642 operating at 720 MHz for implementation.[8] The DM642 has a two-level memory architecture: level 1 has 16-KByte program and data caches and level 2 is a 256-KByte memory that can be configured as SRAM or cache. We allocated 32-KBytes of level 2 memory as cache, and 224-KBytes as SRAM in which we stored time-critical processing parameters. The DM642 is placed on an evaluation module (EVM) with 32-MBytes of SDRAM and 4-MBytes of flash memory. The SDRAM interfaces to the DM642 through a 64-bit wide external memory interface (EMIF) bus. The EMIF bus was originally clocked at 133 MHz. To improve performance we overclocked the memory to 200 MHz. The EVM also has two NTSC video inputs and one NTSC video output which allows us to receive and process the EVS camera outputs and send the processed output to a display.

Retinex Parameters		α	β	σ
Daytime	SWIR	$\alpha = 140$	$\beta = 200$	$\sigma = 40$
	LWIR	$\alpha = 200$	$\beta = 110$	$\sigma = 10$
Night-time	SWIR	$\alpha = 220$	$\beta = 100$	$\sigma = 15$
	LWIR	$\alpha = 180$	$\beta = 200$	$\sigma = 250$

Table 1: Typical EVS Retinex parameters for daytime and night-time flights.

3. FLIGHT TEST ISSUES

The EVS was tested during FORESITE flight demonstrations in August and September of 2005. All flights were performed in good weather conditions. Although the good weather conditions were not ideal for testing the EVS, it still enabled a thorough evaluation of the baseline EVS components. Baseline EVS parameters, such as power, size, and minimum frame update rates, were discussed previously[1] but several new issues, constraints and requirements surfaced during implementation.

First, both of the cameras are flown upside-down underneath the NASA 757 so the images must be rotated 180° for normal viewing. This is usually performed using embedded routines in the cameras but unfortunately the camera integrators were unable to rotate and place the corresponding gamma look-up tables in ROM for the LWIR camera. We decided to perform the rotation of the LWIR image within our image processing routines on the DSP. We modified our Retinex routine to read in the LWIR image data starting at the end of the image data and proceeding to the first pixel. This causes a 180° rotation of the image.

A new constraint was that allocated space and resources on-board the NASA 757 did not allow including components, such as emulators, that would enable real-time data exchange to perform parameter updates. Parameters that must be updated include separate Retinex offset, gain, and scale values for each channel, and registration coefficients. We used the Ethernet port to update parameters in-between flights. This eliminated the need for a JTAG emulator.

A new requirement was that the algorithms, and their associated parameters, must automatically execute at system power-up. We stored the algorithm and parameters in non-volatile flash memory. We then wrote a second-level bootloader to transfer the code from flash to RAM at boot time. The inclusion of Ethernet message processing code expanded the size of the executable beyond one flash page boundary so we also had to develop a new multi-page bootloader algorithm to implement this feature.

Our real-time Retinex algorithm currently processes a 256×256 pixel portion of each input image, but a larger 512×512 sized image is easier to view. The CCD arrays of both imagers are 320×240 pixels, but the NTSC composite inputs received by the DSP board have been upsampled to 640×480 through pixel replication (horizontally) and line duplication (vertically). We modified our core Retinex routine to exploit this fact and generated a 512×512 image by 2:1 subsampling the horizontal and vertical components of the input images and then expanding our processed output into the larger format. This process retains the majority of the original resolution of the cameras.

4. FLIGHT TEST RESULTS

The total sequence of all of the image processing tasks executed at approximately 34 fps. Enhancement and registration parameters were determined empirically and adjusted for different flight conditions such as day or night flights. Typical Retinex parameters for daytime and night-time flights are shown in Table 1. The smaller Retinex LWIR σ value for the daytime emphasizes the detail in the LWIR image. Final transform equations to register LWIR to SWIR images were determined as

$$\begin{aligned} y_1 &= (1.818501)x_1 + (0.855872)x_2 + (0.067616) \text{ and} \\ y_2 &= (41.971947)x_1 + (-0.003245)x_2 + (0.843207). \end{aligned}$$

Figure 5: SWIR image before processing.

Figure 6: Enhanced SWIR image.

Figure 7: Rotated LWIR image before processing.

Figure 8: Enhanced LWIR image.

The floating-point coefficients are scaled in the actual flight code for use in calculations by the fixed-point DSP.

The first data we show was obtained as the NASA 757 was sitting on the runway during pre-flight checkout. Input images from the SWIR and LWIR cameras are shown in Figures 5 and 7, respectively. The LWIR input is actually received from the LWIR camera upside-down but it is shown right-side up for viewing purposes. Figures 6 and 8 show the SWIR image after enhancement and the LWIR image after enhancement and registration. It is easy to see the improved contrast and brightness in the images. Registration can be seen by noting the vertical shift downward at the top of the LWIR image. The band at the bottom of the SWIR enhanced image is a result of padding to extend the vertical dimension of the image from 240 to 256 pixels.

Figure 9 is given for comparison purposes and shows the fusion of the SWIR and the rotated LWIR images prior to registration and enhancement. The final fused output is shown in Figure 10. This image has significantly better contrast, brightness, and sharpness than either of the original inputs, and the fused image prior to processing. The edges of the runway are visible in the far field of the LWIR image while the stand sitting on the runway is visible in the near field of the SWIR image. The fusion of these two images contains both of these features. The tree line which is not visible in the LWIR image can clearly be seen in the fused image. The runway marker on the right side of the image that is just below the taller trees on the right is barely visible in the SWIR. The same object is easily identified in fused image.

Figure 9. Fused images before enhancement and registration

Figure 10: Enhanced, registered and fused image.

Figure 11. SWIR image from night-time flight. Only lights are visible.

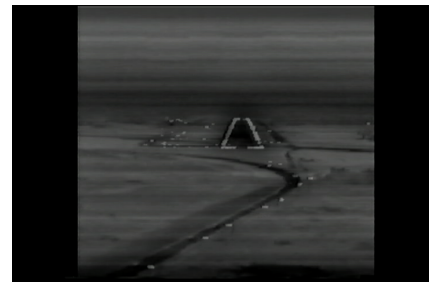**Figure 12.** LWIR image from night-time flight. Roads are visible.

Figure 13. Enhanced, registered and fused output.

Next we show data obtained from a night-time flight on September 9th, 2005. Figure 11 shows an image from the SWIR camera prior to processing. The SWIR sensor is only sensitive to lights around the runway. An output image from the LWIR camera is shown in Figure 12. Roadways and background scenery are visible. The processed output is shown in Figure 13. The enhancement and fusion of the two images clearly yields more information about the ground scenery than either individual input. This image is also slightly darker than the LWIR image because the enhancement parameters were set for this flight during daylight hours, and we were unable to make adjustments during flight. There is also a slight horizontal offset between the road of the runway and the runway lights in the fused image. We were still adjusting registration parameters when this data was captured.

Our final data was obtained from a daytime flight on September 27th, 2005. Figure 14 shows an image from the SWIR camera. Note the blooming in the foreground of the image. Figure 15 shows the associated image from the LWIR camera. Although many of the details in this image are also visible in the SWIR image, there is significantly better contrast in foreground features. The enhanced and fused image in Figure 14, displays the significant improvement achievable through proper processing even in clear weather conditions. We feel the improvements would be even more dramatic in poor visibility conditions.

5. CONCLUSIONS

We have successfully tested the image processing components of the EVS during FORESIGHT demonstrations held in August and September of 2005. The video streams from the LWIR and SWIR cameras of the EVS were enhanced using the single-scale monochrome Retinex, the LWIR data was rotated 180 degrees, the two channels were registered and fused, and the final output was recorded and displayed on monitors during the flight tests. Real-time performance of the system was obtained with all tasks executing at 33.89 fps on a single DM642 DSP.
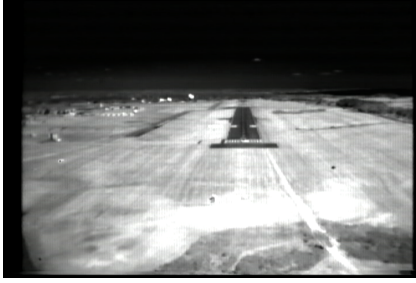
Figure 14. SWIR image from daytime flight. Note blooming in foreground.

Figure 15. LWIR image from daytime flight.

Figure 16. Enhanced, registered and fused output.

For future missions we can improve performance by processing the RS-422 digital signals from both cameras. This will require developing digital data acquisition components for the DSP board and modifying the algorithm to process 14-bit parallel digital pixel data instead of the current 8-bit data. Processing 14-bit data will primarily affect the scaling operations used in the fixed-point DM642 and increase the memory requirements. We will also process the visible-band camera data in future flight tests. This will allow us to quantify if, and under what conditions the visible-band camera can provide more information than the infrared cameras during flight.

6. ACKNOWLEDGMENTS

REFERENCES

1. G. D. Hines, Z. Rahman, D. J. Jobson, G. A. Woodell, and S. D. Harrah, "Real-time enhanced vision system," in *Enhanced and Synthetic Vision, Proceedings of SPIE 5802*, J. G. Verly, ed., March 2005.
2. C. L. M. Tiana, J. R. Kerr, and S. D. Harrah, "Multispectral uncooled infrared enhanced vision system for flight test," in *Proceedings of SPIE*, **4363**, April 2000.
3. G. D. Hines, Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-sensor image registration for an enhanced vision system," in *Visual Information Processing XII, Proceedings of SPIE 5108*, Z. Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds., April 2003.
4. G. Wolberg, *Digital Image Warping*, IEEE Computer Society Press, 1990.
5. D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. on Image Processing* **6**, pp. 451–462, March 1997.
6. D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multi-scale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing: Special Issue on Color Processing* **6**, pp. 965–976, July 1997.
7. G. D. Hines, Z. Rahman, D. J. Jobson, and G. A. Woodell, "Single-scale retinex using digital signal processors," in *Global Signal Processing Conference*, September 2004.
8. Texas Instruments, "TMS320C64x technical overview," Tech. Rep. SPRU395B, Texas Instruments, Dallas, Texas, January 2001.