Real-Time Millimeter Wave Radar Vision System

Barnabas Takács†
WaveBand Corp., Irvine, California, USA

We describe a novel real-time image processing and sensor fusion system for aerial vehicles in need of autonomous landing, guidance and obstacle avoidance. The Intelligent Synthetic Vision System (ISVS) described herein can process and display information merged from multiple image sources including a high resolution millimeter wave radar, a stored terrain with 3D airport database, FLIR, as well as visual reference images. The resulting solution provides safe all-time/all-visibility navigation in terrain-challenging areas and, in particular, real-time object detection during the final phases of landing in all weather conditions.

Journal of Imaging Science and Technology 49: 230–236 (2005)

Introduction

Intelligent Synthetic Vision Systems (ISVS) combine advanced sensor capabilities with real-time rendering of airports and 3D models of the environment to help pilots navigate and land in low visibility conditions or at night. SVS works by rendering 3D synthetic scenery based on Global Positioning Satellite (GPS) data and on-board sensors of the aircraft. Large data sets of Digital Elevation Models (DEM) as well as 3D models of airports can be used to simulate daylight approach and ideal visibility conditions.

To address these difficulties we developed a real-time imaging solution serving as an Intelligent Synthetic Vision System (ISVS). The main purpose of this ISVS is the presentation of realistic daylight scene to the pilot in all environmental conditions. Using the ISVS they are not only able to develop situational awareness with less mental integration of disparate data, but fly and land safely in all weather and lighting conditions.

However, the 3D databases do not provide immediate feedback on obstacles and dangerous events on terrain and landing sites and GPS sensors can only register the plane within unacceptable margins of error. Therefore the use of real-time imaging techniques that use advanced image enhancement, pattern recognition, and sensor fusion techniques are required to provide a fully integrated and reliable solution to this problem.

Original manuscript received July 16, 2004

†Corresponding Author: B. Takacs, Barnabas.Takacs@waveband.com

Supplemental Material—Figures 2 through 5 and 7 through 12 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.

High Performance Millimeter Wave Imaging System

Recent advances in radar antenna technology and personal computing platforms have created a novel opportunity to design all weather imaging radars that run on low cost hardware and can be readily installed in aircraft to provide live images during landing. The 95 GHz millimeter wave (MMW) imaging radar developed by WaveBand consists of a low power transceiver, a radar signal processor and an image processor. Our technical approach is based on the following premises:

- The pilot is most comfortable and less prone to errors if information is presented as a simulated daytime image of the outside world
- Such 3D rendition can be obtained from a synthetic urban or terrain database using known GPS positioning within certain error
- The on-board MMW radar sensor can generate real-time range/cross range display (C-scope) of the out-of-the-window view, including accurately positioned runway obstacles

The overall configuration of the real-time ISVS radar processor is shown in Fig. 1. The system uses a 3D airport and terrain database, a high resolution MMW radar for all weather visibility and an video camera or FLIR device for high resolution detail. During operation, the 3D database is used to render a 3D perspective view of the underlying surface using GPS coordinates provided by the navigation system. In parallel a high resolution MMW imaging radar provides detailed visual information for obstacle avoidance, wire detection and near-surface navigation without the danger of hitting the ground. Finally, an on-board camera provides further image details. These three sources of information are subsequently registered and transformed to create a single, consistent representation to the pilot by means of an advanced feature registration and matching algorithm we developed during the course of the development process.

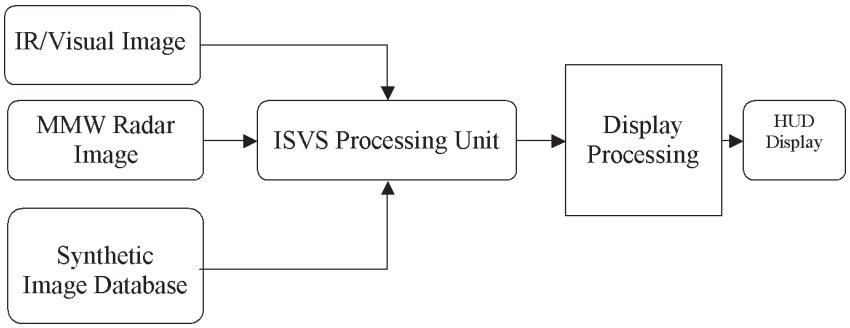

Figure 1. Functional overview of the real-time ISVS.

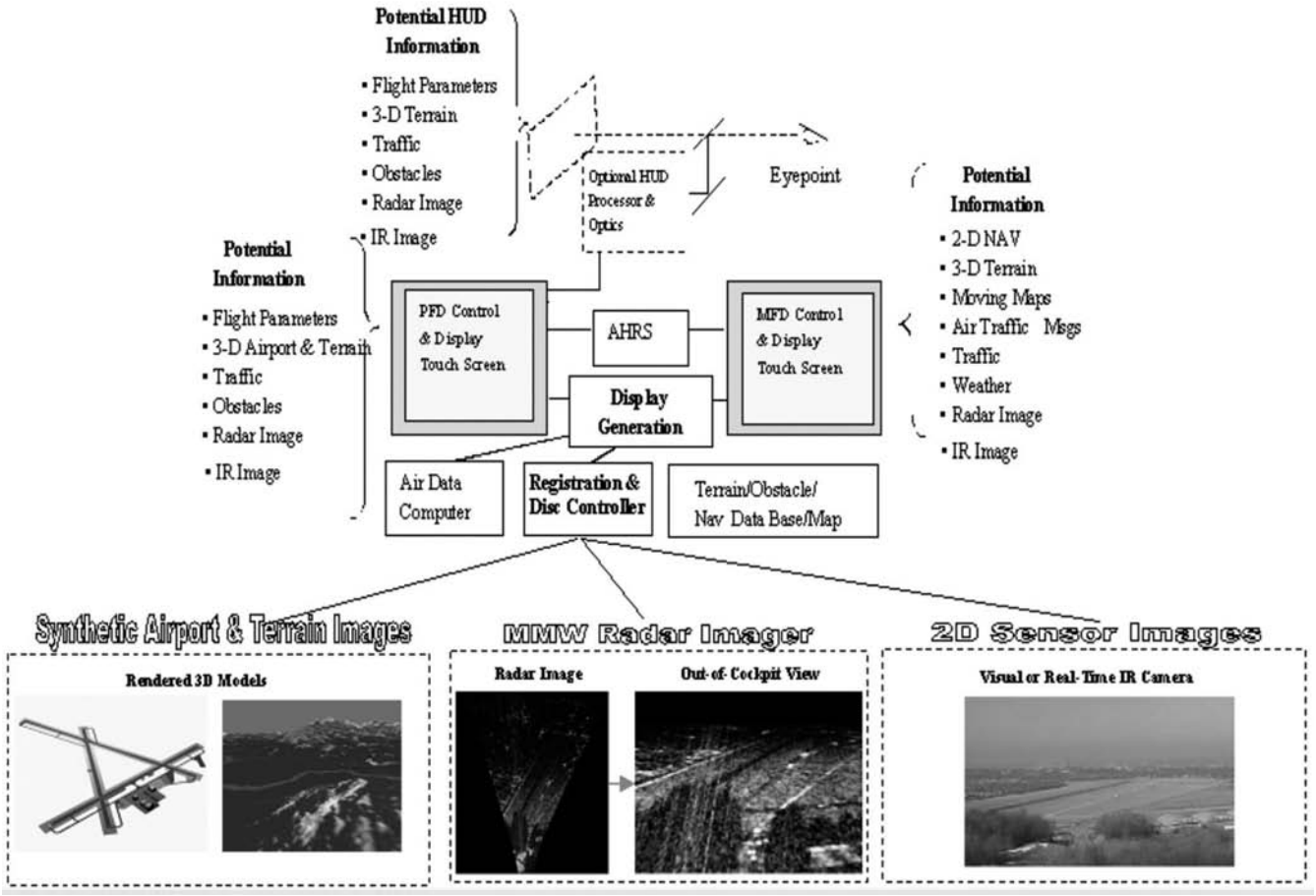

Figure 2. Detailed modular overview of WaveBand's real-time MMW radar SVS. *Supplemental Material—Figure 2 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

In this real-time version of our Intelligent Synthetic Vision System, three complementary sources of information are integrated to present most information to the pilot with minimal overload by the processing module. This output is then transformed by the display processing unit that creates a color transformation compatible with night vision goggles and also places symbology of detected obstacles visually overlaid on the scene. Figure 2 shows more details of the functionality of these modules and how they provide the final view to the pilot during the most critical phase landing.

The real-time image processor provides the processing and coordinate conversion necessary for displaying the image on a head-up-display and achieves a 30 fps scanning rate for the continuous linear scan. We use this MMW imaging radar to produce angular distance measurements of objects in the field. The WaveBand ISVS offers advantages that cannot be provided by either system alone. We used the image of the actual terrain and runway environment developed by WaveBand's 94 GHz MMW sensor[1] to accurately register and locate the GPS-based ISVS image. Our innovative approach dramatically increased the quality of the MMW image since it was rendered in real-time as a daylight scene from an on-board terrain and synthetic database. Furthermore we also devised a number of advanced regis-

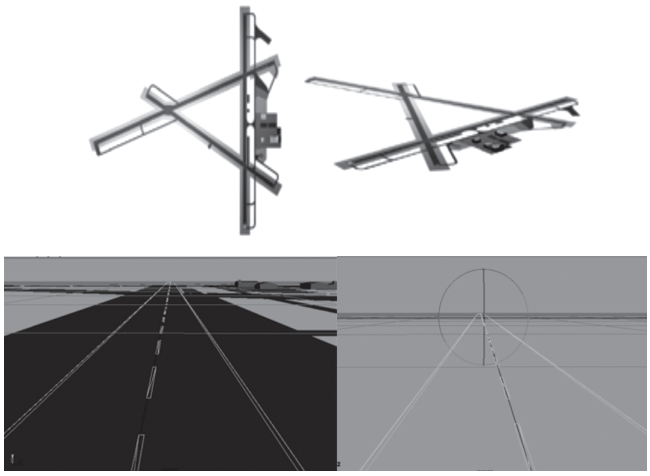

Figure 3. Synthetic runway images from our program used to generate binary templates within the image processing module. *Supplemental Material—Figure 3 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

tration algorithms, such as the use of the binary Hausdorff distance (see below), to help real-time operations. As a result our approach dramatically increases the integrity of the terrain image presented to the pilot when near the runway where image content integrity is most important. In the following sections we briefly review each of the key elements.

Synthetic 3-Dimensional Airport and Terrain Models

To allow the SVS to operate with either man-made objects such as buildings or airfields, we developed a 3D visualization module. This module is capable of displaying live video feed from either visual or IR sensors and 3D model views from DEM data. The position information for the 3D rendition can be obtained using the GPS-based sensors in the aircraft. This data flow is used to control the relative position of a *virtual camera* traveling "above" the digital airport model and terrain surrounding it. This step is a prequisite to our 3D template matching algorithm allowing the ISVS to generate 3D images of the environment within the image processing module itself so that parametric template images can be created and compared with the incoming sensor image on-line. Thus, in order to use these 3D models for direct comparison with the images received from the MMW sensor, we implemented a simple rendering algorithm that creates a wire frame, binary template image. The render algorithm uses a realistic computer graphics (CG) camera model with focal length, pixel aspect ratio, object distance, yaw, pitch, roll and pivot information. The 3D runway is imaged using this camera projection module as it moves in 3D space with respect to the camera. Figures 3 and 4 demonstrate the use of 3D models for a runway, and digital terrain, respectively.

Sensor Fusion and Integration Using Dynamic Binary Image Templates

Since the raw image data from the MMW image, 3D synthetic views and the optional video/IR camera have different resolutions, they are combined and further processed by a registration module. This registration

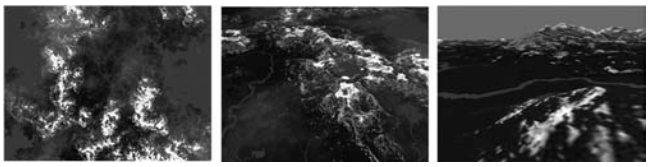

Figure 4. Example of real-time digital terrain images rendered using a virtual camera controlled by GPS signals (From left to right: top view, high approach angle and low-altitude mid-terrain images). *Supplemental Material—Figure 4 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

module is a real-time image processing module that creates a properly registered output image. To create the output image we consider the registration problem in the 2D image domain, whereas binary image features obtained from sensor image pairs are aligned using an iterative matching algorithm, as discussed in the following sections. The sensor image pairs that feed into the registration module are the perspective radar image, and a 2D image obtained from either a rendered view of a 3D synthetic model, or processed images from the optional image sensors. Both input images first pass through a feature detection module. This feature detection module uses a feature extraction algorithm that exploits a model of *neural attention* mechanisms in the *human visual system*.[2] It automatically locates a subset of binary features that can be used for optimal feature matching and tracking. Subsequently, the algorithm matches these detected binary features by minimizing a cost function based on a measure derived from the *Hausdorff distance*[3] to find the best matching image transformation parameters.

The Binary Hausdorff Distance

The use of the Hausdorff distance for binary image comparison and computer vision was originally proposed by Huttenlocher and colleagues.[4] In their study, the authors argue that the method is more tolerant to perturbations in the locations of points than binary correlation techniques since it measures proximity rather than exact superposition. Unlike most shape comparison methods the Hausdorff distance can be calculated without the explicit pairing of points in their respective data sets, A and B. Furthermore, there is a natural allowance to compare partial images and the method lends itself to simple and fast implementation.

Formally, given two finite point sets $A = \{a_1,...,a_p\}$, and $B = \{b_1,...,b_q\}$, the Hausdorff distance is defined as

$$H(A, B) = \max(h(A, B), h(B, A)) \tag{1}$$

where

$$h(A,B) = \max_{a\in A} \min_{b\in B} \|a - b\| \tag{2}$$

In the formulation above $||\cdot||$ is some underlying norm over the point sets A and B. In the following discussion, we assume that the distance between any two data points is defined as the Euclidean distance. $H(A,B)$ can be trivially computed in time $O(pq)$ for point sets of size p and q, respectively, and this can be improved to $O((p + q)\log(p + q))$. The function $h(A,B)$ is called the *directed* Hausdorff distance from set A to B. It identifies the point $a\in A$ that is farthest from any point of B and

measures the distance from a to its nearest neighbor in B. In other words, $h(A,B)$ in effect ranks each point of A based on its distance to the nearest point in B and then uses the largest ranked such point as the measure of distance (the most mismatched point of A). Intuitively, if $h(A,B) = d$, then each point of A must be within distance d of some point of B, and there also is some point of A that is exactly distance d from the nearest point of B. For practical implementations it is also important (due to occlusion or noise conditions) to be able to compare portions of shapes rather than providing exact matches. To handle such situations, the Hausdorff distance can be naturally extended to find the best *partial distance* between sets A and B.[4] To achieve this, while computing $h(A,B)$, one simply has to rank each point of A by its distance to the nearest point in B and take the Kth ranked value. This definition provides a nice property, that is it automatically selects the K "best matching" points of set A that minimizes the directed Hausdorff distance.

Realizing that there could be many different ways to define the directed ($h(A,B)$, $h(B,A)$) and undirected ($H(A,B)$) distances between two point sets A and B, Dubuisson and Jain[5] revised the metric and investigated 24 different distance measures based on their behavior in the presence of noise. In Ref. 3 the author has redefined the original definition of $h(A,B)$ proposing an improved measure, called the modified Hausdorff distance (MHD), which is less sensitive to noise. Specifically, in the formulation

$$h(A,B) = \frac{1}{N_a} \sum_{a \in A} \min_{b \in B} \|a - b\| \tag{3}$$

where $N_a = p$, the number of points in set A. In this article the author argues that even the Kth ranked Hausdorff distance of Huttenlocher presents some problems for object matching under noisy conditions, and conclude that the modified distance proposed above has the most desirable behavior for real world applications.

In the work we presented in Ref. 3, we adopted the MHD formulation of Dubuisson et.al.[5] and further improved its performance by introducing the notion of a *neighborhood function* (N_B^a) and associated *penalties* (P). Specifically, in our solution we assume that for each point in set A, the corresponding point in B must fall within a range of a given diameter. Let N_B^a be the neighborhood of point a in set B, and an indicator $I = 1$ if there exists a point $b \in N_B^a$, and $I = 0$ otherwise. The complete formulation of the "doubly" modified Hausdorff distance (M2HD) can now be written as:

$$d(a,B) = \max\left(I \cdot \min_{b \in N_B^a} \|a - b\|, (1 - I) \cdot P \right) \tag{4a}$$

$$h(A,B) = \frac{1}{N_a} \sum_{a \in A} d(a,B) \tag{4b}$$

$$\mathrm{H(A,B)} = \max(\mathrm{h(A,B)}\ ,\ \mathrm{h(B,A)}) \tag{4c}$$

The notion of similarity encoded by this modified Hausdorff distance is that each point of A be near some point of B and vice versa. It requires, however, that all matching pairs fall within a given neighborhood of each other in consistency with our initial assumption that local image transformations may take place. If no matching pair can be found, the present model introduces a penalty mechanism to ensure that images with large overlap are easily distinguished as well.

The modified Hausdorff measure (M2HD) as used here is ideal for applications, such as object tracking and recognition, where although overall shape similarity is maintained, the matching algorithm has to account for small local distortions and structural noise.[3]

Registering Radar Output to Synthetic or Sensory Images

To address the need of registering the MMW output obtained from our radar system with other imagery representing the same scene, we devised a method that builds upon our image processing modules developed earlier in this project and creates a transparent pipeline to be used both in conjunction with synthetic data sets as well as imagery delivered by other sensors. To achieve this goal the registration problem was considered in the 2D image domain, whereas binary image features obtained from a pair of images are aligned using an iterative matching algorithm. Figure 5 shows the overall diagram of the process. The inputs to our registration module are the i) original radar image is first transformed to a perspective "out-of-cockpit" view (upper/left), and ii) a 2D image obtained either as a rendered view of a 3D synthetic model such as an airport or images from another sensor, such as a visual or IR camera (upper/right). Both input images are first passed through a feature detection module that locates corners and optionally edges and straight lines in order to obtain binary features used for feature matching (lower/left). Finally, using the detected binary features we find the best matching affine transformation parameters via a local search algorithm based on the Downhill Simplex Method,[6] and use them to transform each image into the coordinate space of the other, i.e., the visual image overlay on the radar image or vice versa (lower/right).

Further details of the feature detection algorithms are shown in Fig. 6. Taking advantage of the predominantly man-made structures the MMW landing radar will operate on, we used a set of low-level feature detection algorithms that best capture the rectangular shapes dominating runway and airport images. Specifically, we start visual processing by applying corner detection and optional line detection algorithms using the Hough transform. We discussed these techniques as well as the effects of various parameter settings in detail in our previous reports.[7–9] We run the corner and line detector modules for both input images (radar and sensory/synthetic) and create two corresponding binary feature maps to be used in the matching stage that follows.

The matching algorithm selects one of the input images as a fixed reference and attempts to iteratively find the best matching affine transformation that rotates, scales and translates the other image in such a way that all binary features match with the smallest possible degree of error. As a measure of similarity we used a binary image metric, called the Hausdorff distance described above.[6–8] The main advantage of using a Hausdorff distance in this scenario is that it lacks the explicit need to pair image features up before matching takes place. Therefore an iterative search algorithm can easily be constructed that spans the possible space of affine image transformations and finds the best matching alignment that minimizes the computed Hausdorff measure (see above).

Although the possible space of affine image transformations is 6-dimensional, and there is no guarantee of

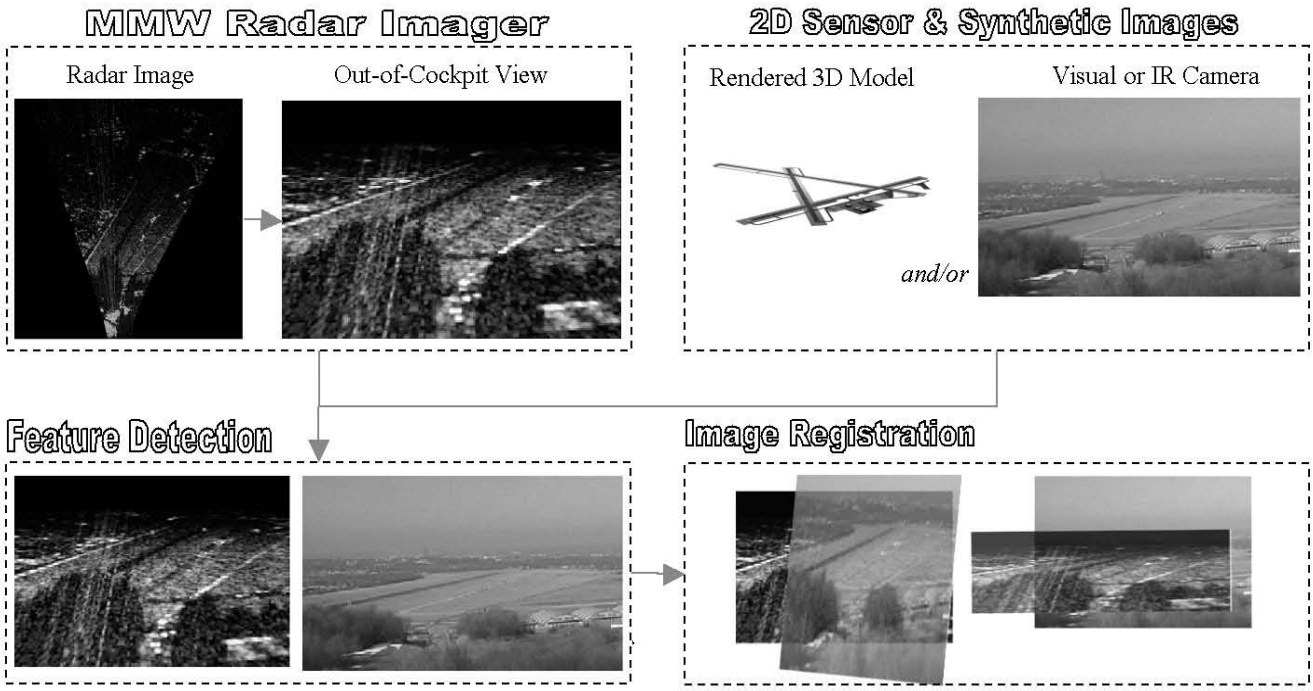

Figure 5. Overview of the image registration algorithm used to match MMW radar images with synthetic images and/or output from other sensors. *Supplemental Material—Figure 5 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

an absence of local minimums in the Hausdorff distance function for a given dataset our solution takes advantages of the natural physical constraints of the moving airplane. Specifically, the affine transformation values are a function of the yaw, pitch, roll as well as *x, y, z* positional parameters of the equipment as measured by GPS and mapped to the synthetic terrain data set via the parameters of the virtual camera. During landing the glide slope of the aircraft is 3 degrees and varies slowly and smoothly, thus eliminating many ambiguities in the matching process. Once the iteration is complete, the final set of transformation parameters are used to express the radar image in the coordinate space of the other sensor, or—by using the inverse—overlaying the visual image on the radar itself. Figures 7 and 8 show the respective outputs using the best matching affine image parameters. Both figures show the registered images in solid (left) and semi-transparent display mode (right).

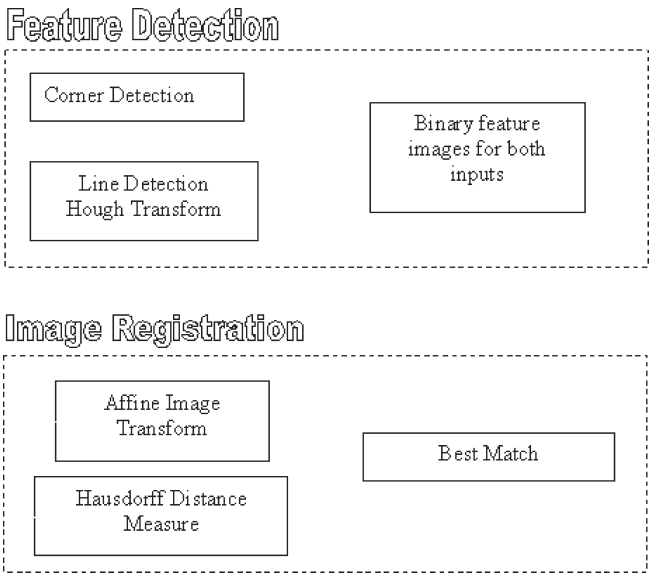

Figure 6. Iterative image registration algorithm using fast and efficient binary image metrics.

Evaluation of the Real-Time Image Processing and Visualization Engine

To create the real-time image processing and 3D rendered images of 3D airfield models and terrain data required by the image registration algorithms, we use a high performance real-time rendering engine that implements a variety of optimization techniques to render radar + visual + synthetic 3D data at high speeds without special demands on computing power. In fact, the render and image processing environment we developed runs on a personal computer and low cost commodity hardware. The performance of the rendering engine is demonstrated in Table I. We have tested the system on multiple computer platforms each exhibiting different configuration and performance characteristics. For testing purposes we focused primarily on CPU and graphics card performance and measured the *overall frame rate* as a function of 3D model *polygon size* to establish optimal performance. We compared two computers labeled Computer1 and Computer2, both running *DirectX* and *OpenGL*. Computer1 was a machine with one of the fastest CPU's and graphics cards commercially available today. Computer2 was an average computer configuration by today's standards. The same performance is expected to cost significantly less in the near future

Figure 7. Final result of the image registration algorithm showing visual image mapped onto the perspective radar output (see text). *Supplemental Material—Figure 7 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

TABLE I. Comparison of Real-Time Image Processing and Render Performance

	Computer1		Computer2		
CPU speed	Pentium4/3 GHz		Pentium4/850 MHz		
Memory	1 GBytes		256 MBytes		
Graphics Card	Radeon-9700		GEForce2-Go		
Memory	2 Gbytes/DDR3		1 Gbytes/DDR3		
	127,709	**97,305**	**62,859**	**33,017**	*polygons*
Computer1-DirectX	6.15	7.73	12.86	20.44	*fps*
Computer1-OpenGL	7.92	8.65	12.07	14.95	*fps*
Computer2-DirectX	31.89	35.72	42.84	52.07	*fps*
Computer2-OpenGL	29.25	31.33	40.27	44.75	*fps*

and therefore it will be widely available on home level computers.

In its current configuration the ISVS runs on two Pentium-based personal computers providing live radar images and image processing at 15–30 (fps) in 256 scanning angles and range resolutions between 2K–4K FFTs. One computer is responsible for computing the radar image, while the second one provides radar transformation, image processing and real-time alignment with other image sources.

Experimental Results

The combined ISVS was tested using in-flight images captured on an airfield in Brackett field, California. Figures 9 and 10 demonstrate the results obtained with our registration method using video imagery of an experimental aircraft approaching Brackett field airport. In both figures, the visual reference, recorded by a video camera mounted on the other wing of the aircraft, is properly aligned with features made visible by the MMW radar. Specifically, we use our perspective radar image transformation and correction algorithms to register the radar image with the video input. In the figures below, the visual image, the radar display and finally those two images overlaid, are shown in the top, center and bottom order, respectively.

Finally, using a Digital Elevation Model Database and satellite imagery a 3D real-time preview of the landing sequence was created as demonstrated in Figs. 11 and 12.

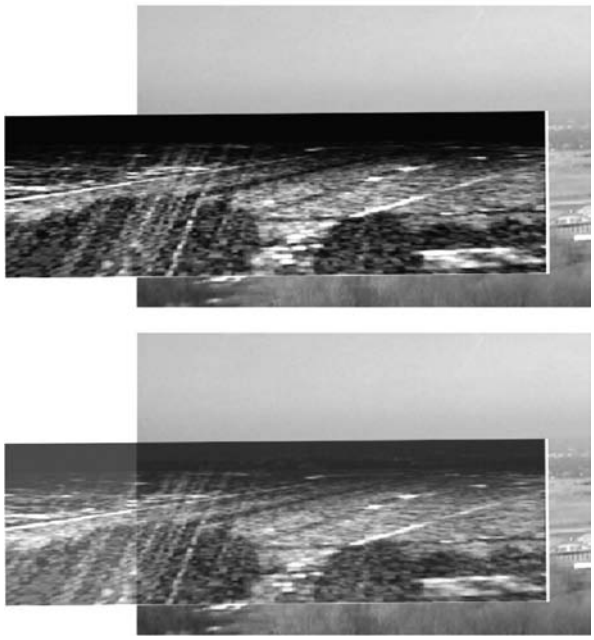

Figure 8. Final result of the image registration algorithm showing radar image mapped into the coordinate system of the visual reference (see text). *Supplemental Material—Figure 8 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

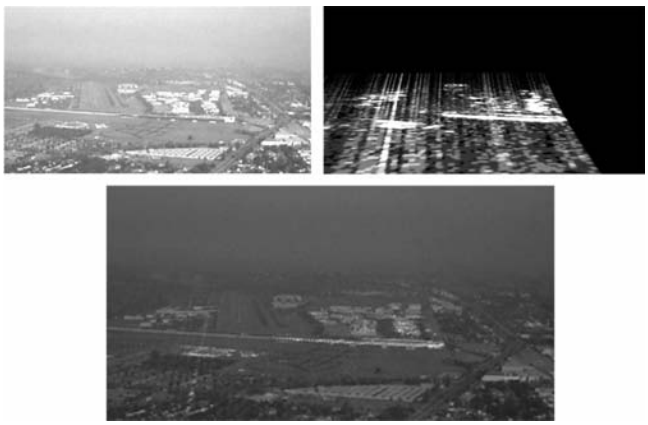

Figure 9. Image processing and registration of MMW radar images to visual reference. *Supplemental Material—Figure 9 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

Conclusion and Future Work

In this article we described a novel real-time image processing application that combines advanced millimeter wave radar imaging with computer vision and digital elevation databases. The resulting Intelligent Synthetic Vision System provides a robust solution to the fusion and overlay of information for pilots during standard maneuvers. By virtue of the MMW radar the system described herein is capable of "seeing through" dense fog and thus enabling navigation and landing in zero–zero visibility conditions.

Future work includes further enhancement of the image processing and recognition algorithms as well as using multiple MMW radar beams to create a fully 3D visual representation of the underlying scenery.

Figure 10. Image processing and registration of MMW radar images to visual reference. *Supplemental Material—Figure 10 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

Acknowledgment. This work was partly supported by NASA under contract # NAS1-02074. the authors wish to thank Digital Elite Inc. and VerAnim Bt.[10] for their support in creating the advanced 3D visualization environment.

References

1. WaveBand Corporation (2005) Technical documents and system description available from www.waveband.com.
2. B. Takács and H. Wechsler, A dynamic, multiresolution model of visual attention and its application to facial landmark detection, in *CVGIP Image Understanding* **71**(1), pp. 63–73 (1997).
3. B. Takács, Comparing face images using the modified Hausdorff distance, *Pattern Recognition* **30**(10), 1623–1636 (1998).
4. D. P. Huttenlocher, G. A. Klanderman and W. A. Rucklidge, Comparing images using the Hausdorff distance, *IEEE Trans. Pattern Analysis and Machine Intell.* **15**(9), 850–863 (1993).
5. M. P. Dubuisson and A. K. Jain, Modified Hausdorff distance for object matching. *Proc. IAPR Int. Conf. on Pattern Recognition,* vol. A., IEEE Computer Society Press, Los Alamitos, CA, 1994, pp. 566–568.
6. H. Cantzler, R. B. Fisher and M. Devy, Improving architectural 3D reconstruction by plane and edge constraining in *Proc. of the British Machine Vision Conf.* (BMVC '02), pp 43 – 52, Cardiff, U.K., September 2002.
7. B. Takács and L. Sadovnik 3-D target recognition and tracking using neural networks trained on optimal views, *J. Opt. Eng.* **37**(3), 819–828, 1998.

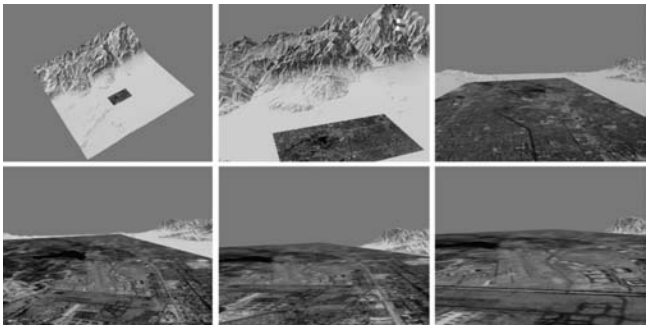

Figure 11. Synthetic image sequence of the initial approach (see text). *Supplemental Material—Figure 11 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

Figure 12. Synthetic radar overlay images during the final approach. *Supplemental Material—Figure 12 can be found in color on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

8. B. Takács and L. Sadovnik 3D target recognition using quasi-optimal visual filters, *SPIE Aerospace/Defense Sensing, Simulation and Controls*, SPIE Press, Bellingham, WA, 1998.
9. B. Takács, L. Sadovnik, V. Manasson, M. Wade, L. A. Klein, D. Wong, B. Kiss, B. Benedek, and G. Szijarto, Real-time visualization using a 2d/3d imaging mmwave radar, *Proc. SPIE* **5297** (2004).
10. Digital Elite Inc., www.digitalElite.net