Intelligent Multimodal and Hyperspectral Sensing for Real-Time Moving Target Tracking

Tao Wang[1], Zhigang Zhu[2]

Abstract—**Real time moving target tracking and identification with hyperspectral imagery is still very challenging with conventional sensors and algorithms. The increased information content of hyperspectral imaging has enabled improved classification and quantification of targets of interest. However, recording hyperspectral data for target classification is very time consuming. We design a sensor platform with multi-modalities, consisting of a dual-panoramic peripheral vision system and a narrow field-of-view hyperspectral fovea. Thus, we only need to capture hyperspectal images in regions of interest. This design is inspired by the human vision system where the periphery vision of the retina is used to detect motion and the fovea of the retina is used to recognize objects.**

The proposed intelligent sensors design is also supported by real-time algorithms for target detection, tracking and identification. Only hyperspectral data for areas of interest are captured for target classification and recognition. Important issue such as multimodal sensing component cooperation, region of interest extraction, target tracking, hyperspectral image analyzing and target signature identification are discussed.

Keywords—Hyperspectral Imaging, Multi-Modal, Regions of Interest, Tracking, Spectral classification.

I. Introduction

Tracking and recognizing moving targets in real-time is one of the important tasks for monitoring and surveillance. Traditional motion tracking methods with single sensor modalitieshave limitation in real-world scenarios. When using panchromatic video or color video data, blob tracking or kernel-based tracking [1] may perform well for targets with uniquely different colors or shapes. However, typical urban setting includes objects of similar sizes, shapes and colors and therefore they are hard to detect and identify. Multi-modal and/or hyperspectral sensors have been designed and used in such difficult situations. Hyperspectral imaging (HSI) involves data collection in dozens or hundreds of narrow, adjacent spectral bands. With extended information of the recorded data, image classification can be greatly improved. However, recording hyperspectral data is very time consuming. Also, for those applications that are time-varying and transient, or when there is a relative motion between objects and surroundings, the use of time-sequential HSI introduces the possibility of spatial mis-registration of images and a consequent distortion of the recorded spectra [2]. Thus, tracking moving objects with HSI in real time using conventional hyperspectral imaging techniques is problematic. To ease the spectral distortion, it will be the best that all voxels of the spectral data to be recorded simultaneously. This can be done using HSI snapshot techniques, as in [2]. But the spectral resolution is limited to few tens of bands due to the information bottleneck constituted by the space-bandwidth product of the detector arrays. To achieve two-dimensional, video-rate, hyperspectral imaging (of hundreds of spectral bands) with megapixel spatial resolution requires images to be recorded at the order of 10 Gpixels s^{-1} [3]. It is very expensive to capture and to process such high-rate data in real-time.

Recently, a great deal of effort has been put into adaptive and tunable multi-spectral or hyperspectral sensor designs with a goal to address the challenging problems of detecting, tracking and identifying targets in dynamic scenes. Representative large programs include: the DARPA's Adaptive Focal Plane Array (AFPA) Program [4], ARL's Advanced Sensor CTA [5], and NSF's Center for Mid-Infrared Technolgies for Health and Environment [6]. Even commercial real-time, spectral tunable, polaralization-enabled, hyperspectral imaging system are available [7]. In the design of a useful sensor, its field of view (FOV) is another important factor. Real-time omnidirectional (or panoramic) vision systems have been implemented by catadioptric designs using hyperbolic, parabolic or ellipsoidal mirror designs, as in [8] and [9]. Real-time panchromatic peripheral view and hyperspectral fovea have been proposed [3] to balance the requirements of both the FOV and real-time hyperspectral imaging. All of these works represent the state-of-the-art of adaptive sensor designs that have advanced the conventional sensor concepts. However, most of the sensor designs only have very limited data exploitation capabilities for optimizing hyperspectral data acquisition.

This work is supported by AFOSR under the Discovery Challenge Trusts (DCTs) Program, Award #FA9550-08-1-0199, and in part by AFRL/SN under Award No. FA8650-05-1-1853 and by NSF under Grant No. CNS-0551598.

1. Tao Wang is with Department of Computer Science, The City College of New York, 138th Street at Convent Avenue, New York, NY 10031; Department of Computer Science, The CUNY Graduate Center, 365 Fifth Avenue, New York, NY 10031. (phone: 646-217-8457; e-mail: twang@cs.ccny.cuny.edu).

2. Dr. Zhigang Zhu is with Department of Computer Science, The City College of New York, 138th Street at Convent Avenue, New York, NY 10031; Department of Computer Science, The CUNY Graduate Center, 365 Fifth Avenue, New York, NY 10031. (phone: 212-650-8799; fax: 212-650-6248; e-mail: zhu@cs.ccny.cuny.edu).

In this paper, we present a novel multimodal sensor design that has both omnidirectional (panoramic) view and hyperspectral fovea, and that integrates sensing and processing for real-time target detection and signature acquisition. This design is useful for a wide range of applications in real-time target recognition [10] and surveillance [3]. This sensor platform improves or differs from previous designs [2,3,5] in literature in three aspects:

(1) A panoramic view is provided instead of a normal wide-angle view.

(2) A dual scanning system is designed to obtain moving targets in a very effective and efficient manner.

(3) Active control of the hyperspectral sensor is added to facilitate signature acquisition of targets of various locations that can only be determined in real-time.

We note that not every sensor design can be fit in all different situations. The intention of this design is to test a system approach we have proposed to facilitate adaptive multimodal sensor designs [11], in order to reduce development time and system cost while achieving optimal results. This can be realized through an iterative process that incorporates simulation, evaluation and refinement of critical elements. Therefore, in addition to presenting a novel sensor design, this paper also provides some discussions of this design example that follows the philosophy of our system approach, by using scene and sensor simulation, and by integrating data exploitation into data acquisition.

The paper is organized as the following. Section 2 presents the design of the bio-inspired adaptive multimodal sensor platform - the dual-panoramic scanner with a hyperspectral fovea (DPSHF) for the task of tracking moving targets in real time. Section 3 discusses the target extraction method using conventional background subtraction. Section 4 illustrates our algorithm to estimate the next position of a moving target for adaptive fovea data collection. Section 5 discusses spectral classification in recognizing simple moving targets. Section 6 describes a simulation environment, the parameter configuration of the sensor platform, and provides some experiment results. Conclusions and discussions will be provided in Section 7.

II. SENSOR DESIGN

Our sensor platform design is inspired by the biological concept of the human vision system. The structure of a human eye is to focus light onto the retina. The peripheral region of the retina can respond well in dim light conditions. Its wide FOV facilitate the detection of motion of a target, and then the center of the retina, also know as fovea, can be turned to the specific target. The color information can then be obtained and compared with the database in the brain for recognition.

Our proposed sensor platform, the dual-panoramic scanners with a hyperspectral fovea (DPSHF) (Fig. 1), consists of a dual-panoramic (omnidirectional) peripheral vision and a narrow FOV hyperspectral fovea. It is an extension of the human vision system in that the peripheral vision is omnidirectional rather than just wide FOV, and the foveate vision is hyperspectral rather than just color. The intelligent sensor works as the follows. In the first step, two panchromatic images with 360-degree FOV are generated by rotating two line scanners around a common rotating axis, pointing apart to two slightly different directions. The angle difference between the two scanners can be adjusted for detecting and tracking moving targets with different velocities and distances. An initial angle is used at the beginning. Then the detecting results from the two scans can determine what the new angle difference should be - either decreased if a target is moving too fast, or increased if the target is moving too slow. With the use of line scanners in our sensor design, the geometric distortion problem can be avoid since the platform is motor-controlled and stabilized. There are two advantages of using line scanners that will be further amplified. First, a line scanner can have a full 306-degree horizontal FOV. Second, resulted images are inherently registered.

Moving targets can then be easily and quickly determined by the differences of the two panoramic images generated from two scanners. The next position and the time of a moving target can be estimated from the difference of two regions of interest (ROIs) that include the target. In real-time processing, the comparison is started whenever the second scan reaches the position of the first scan, therefore, only a small portion of images is used before full-view panoramas are generated. The detail of the target detection processing algorithm will be discussed in Section III.

Then, we can turn the hyperspectral imager with a specific focal length calculated based on the size of the object, and to the predicted region that includes the moving target. Thus, hyperspectral data is recorded more efficiently for only the ROIs that include possible moving targets. This can be done, for example, by using a control program implementing the processing algorithm on a FPGA chip. The two line scanners and hyperspectral imager are aligned so that they all share a single effective viewpoint. The spectral data can be efficiently recorded with a foveal hyperspectral imager (FHI) [3] which maps a 2D spatial image into a spatial 1D image. This is implemented by using a micro mirror as a fovea that intercepts the light onto a fiber optical reformatter (FOR) [12], which then forms a 1D array onto a dispersive hyperspectral imager (DHI) [13], producing a 2D hyperspectral data array with one dimension as spatial and the other as spectral. Then the result is combined with the panchromatic images generated by the dual-panoramic scanners to produce a co-registered spatial-spectral image.

The FOR is constructed from a coherent array of ribbon fibers. The reformatter maps an N x N array of fibers into a one-dimensional 1 x N^2 array. The fibers diameter determines the maximum spatial resolution of FHI. With 50 μm in diameter and arranged on a 56 μm, the spatial resolution can be 14 x 14 pixels. With reduced diameter of 10 μm, the spatial resolution can be significantly improved to 70 x 70 pixels. The blurring effect from cross-coupling of optical fibers was not a significant magnitude as shown in [2].

The advantage of using FHI is that we can have very high spectral resolution up to 1 nm. However, the spatial resolution of FHI is limited by FOR to 14x14 pixels (available) or to 70x70 pixels (improvement). One solution is to use image replication imaging spectrometer (IRIS) [14], which gives higher spatial resolution with the trade-off of a smaller number of spectral bands. In using IRIS, one approach is to select fewer narrow spectral bands to be sampled. Therefore, we can only retain spectral bands of interest by analyzing the spectral profile of a model object and then determining the significant bandwidth to be selected for the spectral filter. An acousto-optic tunable filters (AOTF) [15] [16] can be one of the solutions that can be rapidly tuned to pass specific wavelengths of light by varying the frequency of an acoustic wave propagating through an anisotropic crystal medium [17]. The AOTF operates very fast, usually in microseconds, so that it can be used in real-time processing. Indeed, we only need to tune the filter once as soon as the tracking target is decided.

III. Moving Target Detection

Most moving target detection algorithms use background subtraction as the first step. Here we use a simple background subtraction method to determine a region of interest for estimating the next position of each moving target. The initial background is acquired from the average of several scans captured by rotating one scanner pointing to the scene without any moving targets. Each of the two line scanners will generate a sequence of 1D image lines at each rotation angle. The panorama is constructed by combining all image lines. The advantage of using line scanner here is to avoid registration problem. The moving target can be easily extracted from subtracting each panorama from the background whenever the following relation holds for a pixel:

$$\left|I_{t+1} - B_t\right| > \tau \tag{1}$$

where τ is a "predefined" threshold which should be relatively small so that any small regions possibly including a moving target will be detected. In real-time process, whenever the second scan reaches the position of the first scan, a portion of spatial region is generated, thus, we can start the background subtraction without waiting the full panoramas to be compared. However, the background subtraction result (i.e., the difference image) has a lot of noises due to change of time and reflection of the light over time. Most significant noisy regions are resulted from changing of the illumination intensity over time. Since we have some intuitive knowledge that the moving target we want to track should be in some solid shapes after background subtraction, those regions have sparse points or irregular shapes may be discarded. We apply a morphological noise removal technique that consists of the opening and the closing operations [18]. Therefore, small sparse noises can be removed using the opening operation, while small holes can be filled using the closing operation. Those noises can be further reduced when we keep updating background information after each 360-degree rotation with

$$B_{t+1} = \alpha I_t + (1-\alpha) B_t \tag{2}$$

where α is kept small to prevent artificial "tails" forming behind moving objects [19]. In fact, lines of background are updated once each dual-lines is scanned before combination, therefore only those pixels belonging to the background are updated using the above formula. The background subtraction approach using two scans at different times can find the speed of a moving target easily, which will be described in the next section. Note that real-time target detection is achieved since the scanning and detection are performed simultaneously and continuously. Furthermore, with the dual panoramic scanners, even if background subtraction fails, targets can still be detected from the difference of two scans scanning the scene at slight different times.

IV. Moving Target Tracking

After the moving target detection step via background subtraction, we find two ROIs of the same target viewed at two different times resulting from the two different scans. The location and size differences of the two regions can determine the relative bearing angle of the hyperspectral fovea imager to zoom on to the moving target. The extracted regions positions from first scan and second scan can indicate which direction the target is moving to. Also, the size of two regions can indicate whether the target is moving closer to the sensor or farther. Therefore, we can calculate the next position where the target will be. Then, the ratio of the previous two regions can be used to estimate the new focal length of the hyperspectral imager.

The angle difference of two scans for two ROIs at different times t_i and t_{i+1}, can be used to predict the position of the next ROI having the moving target at the time, t_{i+2}, when the hyperspectral imager can be in place. Therefore, given the time t_{i+2}, we can estimate the panning and tile angles of the hyperspectral imager. Note that only the angles relative to the center of a region are needed.

The turning angles (i.e., panning and tilting) of the hyperspectral imager should be:

$$\theta_{t_{i+2}}^{(x,y)} = \theta_{t_{i+1}}^{(x,y)} + \frac{t_{i+2}^{(x,y)} - t_{i+1}^{(x,y)}}{t_{i+1}^{(x,y)} - t_i^{(x,y)}} (\theta_{t_{i+1}}^{(x,y)} - \theta_{t_i}^{(x,y)}) \tag{3}$$

where the superscript x and y correspond tothe panning angle (in the x-direction) and the tilting angle (in the y-direction), respectively. The angle θ_{t_i} corresponds to the angle position of a ROI at a time t_i as shown in Fig. 3.

The focal length of the hyperspectral fovea is inversely proportional to the desired FOV of the hyperspectral imager , α , in order to have the target in the full view of the FOV. The FOV angle can be estimated as

$$\alpha = \frac{R_{t_{i+2}}}{P^l} \tag{4}$$

where $R_{t_{i+2}}$ is the predicted size of the target region at t_{i+2}, and P^l is the number of line scan per angle. The relationship between $R_{t_{i+2}}$ and the previous two regions of the same target at different times can be expressed as

$$\frac{R_{t_{i+1}}}{R_{t_i}}(t_{i+1} - t_i) = \frac{R_{t_{i+2}}}{R_{t_{i+1}}}(t_{i+2} - t_{i+1}) \tag{5}$$

Then a hyperspectral foveal shot of a ROI from the calculation can be taken.

Thus, hyperspectral data is recorded in a more efficient way, only for regions of interest. It is possible for some regions not having true moving targets inside. This might be verified by simply subtract ROIs of the two scans from the matching regions in the background. If one of them has no motion change, we confirm that there is no moving object. This is an additional advantage of having two panoramic scanners.

V. Target Recognition

Recognizing moving target needs to compare a target spectrum associated with each pixel to its training spectrum. In our experiments, a spectral library was pre-built with some existing models. Currently, four vehicles with different spectral signatures captured in the same background are used. Various endmembers which represent spectral extreme that best characterize a material type of a target were selected, and their spectral curves were stored in the library database. We use the sequential maximum angle convex cone (SMACC) [20] to extract spectral endmembers and their abundance for every model target. In comparison to the conventional pixel purity index (PPI) [21] and N-FINDER [22], SMACC is a much faster and more automated method for finding spectral endmembers. Simply speaking, SMACC first finds extreme points or vectors that cannot be represented by a positive linear combination of other vectors in the data as a convex cone, and then a constrained oblique projection is applied to the existing cone to derive the next endmembers. The process is repeated until a tolerance value is reached, for example, max number of endmembers. Each endmember spectrum, defined as *H*, can be presented mathematically as a combination of the product of a convex 2D matrix contains endmember spectra as columns and a positive coefficient matrix:

$$H(c,i) = \sum_{k}^{N} R(c,k)A(k,j) \tag{6}$$

where *i* is the pixel index, *j* and *k* are the endmember indices, and *c* is the spectral channel index. Some endmembers might have less spectra differences in term of redundancy. Those can be coalesced based on a threshold so that the most extreme spectra are identified and used to represent the entire coalesced group of endmembers.

SMACC assumes hyperspectral data are calibrated to either radiance or reflectance. Assume the effective radiance reaching the sensor is $L(l,\theta,\phi,\lambda)$, where *l* is slant range from sensor to target, θ, ϕ and λ are the zenith angle, the azimuth angle and the wavelength, respectively. It follows the first-principle physical equation in [23]. If we are only interested in the reflective bands, the general equation can be further simplified as:

$$L(l,\theta,\phi,\lambda) = f\left(L_s(l,\sigma,\lambda),\ L_{ds}(\theta,\phi), L_{bs}(\theta,\phi,\lambda), L_{us}(l,\theta,\lambda)\right) \tag{7}$$

where σ is angle from the normal to the target to the sun, L_s is the solar radiance, L_{ds} is the downwelled radiance from the sky due to the atmospheric scattering, L_{bs} is the spectral radiance due to reflection from background objects, and L_{us} is the scattered atmospheric path radiance along the target-sensor line of site.

The general equation of radiance reaching the sensor shows the importance of various factors affect the results for further analysis. Even with the simplified term in Eq. (7), the reflectance measurements might be different from time to time for same target signature due to variations in illumination intensities and directions. To reduce those effects, the magnitudes of all target and training spectra were normalized using the continuum removal technique [24], which has been successfully used to highlight the spectral shape and magnify a small difference. A continuum is a mathematical functions used to isolate a particular absorption feature for analysis. The idea of a continuum is formed by defining high points of the spectrum (i.e., local maxima) and fitting straight line segments between these points. The final spectrum is resulted from dividing the continuum curve into its original spectrum (Fig. 4). It is not always advantageous to place every peak in a spectrum on the continuum curve, especially if the spectrum contains noise or poorly defined features. However, for recognizing some man-made objects like vehicles, it can be suitable for only determining significant parts of features.

With normalized data, a rapid matching filter [25] for target spectra was applied to match the target signature with the known signatures that have already been built in the spectral library. A least square difference (LSD) was calculated between each target spectrum and the reference endmembers. The total root-mean-square (RMS) error was used to form an RMS image for each endmember. One general RMS using the arithmetic mean is the Minimum Euclidean Distance (MED) [26]:

$$D(s,u_c) = \left\| s - \mu_c \right\| = \sqrt{\sum_{n=1}^{N} (s_n - \mu_{cn})^2} \tag{8}$$

where *s* is an unclassified pixel, μ_c is the mean of the class *c*, and *N* is the number of spectral dimensions. The MED computes the Euclidean distance between the current sample

s and the mean μ_c for all c, $(c = 1 \ldots C)$, and assigns the sample to the class with the minimum $D(s, u_c)$. The problem with the MED is that it does not account for the variance of the data. Other methods may be applied for certain cases, but will not be discussed here.

As a result, each pixel was classified either to a known object if the target spectrum was matched with the library spectrum of that object, or to an unknown object, for instance, the background. The matching results for all endmembers of a single object were combined in order to recognize the target. To distinct multiple objects from database, the results of different group of endmembers of different targets were compared.

VI. Experiment Trough Simulation

In order to test the sensor design concept and to integrate sensing and processing before a real sensor is fabricated, we simulate a realistic scene and model our sensor platform in DIRSIG [27], the Digital Imaging and Remote Sensing Image Generation tool, developed at the Rochester Institute of Technology (RIT). By scene simulation and sensor modeling, we hope to reduce the cost and development time in new sensor designs, together with the support of the tracking algorithm and the classification method.

The DIRSIG model is a complex synthetic image generation application which produces simulated imagery in the visible through thermal infrared regions [28]. The model is designed to produce broad-band, multi-spectral and hyperspectral imagery through the integration of a suite of first principles based radiation propagation sub-models. One important sub-model DIRSIG uses is MODTRAN [29], which is a radiation propagation model to generate a complex set of look-up tables to characterize the exoatmospheric irradiance, emitted and scattered radiances and path transmission.

Before scene simulation and sensor modeling, we need to set up different scenarios and configure the sensor parameters. One simple scenario we used involves a moving Humvee in an urban scene (Fig. 5). The average speed of the Humvee is about 33.6 mile per hour moving from east to west, passing through the main building in the scene. Our sensor platform is placed in front of this building. The scan speed of each line scanner can be set from 60 Hz to 100 Hz selectable, thus one entire 360° scan take from 6.0 seconds down to 3.6 seconds. This time constraint is not a problem for real-time target detection since the detection and scanning are continuous and simultaneous. The number of pixels per line in the vertical direction is set to 512 to match the horizontal scanning resolution. Only one spectral channel is captured for the dual line scanning. The focal length is fixed at 35mm for both line scanners, and the angle between the pointing directions of the two scanners is 10° so that the time the second scan reaches the position of the first scan is only about 0.1s. In theory, the time difference between two scans should be much less than one second to avoid a lot of uncertainty of action changes in moving vehicles. Two scanners are used so that (1) the more accurate direction and focal length of the hyperspectral fovea can be estimated; and (2) moving target detection can still be performed when background subtraction using a single scanner fails due to cluttered background, multiple moving targets, and the ego-motion of the sensor platform.

Fig. 2 shows the background subtraction result from two scans. With the help of noise removal, two ROIs are further extracted (Fig. 6). The focal length of hyperspectral imager is automatically adjusted according to the target detection results generated from two line scanners. To simulate the hyperspectral imager, we model a frame array sensor in DIRSIG with a spatial resolution of 70 x 70. The spectral resolution is 0.01 μm and the range goes from 0.4 μm to 1.0 μm. Different portion of bandwidth can be selected and determined by analyzing model spectral profile.

Four model vehicles are used. Their endmembers that are extracted using SMACC under different orientations are stored in the library database (coalesce value = 0.1), as shown in Table I.

Table I. Numbers of endmembers of four models

	Endmembers			
Model	Front	Side	Back	Total
Humvee	3	4	2	5
Truck 1	2	3	2	3
Truck 2	2	3	2	3
Small Car	2	2	2	2

Note that the total endmembers are not the sum of all endmembers from the front, the side and the back. It is the number of unique endmembers that are not duplicated in all orientations. For example, from the side, the Humvee model has 4 endmembers including those of the side body, front cover, top cover and tires; however, it does not have any signature information from the back, where back body and tires are included. Therefore, the total significant endmembers for the Humvee from all orientation are 5.

After normalizing, the false-color image was generated for each pixel that belongs to the known object in the database. Fig. 7 demonstrated a simple spectral matching result of spectral curve of target pixels and sample pixels. For demonstration purpose, only two library spectral curves (Fig. 8d) were shown here to be compared with target spectra to recognize the object in this case. The results of matching differences of the two curves were shown in Fig. 7a and 7b. Their statistical data with 0.0 mean are shown in Table II.

Table II. Results of matching filtering

	Curve A	Curve B
Min	-1.25	-0.405
Max	1.416	1.044
Std	0.281	0.122

To recognize the whole object, all matching results were combined to a single image profile and a binary decision was

made using a threshold found by experience (Fig. 7c). It shows that even though we had a large area of shadow on the top of the Humvee caused by trees, we were still able to detect large parts of the Humvee body. The resulted false-color spectral image was then embedded into the panchromatic peripheral image which was captured with HSI simultaneously to produce a co-registered composite image that indicated the next position of the moving target. Thus, tracking and recognizing moving object can be demonstrated in real time. In Fig 8 we show a single image snapshot captured the same time as the HSI was recorded, and the result was combined to produce a co-registered composite image.

VII. CONCLUSIONS

We mainly described our bio-inspired multimodal sensor design that enables efficient hyperspectral data collection for tracking moving targets in real-time. This design and the related processing steps are tested through sensor modeling and realistic scene simulation. By simulation, various components can be reconfigured or replaced for specific situations or tasks. The image processing algorithms are designed only to demonstrate the basic idea of effectively capturing hyperspectral data in regions of interest based on data exploitation. Needless to say, more sophisticated algorithms need to be developed for more challenging tasks. We only described one spectral classification method for recognizing the object. More precise and efficient hyperspectral classification routines may be applied.

Future work includes the evaluation of a similar sensor design mounted on a moving platform. Geometric distortion will then need to be considered. More complex situations (both background and targets) will be simulated and tested. Then more complicated results need to be evaluated precisely. However, the theoretical underpinnings of this work and the simulated experiments suggest that the capability of our system approach is appropriate for intelligent adaptive multimodal sensors design.

REFERENCES

[1] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-Based Object Tracking", *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 25, no. 5, May 2003.

[2] A R Harvey, J Beale, A H Greenaway, T J Hanlon and J Williams, 2000 "Technology options for imaging spectrometry", *Imaging Spectrometry VI Proc. SPIE* vol 4132, ed M Desour and S Shen (Bellingham, WA: SPIE)

[3] Fletcher-Holmes, D. W. and A. R. Harvey, 2005. "Real-time imaging with a hyperspectral fovea," *J. Opt. A: Pure Appl. Opt.* 7, 2005, pp. S298-S302

[4] R. Balcerak, 2007. DARPA's Adaptive Focal Plane Array (AFPA) Program, Available: http://www.arpa.mil/mto/programs/afpa/index.html

[5] A.C. Goldberg, B. Stann, N. Gupta, 2003, "Multispectral, hyperspectral, and three-dimensional imaging research at the U.S. Army research laboratory", *Proceedings of the Sixth International Conference of Information Fusion*, 2003. Vol. 1, 2003, pp: 499-506

[6] MIRTHE, Center for Mid-Infrared Technologies for Health and the Environment. Available: http://www.mirthecenter.org/

[7] POC, POC's Real-Time Hyperspectral Imaging, http://www.poc.com/emerging_products/rthi/default.asp

[8] Z. Zhu, E. M. Riseman, A. R. Hanson, 1999. "Geometrical modeling and real-time vision applications of panoramic annular lens (PAL) camera," *Technical Report TR#99-11*, Computer Science Department, University of Massachusetts Amherst, February, 1999. Available: http://www-cs.engr.ccny.cuny.edu/~zhu/UMASSTR99-11.ps.gz)

[9] S. K. Nayar and S. Baker, 1997. "Catadioptric image formation", *Prof. DARPA Image Understanding Workshop*, May 1997: 1431-1437

[10] Hsuan Ren, Chein-I Chang, "Automatic spectral target recognition in hyperpectral imagery", *Aerospace and Electronic Systems, IEEE Transaction on*, Oct, 2003, vol 39, issue 4, pp 1232-1249

[11] Z. Zhu, "A System Approach to Adaptive Multi-modal Sensor Designs", Available: http://www-cs.engr.ccny.cuny.edu/~zhu/Research/MultimodalSensors.html

[12] Fibreoptic Systems Inc. 60 Moreland Rd, Unit A, Simi Valley, CA 93065, USA (http://www.fibopsys.com/)

[13] Headwall Photonics, Inc., 601 River Street, Fitchburg, MA 01420, USA.(http://www.headwallphotonics.com/)

[14] A. R. Harvey, D. W. Fletcher-Holmes, A. Gorman, K. Altenbach, J. Arlt, N. D. Read, "Spectral imaging in a snapshot," *Spectral Imaging: Instrumentation, Applications, and Analysis III*, Proc. SPIE Vol. 5694, p. 110-119, (2005)

[15] N. Gupta, "Acousto-Optic Tunable Filters," *Opt. Photon.* News 8, 23-27 (1997)

[16] M. S. Gottlieb, "Acousto-Optic Tunable Filter," in *Designs and Fabrication of Acousto-Optic Devices*, .A. P. Goutzoulis and D. R. Pape, eds., Marcel Dekker, New York, pp. 197-283 (1994).

[17] H. H. Morris, C. C. Hoy, and P. J. Teado, *Appl. Spectrosc*. 48, 857 (1994)

[18] E. R. Dougherty, "An Introduction to Morphological Image Processing", SPIE Press, Ltd, US, 1992

[19] A. M. McIvor, 2000 "Background Subtraction Techniques," *In Proc. of Image and Vision Computing*. 143-153, 2000

[20] J. Gruninger, A. J. Ratkowski and M. L. Hoke. "The Sequential Maximum Angle Convex Cone (SMACC) Endmember Model". Proceedings SPIE, Algorithms for Multispectral and Hyper-spectral and Ultraspectral Imagery, Vol. 5425-1, Orlando FL, April, 2004.

[21] J. W. Boardman, F.A. Druse, and R.O. Green, "Mapping target signatures via partial unmixing of AVIRIS data." *Fifth Annual JPL Airborne Earth Science Workshop*, Vol. 1, AVIRIS Workshop, pp. 23-26, 1995

[22] M. F. Winter, "N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data." *Proceedings SPIE*, Vol. 3753, pp. 266-275

[23] J. R. Schott, "Remote Sensing: the image chain approach, 2nd ed". Oxford, NY: Oxford University Press, 2007.

[24] F. A. Kruse, and A. B. Lefkoff, "Knowledge-based geologic mapping with imaging spectrometers." *Remote Sensing Reviews, Special Issue on NASA Innovative Research Program (IRP) results*, Vol. 8, pp. 3 – 28, 1993.

[25] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, pp. 779-785, 1994

[26] J. Blackburn, M. Mendenhall, A. Rice, P. Shelnutt, N. Soliman, J. Vasquez, "Feature Aided Tracking with Hyperspectral Imagery", *Signal and Data Processing of Small Targets 2007*. Edited by Drummond, E. Oliver, Teichgraeber, D. Richard, *Proceedings of the SPIE*, Vol 6699, pp. 66990S-66990S-12, 2007.

[27] DIRSIG, http://dirsig.cis.rit.edu/, last visited August, 2008

[28] J. R. Schott, S. D. Brown, R. V. Raqueño, H. N. Gross, and G. Robinson, "An advanced synthetic image generation model and its application to multi/hyperspectral algorithm development", *Canadian Journal of Remote Sensing*, Vol. 25, No. 2, June 1999

[29] MODTRAN4 http://www.kirtland.af.mil/library/factsheets/factsheet.asp?id=7915

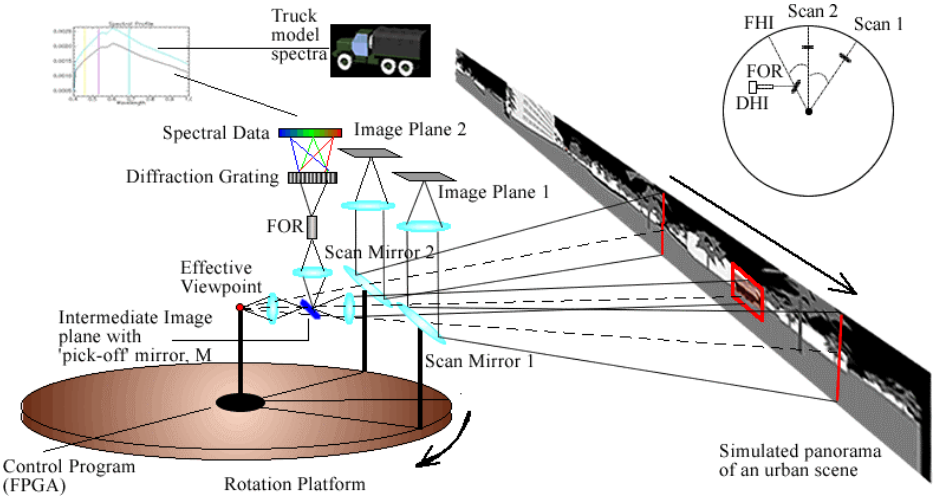

Fig. 1. The design concept of dual panoramic scanners with a hyperspectral fovea (DPSHF). The dash lines indicate the single viewpoint of both the foveal hyperspectral imager and the two line scanners. Note that the actually configuration is not shown in diagram. It can be adjusted during simulation and will be described in section VI.

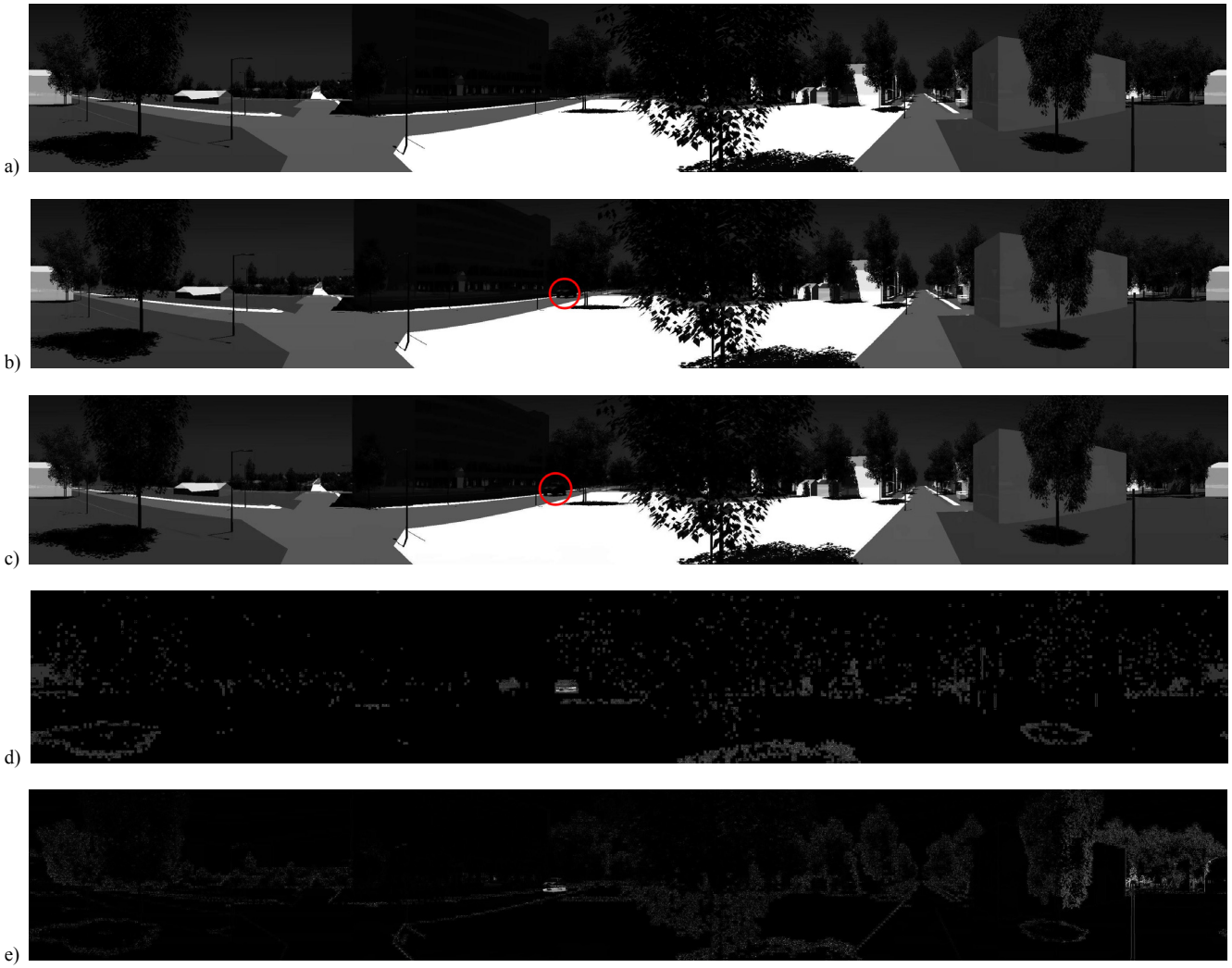

Fig. 2. All 360° panoramic images (512 x 3600) shown here are integrated from line scans captured by the dual-panoramic scanners. a). An initial background scene is simulated in single band and is the average calculated from several scans. b). Panoramic image from the first scanner, with the moving target indicated inside a red circle. c). Panoramic image from the second scanner, again the same moving target indicated inside a red circle. d). Background subtraction result for the image captured by the first scanner. e). Background subtraction result for the image captured by the second scanner. (Note: in order to view the shapes of the targets clearly and make the background noise more visible an offset of 50 was added to the intensities of all pixels in d and e except the pure black (i.e,, zero intensity)).

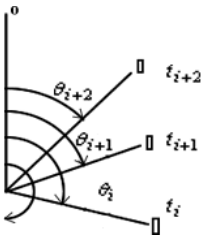

Fig. 3. Angle estimation of a moving target based on different times. The platform is rotating clockwise where the target is moving in the opposite direction. The formula is also applied for a target moving in other directions.

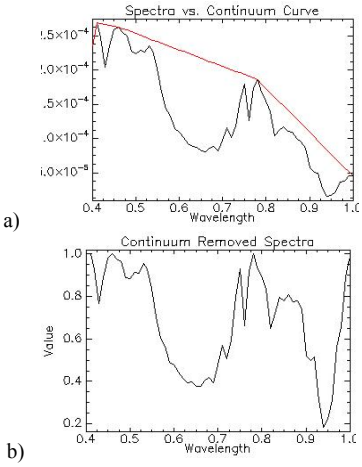

Fig. 4. a) The top straight lines show the continuum curve, C; where the bottom curve shows the original spectrum, S. b). The result continuum-removed spectra $S_{cr} = S / C$

Fig. 5. A simulated urban scene image at latitude=43.0° and longitude=77.0°, captured at 1000 meters above the scene. A Humvee is placed in the center road moving from right to left. The bigger red circle shows the starting position, and the red lines show the moving trajectory. The sensor platform showing in blue square is placed in front of the main large building, and about 5 meters high from ground. The sun position at 14:00 is 21.8° from the z-axis and 299.4° clockwise from the y-axis, where the y-axis points from left to right and the z-axis points out of the paper in the above figure.

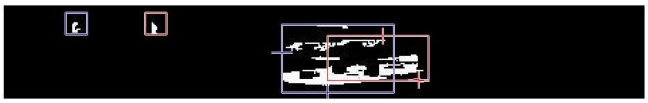

Fig. 6. ROIs extracted from two scans. The same small potions of the two previous panoramas (Fig 2d and fig 2e) are cropped and zoomed to show a close-up view of those ROIs where the targets are detected. Red boxes show the ROIs from the first scan. Blue boxes show the ROIs from the second scan.

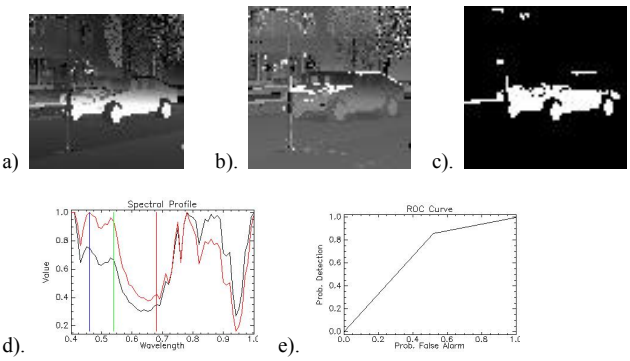

Fig. 7. Spectral matching result of target pixels with model spectra. a) Matching result of the first curve in model spectral library. The brighter a pixel is, the less the difference is. b) Matching result of the second curve in model spectral library. The darker a pixel is, the larger the difference is. c) Combination of all matching result, binary decision is made at the threshold 0.6. d) Two selected normalized spectral curve of a model Humvee. e). ROC curve of a classification result from matching filter.

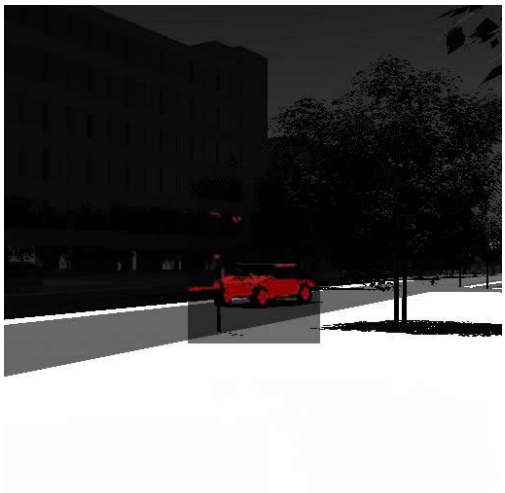

Fig. 8. Spectral classified image was embedded into a panchromatic peripheral image snapshot (512 x 512) showing the next movement of the target Humvee. Pixels inside the fovea are assigned false-color red and black in watershed to indicate similarities to the sample spectra of the Humvee model and non-target background points, respectively.