Studies on Image Fusion Techniques for Dynamic Applications

D. W. Repperger[1] A. R. Pinkus[1] J. A. Skipper[2] R. Woodyard[2]

[1]Air Force Research Laboratory
AFRL/RHCV, WPAFB,
Dayton, Ohio 45433-7022
d.repperger@ieee.org

[2]Wright State University
Department of Biomedical, Human Factors and Industrial Engineering,
Dayton, Ohio 45435

***Abstract*— A survey of present methods and current techniques being pursued by the US Air Force for image fusion and registration is conducted. Formulating the problem within a signal detection theory framework provides a unique thrust.**

I. INTRODUCTION

Modern image fusion techniques need to become increasing sophisticated as the hardware that gleans data from complex visual scenes becomes ever-increasing more powerful in its ability to extract information from an exogenous environment. Once the data gathering hardware systems are calibrated and registered (with each other), the difficult problem is then to integrate or fuse the data in a manner which improves the quality of the overall visual rendering that may be presented to a decision maker. If the fused image is "information rich" in some sense and better than any individual or constituent image, then the fusing of the information has been a productive task and adds value to the overall process of object identification.

The first step in image fusion is the static registration of the individual data gathering sources. There exist a number of techniques for this task, and the literature is reviewed on some of the more popular methods. One modern example even includes the concept of "image registration energy" [1], whereas a minimum energy framework is synthesized to develop the coordination between the individual camera sources. There also exist alternative image registration methods [2] based on principals of optimality employing least squares methods which have certain computational advantages. The registration problem has to be validated and calibrated against a test bed system.

After the initial registration problem has been addressed, there exists a number of ways the actual fusion of the images can be conducted. First the static case is discussed. One well known example is the Laplacian Pyramid method [3] in which different camera sources of varying resolution size, field of view, and other dynamic attributes are combined. Such a procedure involves extracting key control points within an image and coordinating the control point data in a manner which is beneficial overall. Again, the recurring point is that the fused image must always result in an improvement over any individual or constituent image, alone. The use of statistical methods [2], such as maximum likelihood, can be very beneficial if proper definitions of the density functions can be determined which make up the objects to be identified.

The final step would be the image fusion under dynamic conditions [4]. As the image gathering sources move with time, in order to identify an object, a model of relative image motion must be incorporated in the algorithm. This also affects the algorithms used for the image fusion problem under motion. Several techniques to provide adequate models of objects in motion are incorporated into the fusion algorithm. Again, statistical methods such as maximum likelihood have the advantage that they provide statistical optimality for certain types of formulations. For example, for minimization of the type 2 error, for a given and fixed type 1 error in identification, the Neyman-Pearson method provides optimality advantages for the dynamic application of image fusion techniques.

Another important variable to consider in the image fusion methodologies is the computational efficiency of the various algorithms employed. Certainly, as the object of interest to be identified changes dynamically, and data are quickly captured, the efficacy of the dynamic fusion technique is heavily predicated on the numerical efficiency of the procedures employed. A survey of the extant methods and the possible computational bottlenecks that may occur due to different analysis procedures are discussed.

Finally, a comparison of the image fusion methods will be investigated across various application domains. For example, [5,6] there are a number of successful methods which may provide advantages in situations where the identification satisfies optimality principles (least squares). A discussion of the present methods and the ongoing work at the AFRL will be presented.

II. FORMULATION OF THE PROBLEM

Fig. (1) displays the problem of interest. Multi camera sources may be viewing a moving object of interest. In Fig. (2) the multi cameras are translating while the entity of interest is stationary (in a relative sense). Fig. (3) shows how

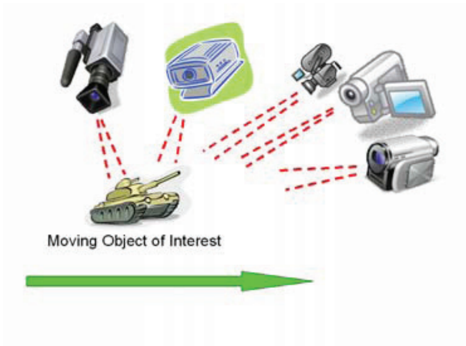

Fig. (1) – Stationary Multi Cameras

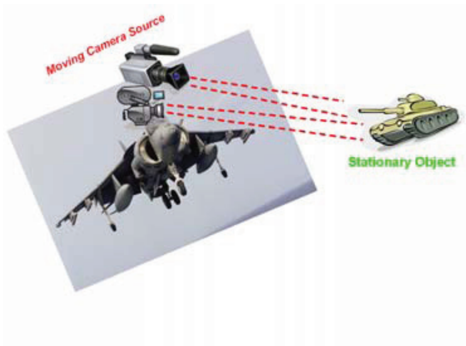

Fig. (2) – Multi Cameras are Moving

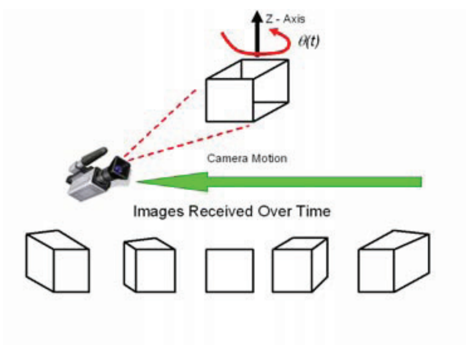

Fig. (3) – A Cube Object in Different Orientations

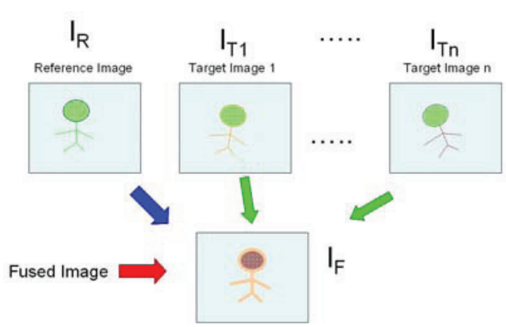

Fig. (4) – The Classical Image Fusion/Registration Problem

an object may provide a different image, depending on the orientation, even for the same entity. In Fig. (4) the classical registration/fusion problem is posed, i.e. the fused image is more "information rich" than any of its constituent images. The question is how to best register and fuse the individual images? This procedure also generalizes to hyper spectral data such as portrayed in Fig. (5). In that diagram, various

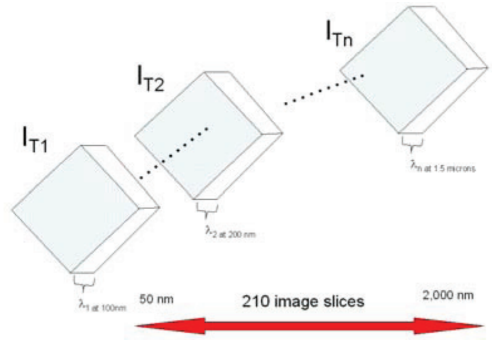

Fig. (5) – The Hyper Spectral Imaging Problem

slices of an image have particular signatures inherent of an object. In Fig. (6), a possible paradigm is discussed using maximum likelihood methods [2,4]. The problem is then cast within the framework of optimal estimation, as considered in Fig. (7).

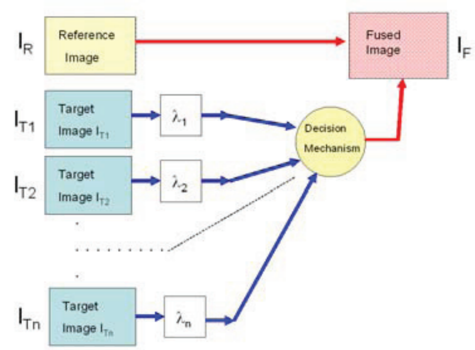

Fig. (6) – Maximum Likelihood Estimation Methods

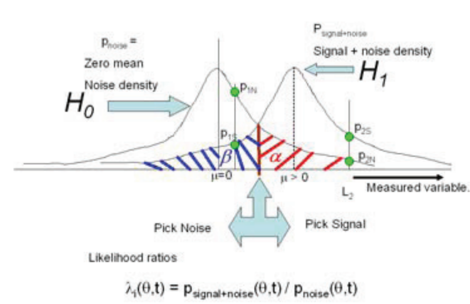

Fig. (7) – A Signal Detection Theory (SDT) Viewpoint

Some of the immediate questions the registration/fusion problem brings to light include:

(1) In Fig. (6), how to best select within and across the candidate sources target (images) to optimize the information received at the final fused image?

(2) If certain image sources are giving similar (correlated) information, how to "weed out" those sources that are not beneficial in adding new information to the overall fused image? This is to be considered even though the quality of the similar images are high, they may unfairly over weigh one source of information.

III.A SIGNAL DETECTION THEORY (SDT) FRAMEWORK

With reference to Fig. (8), the approach here will develop analogies between the classical problem of detecting whether a signal or noise alone has been received. In Fig. (8) it is desired to determine when the signal *S(t)* went high? Only measurement of the signal + noise is available and the goal is to determine if the signal *S(t)* is in the high or low state.

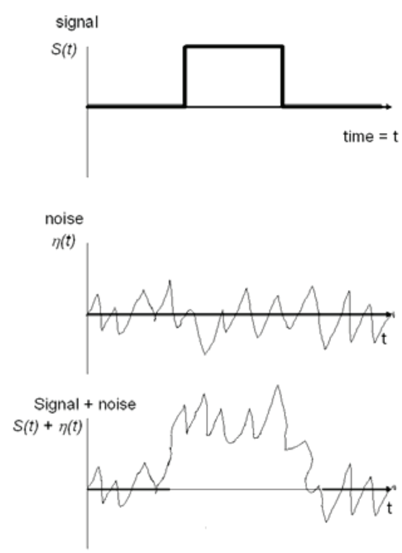

Fig. (8) – Signal, Noise, and Signal plus Noise

IV. FORMULATING THE IMAGE FUSION PROBLEM IN SDT

Since brevity must be the style here, the goal will make the selection of the target images from Fig. (6) and their respective weightings analogous to signal detection theory. The rationale for this approach is due to the fact that an extensive background knowledge exists in the SDT area which can be leveraged to assist in the formulation of the image fusion problem. To briefly outline the procedure, the following steps will cast the image fusion into a framework such that SDT principles can be employed.

<u>Step 1:</u> In Figs. (4,6) develop a set of difference images:

$$D_{1i} = I_R - I_{Ti} \qquad (1)$$

Where I_R is the reference image (usually the visual image) and the one the decision maker would normally have available if no other target images were collected. The matrix D_{1i} will be treated if it was the signal + noise variable and the question to be asked is whether D_{1i} is noise alone or signal plus noise when the signal may be high?

<u>Step 2:</u> To determine if D_{1i} is noise alone or signal + noise when the signal is high, a correlation operation on the matrix can be defined as follows:

Let the correlation function generate a matrix γ_{1i} defined as follows on the difference matrices D_{1i}:

$$\gamma_{1i} = \mathrm{corr}(D_{1i}) \qquad (2)$$

Now let ε_{1i} denote a norm squared of the matrix γ_{1i} such as the Frobenius norm:

$$\varepsilon_{1i} = \|\gamma_{1i}\|^2_F \qquad (3)$$

Note the Frobenius norm of a matrix A may be easily calculated as follows:

$$\|A\|^2_F = \sum_{i,j}^{n,m} (a_{ij})^2 \qquad (4)$$

<u>Step 3:</u> The first pass through the difference matrices D_{1i} has now been completed. A minimum threshold for ε_{1i} is defined based on a threshold scalar quantity $\Delta > 0$. The quantity Δ is determined from an image with only noise. We can now assign a certain number of the λ_i values via the following Rule 1:

<u>Rule 1:</u> For image I_{Ti} if $\varepsilon_{1j} = \|\gamma_{1j}\|^2_F < \Delta$ then assign

$$\lambda_j = \varepsilon_{1j} \qquad (5)$$

Thus the image I_{Ti} has been eliminated and its λ_j has been determined for Fig. (6).

<u>Step 4:</u> To continue the sieve operation, assume I_{T2} has not been eliminated. If, at this point, I_{T2} had been eliminated, then move up to I_{T3} or I_{T4} or the next highest target image that has not been eliminated. Again, suppose I_{T2} has not been eliminated, then define a second set of difference matrices as follows:

$$D_{2j} = I_{T2} - I_j \qquad j=3,\ldots,n \qquad (6)$$

and repeat the following calculations:

$$\gamma_{2j} = \mathrm{corr}(D_{2j}) \qquad (7)$$

$$\varepsilon_{2j} = \|\gamma_{2j}\|^2_F \qquad (8)$$

And invoke Rule 2:

<u>Rule 2:</u> For image I_j if $\varepsilon_{2j} = \|\gamma_{2j}\|^2_F < \Delta$, then assign

$$\lambda_j = \varepsilon_{2j} \qquad (9)$$

Again, this will eliminate some of the I_j images.

To continue, presume I_3 has not been eliminated at this point. If it has, then go to the next highest target image. Again define a third difference matrix:

$$D_{3k} = I_3 - I_k \qquad k = 4,\ldots,n \tag{10}$$

and repeat the following calculations:

$$\gamma_{3k} = \mathrm{corr}(D_{3k}) \tag{11}$$

$$\varepsilon_{3k} = \|\gamma_{3k}\|^2_F \tag{12}$$

And invoke Rule 3:

Rule 3: For image I_k if $\varepsilon_{3k} = \|\gamma_{3k}\|^2_F < \Delta$, then assign

$$\lambda_k = \varepsilon_{3k} \tag{13}$$

Again, this will eliminate more of the I_k images.
The process now repeats itself with the next set of difference matrices (presuming I_4 has not been eliminated up to this point).

$$D_{4l} = I_4 - I_l \qquad l = 5,\ldots,n \tag{14}$$

until all λ_k are defined. It may take up to n-1 rules, but all λ_j will eventually be assigned in proportion to the contribution they add in a correlation sense. Also, if ε_{3k} in equation (12) $\geq \Delta$, then the remaining λ_j may be set to their ε_{3k} values.

V. STATISTICAL TESTING WITH CHI SQUARE

Another advantage of the procedure outlined herein is that the classical decision mechanism as portrayed in Fig. (7) can now be modified by a Chi square test. This is because the square of the Frobenius norm is used in the decision process. To illustrate the new testing scenario, Fig. (9) shows how this decision process will be implemented. To confirm the efficacy of the method presented so far, a Monte Carlo simulation was conducted. Fig. (10) shows the candidate

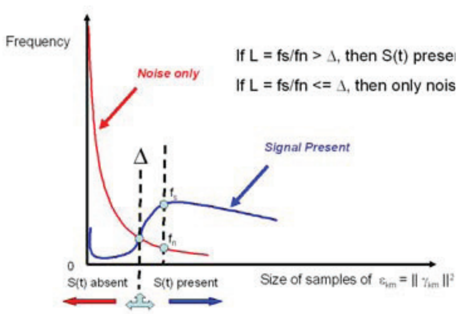

Fig. (9) – Using a Chi Square Likelihood Ratio Test

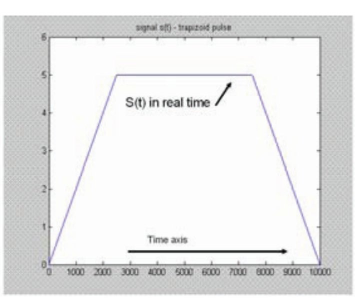

Fig. (10) – The Candidate signal S(t) to be Detected

signal S(t) which is desired to be detected in the high state. Fig. (11) shows the zero mean white Gaussian noise to be added. Fig. (12) is the combination of both plots. In Fig. (13) the algorithm described herein was employed on the Monte Carlo run of 10,000 points. It is clear the decision rule empirically determined in Fig. (13) via Monte Carlo simulation closely follows the theoretical version as seen in Fig. (9).

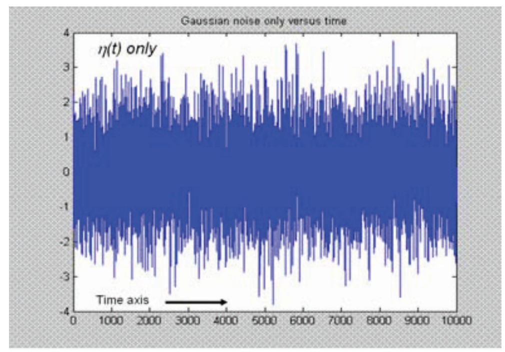

Fig. (11) – The Normal White Gaussian noise Added

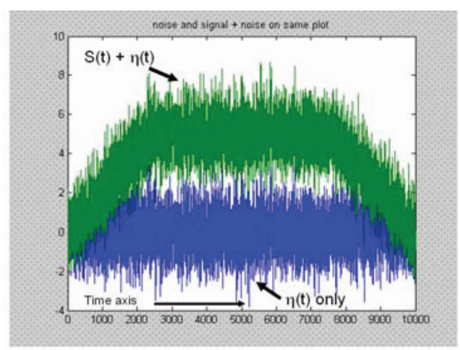

Fig. (12) – Both the Test Noise and Signal + Noise

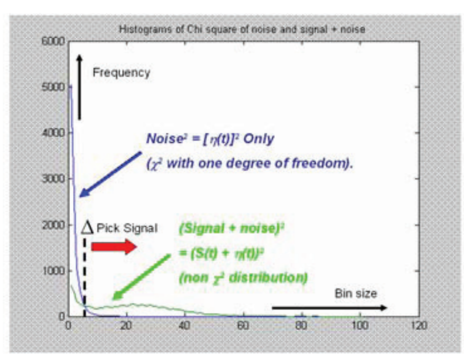

Fig. (13) – The Empirically-derived Histogram Plots

VI. CONCLUSIONS AND DISCUSSION

Comparing the decision rule employing Fig. (9) in lieu of Fig. (7) has some unique advantages. First, only positive quantities are used in the statistical testing in Fig. (9) which is easier to implement. This rule is computationally simple which is required in the numerous testing of the difference images in real time. The same principals of optimality still apply in Fig. (9) as in Fig (7) using the Neyman-Pearson test. The simplicity of the Frobenius norm being calculated from only the matrix coefficients in equation (14) also has its computational benefits.

VII. FUTURE WORK AND RECOMMENDATIONS

The work presented so far has involved simulated data. Data from actual images are currently being evaluated. At issue is the computational cost of providing timely decisions, which is a major consideration of the choice of the final algorithms selected.

REFERENCES

[1] H. Wu, Y. Chen, and T. Fang, "Coupled Bayesian Framework for Dual Energy Image Registration," *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR06), 2006, vol. 2, pp. 2475-2482.

[2] D. W. Repperger, A. R. Pinkus, R. G. Roberts, and R. D. Sorkin, "Image Registration/Fusion Device Based on Two Types of Optimality," AF Invention # AFD 972, November 30, 2007. Patent in progress.

[3] A. Toet, "Image Fusion by a Ratio of Low Pass Pyramid," *Pattern Recognition Letters*, **9** (1989), pp. 245-253.

[4] D. W. Repperger, A. R. Pinkus, R. G. Roberts, and R. D. Sorkin, "Dynamic Image Registration/Fusion Device Based on The Orthogonally Principal, Optimal Estimation and Maximum Likelihood Methods," USAF Invention submitted , May, 2008.

[5] Arun, K. S., Huang, T. S., and Blostein, S. D., "Least-squares fitting of two 3-D point sets," *IEEE Trans. PAMI, 9, 5* (Sept. 1987), pp. 698-700.

[6] B. Kamgar-Paris, "Matching sets of 3D Line Segments with Application to Polygonal Arc Matching," *IEEE Trans. PAMI, (*19) 10, (Oct. 1997), pp. 1090-1099.