Registration of Image Cubes Using Multivariate Mutual Information

Jeffrey P. Kern
Sandia National Laboratories, P.O. Box 5800
Albuquerque, NM 87185-0972

Marios Pattichis
Assistant Professor of Electrical & Computer Engineering, University of New Mexico
Albuquerque, NM 87131-1356

Samuel D. Stearns
Research Professor of Electrical & Computer Engineering, University of New Mexico
Albuquerque, NM 87131-1356

Abstract **- A new method for simultaneously registering a collection of multispectral or hyperspectral images (spectral image cubes) using mutual information is presented. In this paper, we derive a new algorithm based on mutual optimization between pairs of images and extend it to any finite number of images. We present results on the convergence and stability of the new algorithm, showing that the algorithm registers each image to a weighted average of its alignment with every other image. This method has application to other multi-image applications as well.**

I. Introduction

When registering image pairs, a common approach is to designate one image as the reference and apply a spatial shift, rotation, or warping function to the second image in order to bring it into alignment with the reference. Some measure of similarity between the images is used to determine when the optimal alignment has occurred. This measure might be based on point or boundary matching, the cross-correlation between the images, or some other statistical measure. A good survey of these image registration techniques can be found in [1].

Three related phenomena tend to impair the success of these similarity measures when registering multispectral or hyperspectral imagery. The first is a general lack of similarity between widely separated wavelengths. This is particularly pronounced when comparing images collected at wavelengths shorter than 4-5 μm to those at longer wavelengths. This is because image contrasts are primarily due to reflectance of solar energy in the shorter wavelengths (reflectance bands) and primarily due to thermal emissions in the long wave infrared (LWIR, also referred to as thermal bands). A second impairment to registration occurs because thermal emissions are often weak leading to very subtle contrast variations. We will address these phenomena later in the paper.

The third is a phenomenon, known as contrast reversal (or contrast inversion), is caused by differences in the reflectance of various materials as a function of wavelength. To illustrate this, refer to Figure 1. Note that the two images tend to be light or dark in the same regions except for the grassy areas on the left and the lower-right. These areas are dark in the visible image on the left and light (reversed) in the infrared image on the right.

A great deal of literature exists that describes correlation-based registration methods. These methods are based on using the correlation coefficient between corresponding pixels in two images as measure of similarity

$$C(u,v) = \frac{\frac{1}{N}\sum_{i=1}^{N} u_i v_i - \left(\frac{1}{N}\sum_{i=1}^{N} u_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} v_i\right)}{\sqrt{\left(\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} u_i\right)^2\right)\left(\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} u_i\right)^2\right)}} \quad (1)$$

where u_i and v_i represent the values for each pixel in two images U and V, and i is the index of each pixel in the image.

The difficulty with correlation-based similarity measures is that they are degraded by contrast reversals. In extreme

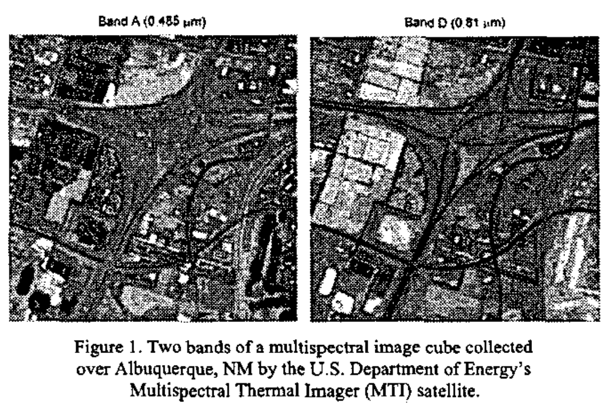

Figure 1. Two bands of a multispectral image cube collected over Albuquerque, NM by the U.S. Department of Energy's Multispectral Thermal Imager (MTI) satellite.

0-7803-8104-1/03/$17.00 ©2003 IEEE

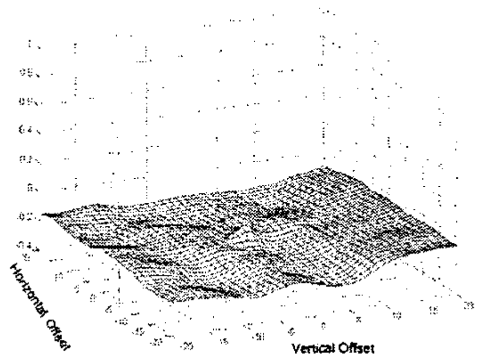

Figure 2. Correlation coefficient of the two images in Figure 1. The right image was shifted horizontally and vertically over the range shown and the correlation coefficient was calculated between corresponding pixels at each shift.

cases, this can degrade to zero as illustrated in Figure 2. Normally, a clearly defined peak signifies the optimal registration. However, in this case, none exists. Attempts to overcome the contrast reversal problem have typically involved pre-processing the images to extract or enhance points or edges [2][3][4][5].

In recent years, similarity measures based on information theory have shown great promise for registering images. An information theoretic measure, that was independently proposed by Collignon et al [6] and Viola [7], is the mutual information between two images U and V

$$I(U;V)=\sum_{u,v} f_{UV}(u,v)\log\frac{f_{UV}(u,v)}{f_U(u)f_V(v)} \tag{2}$$

where $f_U(u)$ and $f_V(v)$ are the marginal probability mass functions of the pixel values for each image, and $f_{UV}(u,v)$ their joint probability mass function. Mutual information is a measure of the dependence or redundancy between two random variables. Because mutual information is not a measure of correlation, it is not degraded by contrast reversals in the same manner as correlation-based approaches. It has been shown that mutual information registration is also a very accurate similarity measure [8] and robust against random noise [9]. An example of the mutual information between the two images in Figure 1 is illustrated in Figure 3. In this case, a prominent peak occurs at the correct registration. In addition to shift operations, registrations involving rotation, changes in scale and warping can also be performed by this method. An excellent survey of mutual information registration can be found in [10].

Mutual information can be written in terms of the joint and marginal entropies of the two images

$$I(U,V)=H(U)+H(V)-H(U,V) \tag{3}$$

where

$$H(U)=-\sum_u f_U(u)\log f_U(u) \tag{4}$$

$$H(V)=-\sum_v f_V(v)\log f_V(v) \tag{5}$$

$$H(U,V)=-\sum_{u,v} f_{UV}(u,v)\log f_{UV}(u,v)\cdot \tag{6}$$

II. Multivariate Mutual Information

The use of mutual information as a similarity measure improves image registration, but can still fail when the images lack sufficient similarity, such is often the case between visible and LWIR imagery. In the case of spectral imaging, a number of images (typically greater than two) are available. Therefore, the opportunity exists to improve registration by choosing image pairs to register in an intelligent fashion rather than arbitrarily. If this is done, some interesting questions arise. Is there a single optimal reference image to register the others to? And if such an image exists, how is it chosen? If two images are equally similar to a third, how should we choose between them. Rather than attempt to choose a single reference image, perhaps a more optimal approach is one where a weighted average between every image pair is used as the similarity measure. Such a method is proposed here.

In information theory, mutual information is the reduction in code length obtained by encoding multiple signals simultaneously rather than separately. When applied to image registration, mutual information is the amount of dependency or redundant information that exists in a collection of images. Just as with two images, it seems reasonable that the optimal alignment should occur at the point where this redundancy is maximized. In this paper, we propose a multivariate extension of the mutual information function to register a collection of spectral images. This approach makes efficient use of the available information in every image in the set (referred to as a *cube* because the related 2D images can be thought of as stacked in a third dimension, i.e. wavelength).

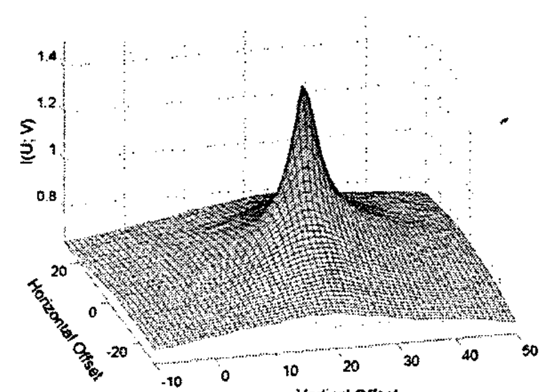

Figure 3. Mutual information surface for the images in Figure 1 as a function of horizontal and vertical offset.

Studholme, et al [11], introduced a useful definition of multivariate mutual information which we will extend to a finite number of images $U_1, U_2, \ldots, U_P$. For an image cube containing P images, the redundant information contained in two or more images is represented as

$$I(U_1;U_2,\ldots;U_P)=\sum_{i=1}^{P} H(U_i)-H(U_1,U_2,\ldots,U_P). \quad (7)$$

Unfortunately, P-variate entropy is difficult to calculate accurately for large P and therefore is prone to error. To solve this problem, it helps to visualize the multivariate mutual information graphically. Figure 4 illustrates the mutual information between three images in a conceptual manner. The areas of the three circles represent their entropy, the combined area can be thought of as the joint entropy, and the gray region in the center represents the mutual information.

It appears from this diagram that the sum of the three pair-wise mutual information values would be a reasonable approximation of three-variable mutual information. In the approximation the center triangular region would be counted three times rather than once. In [12], it was found that this was a reasonably good approximation. We propose that

$$I(U_1;U_2;\ldots;U_P)\approx \sum_{i,j=\{1,\ldots,P;\ i\neq j\}} I(U_i;U_j) \quad (8)$$

is a sufficiently accurate approximation for registering P images and that by maximizing the summation we can register every image in the cube in a manner that makes efficient use of the available information.

Equation (8) requires the simultaneous optimization of $P(P-1)/2$ equations, assuming symmetry. To reduce the number of optimizations, we could envision an approach where two images are registered, then a third, then a fourth, and so on. However, the registration of the first images would not take into account the new information provided by the latter images. A better approach is one that is iterative and allows each registration to consider the total information available. In order to do this, we first consider the previous work in maximizing mutual information by iterative means.

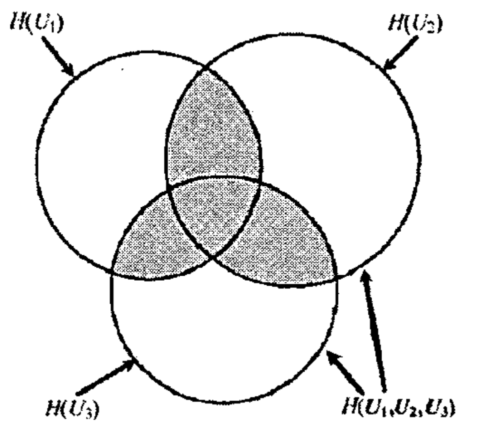

Figure 4. Conceptual diagram of mutual information between three random variables. The information common to two or more images is represented by the gray area.

III. Maximization of Mutual Information

The search space for an optimal registration solution can often be impractically large necessitating the use of optimization methods to find the solution in an efficient manner. For example, the plot in Figure 3 required the computation of 61 x 61 = 3721 points. An alternate solution is to find the peak mutual information using a stochastic gradient optimization algorithm similar to those used in adaptive signal processing and neural networks. Suppose T is a vector representing the horizontal and vertical shift parameters in Figure 3. It can be shown [7] that an optimal solution is found by iteratively updating this vector in the direction of the gradient according to

$$T^{(k+1)}=T^{(k)}+\eta\frac{d}{dT}I(U;V) \quad (9)$$

where η is a small-valued *step-size parameter* or *learning rate*. Further efficiency is achieved by estimating the gradient based on a small number of image samples. Figure 5 illustrates a stochastic gradient search over the same surface as Figure 3. This search required 406 iterations as opposed to the 3721 iterations required for an exhaustive search. Furthermore, the gradient estimate at each iteration used only 100 samples from each image rather than the entire image (256 x 256 = 65536).

The derivative of the mutual information with respect to the transformation (dI/dT) was derived by Viola [7]. Simpler methods can be used, however, that do no require this derivative to be calculated explicitly, such as that by Spall [13].

Equation (9) is analogous to the method of steepest descent used extensively in adaptive signal processing and neural networks, differing only by a sign. It was shown in [12]

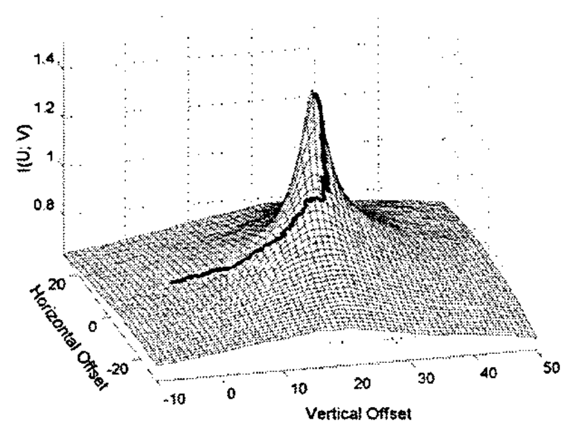

Figure 5. Same surface as Figure 3 overlaid with points searched by a stochastic gradient optimization algorithm.

that the mutual information surface is reasonably quadratic in the vicinity of its peak and that the derivations for stability and convergence developed for adaptive signal processing are applicable to optimizing mutual information.

This has important implications. The mutual information in the vicinity of the peak T^* can be approximated as a Taylor series expansion

$$I(T) \approx I(T^*) + \frac{1}{2}(T - T^*)' H (T - T^*) \quad (10)$$

where H is the Hessian matrix

$$(H)_{ij} = \left. \frac{\partial I}{\partial T_i T_j} \right|_{T=T^*} . \quad (11)$$

Within this region, the eigenvalues λ_i of H describe the quadratic curvature of the peak. Assume for simplicity that the peak is symmetric, H is diagonal and the eigenvalues are all equal to λ. Then $|\lambda|$ is a measure of the "sharpness" of the peak†. Furthermore, the process converges toward the optimal transformation T^* according to

$$T^{(k)} \approx (1 + \eta\lambda)^k T^{(0)} + T^* \quad (12)$$

assuming η is chosen to be less than $2/|\lambda|$. The importance of λ will become clear in a moment. This theory can be extended to the case where λ_i are not equal and H is not necessarily diagonal.

IV. Optimization of the Image Cube

To register an image cube by maximizing the multivariate approximation in (8), equations (9) through (12) can be extended as follows. Suppose image pairs U_i and U_j are randomly selected from the set of P images during each iteration. Then (9) can be restated as

$$T_j^{(k+1)} = T_j^{(k)} + \eta \frac{d}{dT_j} I(U_i; U_j) \quad (13)$$

where

$$i, j = \{1, 2, ..., P; \quad i \neq j\} \quad (14)$$

and T_j is the transformation performed on image U_j. In each iteration, the transformation that determines the registration of image U_j is updated a manner that improves its alignment with image U_i, where U_i can be any one of the other images.

The net effect is that after k iterations each transformation is equal to

$$T_j^{(k)} \approx \prod_{i,j=\{1,...,P;\, i\neq j\}} (1 + \eta\lambda_{ij})^k T^{(0)} + \frac{\sum_{i,j=\{1,...,P;\, i\neq j\}} \lambda_{ij} T_{ij}^*}{\sum_{,j=\{1,...,P;\, i\neq j\}} \lambda_{ij}} . \quad (15)$$

For $k \to \infty$ and small learning rates, i.e. $\eta < 2/\max|\lambda_{ij}|$, the product term approaches zero and the transformation converges to

$$T_j^* \approx \frac{\sum_{i,j=\{1,...,P;\, i\neq j\}} \lambda_{ij} T_{ij}^*}{\sum_{i,j=\{1,...,P;\, i\neq j\}} \lambda_{ij}} \quad (16)$$

This result is very significant because the final solution is a weighted average of the pair-wise solutions. Furthermore. the weights are equal to the sharpness of the mutual information peaks for the pair-wise solutions. In our experiments, we saw a tendency for the largest mutual information peaks also to be the sharpest. To the extent that this trend holds, the implication of (16) is that the algorithm will align each image with those that it is most similar with. Furthermore, it is not necessary to designate any of the images a the reference image. Rather, the reference is mutually agreed upon.

V. Experiments

For multispectral imagery collected using a common aperture, it is common for the bands to be misregistered by a few pixels. This algorithm was performed 14-band image cube for the purpose of determining the correct band alignment. One band of the cube is illustrated in Figure 6. For this experiment, the registration was limited to horizontal and vertical shifts and the gradient was calculated using a random sample of one hundred pixels chosen for each iteration.

Figure 7 illustrates the convergence of the fourteen transformations T_j. Each strand corresponds to the transient behavior of one transformation as it converges. For ease of

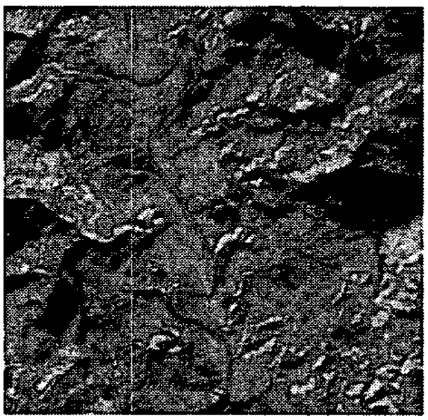

Figure 6. One band (0.485 μm) of a 14-band image collected by the MTI satellite over Burning Hills, Utah.

† Note that since the mutual information peak is concave, the second derivatives and therefore λ are negative.

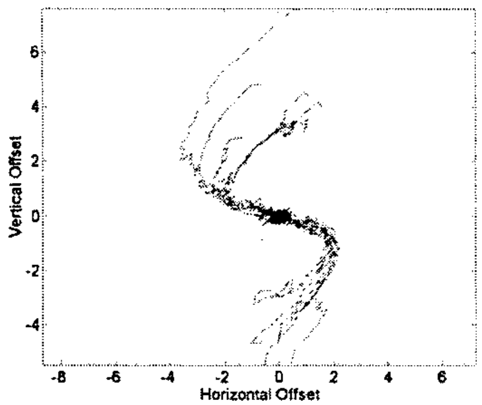

Figure 7. Registration of 14 bands of a multispectral image cube collected by the MTI satellite. Each line represents the convergence of one image toward the mutual optimum.

viewing, the final value for each sequence was subtracted from every other value in that sequence so that each strand would converge toward zero rather than away from zero.

Ten independent trials were performed. Samples were chosen using a different random seed for each trail. The ten trials agreed to within a standard deviation of 0.25 pixels. For this particular image cube, it was possible to register every band to a single reference band in a pair-wise fashion. The multi-image registration also agreed with the pair-wise registration to within 0.25 pixels.

A second test was performed using the three spectral bands illustrated in Figure 8. In this case, the visible and LWIR bands failed to register as a pair. Introducing a mid-wave infrared (MWIR) band resulted in a successful three-image registration. Once again, this experiment was repeated ten times. The results were consistent to within 0.7 pixels.

VI. Conclusions and Future Work

Multivariate mutual information shows promise as a technique for improving automated registration of spectral imagery. This technique is not only applicable to spectral imagery, but may be useful for registering other image sequences where it is difficult to identify a single "best" reference image, such as images collected under different lighting conditions or images collected by multiple instruments. Though these experiments were limited to horizontal and vertical shift transformations, it is possible to extend the technique to higher order functions. Piecewise linear transformations were found to work well for multispectral applications.

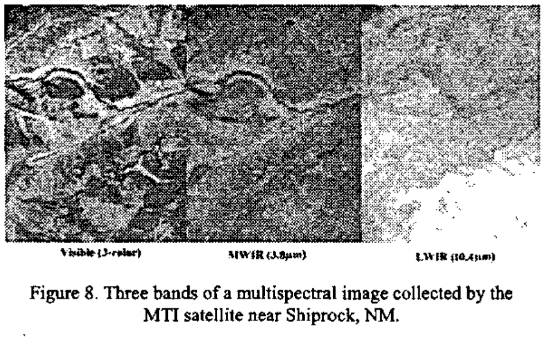

Figure 8. Three bands of a multispectral image collected by the MTI satellite near Shiprock, NM.

An extensive evaluation, involving 90 image cubes (1260 total bands), is currently underway. Early results indicate a significant improvement over bivariate techniques.

VII. Acknowledgments

We would like to thank Sandia National Laboratories and the United States Department of Energy for providing the imagery used in this research. *Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.*

[1] L. G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24(1), pp. 325-376, 1992.

[2] J. Ton and A. K. Jain, "Registering Landsat Images by Point Matching," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27(5), pp. 642-51, 1989.

[3] H. Li, B. S. Manjunath, and S. K. Mitra, "A Contour Based Approach to Multisensor Image Registration," *IEEE Transactions on Image Processing*, vol. 4, pp. 320-334, 1995.

[4] A. Roche, X. Pennec, G. Malandain, N. Ayache, "Rigid Registration of 3-D Ultrasound With MR Images: A New Approach Combining Intensity and Gradient Information," *IEEE Transactions on Medical Imaging*, vol. 20(10), pp. 1038-49, 2001.

[5] H. S. Stone and R. Wolpov, "Blind Cross-Spectral Image Registration Using Prefiltering and Fourier-Based Translation Detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40(3), pp. 637-650.

[6] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated Multi-Modality Image Registration Based on Information Theory," *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, eds., pp. 263–274, Kluwer Academic Press, 1995.

[7] P. A. Viola, *Alignment by Maximization of Mutual Information*, doctoral dissertation, Massachusetts Institute of Technology, 1995.

[8] P. Thévenaz and M. Unser, "Optimization of Mutual Information for Multiresolution Image Registration," *IEEE Transactions On Image Processing*, vol. 9(12), pp. 2083-2099, 2000

[9] K. Johnson, A. Cole-Rhodes, I. Zavorin, and J. Le Moigne, "Mutual Information as a Similarity Measure for Remote Sensing Image Registration," *Proceedings of the SPIE*, vol. 4383, pp. 51-61, 2001.

[10] J. P. W. Pluim, J. B. Maintz and M. A. Viergever, "Mutual-Information-Based Registration of Medical Images: A Survey," *IEEE Transactions on Medical Imaging*, vol. 22(8), 2003.

[11] C. Studholme, D. L. G. Hill and D. J. Hawkes, "Incorporating Connected Region Labelling into Automated Image Registration Using Mutual Information," *Mathematical Methods in Biomedical Image Analysis*, A. A. Amini, F. L. Bookstein, and D. C. Wilson, eds. Los Alamitos, CA, *IEEE Computer Society Press*, pp. 23-31, 1996.

[12] J. P. Kern, *Spectral Image Registration Using Multivariate Mutual Information*, doctoral dissertation, University of New Mexico, 2003.

[13] J. C. Spall, "Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation," *IEEE Transactions on Automatic Control*, vol. 37, pp. 332-341, 1992.