Using nonlinear wavelet compression to enhance image registration

Ronald A. DeVore[1], Wei Shao[2], John F. Pierce[3],
Emre Kaymaz[4],Bao T. Lerner[5], William J. Campbell[6],

[1]Dept. of Mathematics,
University of South Carolina,
Colombia, SC 29208

[2]Advanced Scientific Research,
Western Atlas Logging Services,
Houston, TX 77042,

[3]Dept. of Mathematics,
U. S. Naval Academy,
Annapolis, MD 21402

[4,5]KT-Tech, Inc.
9801 Greenbelt Road, Suite 314,
Lanham, MD 20706

[6]Information Science and Technologies Branch, Code 935,
NASA Goddard Space Flight Center,
Greenbelt, MD 20771

ABSTRACT

We present a method for automatically registering images based on nonlinear compression. The method involves three steps: (i) analysis of the complexity of the images, (ii) high level compression for extracting control points in the images, (iii) registration of the images by matching control points. The first step analyzes the complexity of the given images. It numerically computes from any image a complexity index which determines the efficiency at which the image can be compressed. This index is used in the second step of the algorithm to select coefficients in the wavelet representation of the image to produce a highly compressed image. The wavelet coefficients of the highly compressed image are then transformed to pixel values. Only a few pixel values (called control points) are nontrivial. The third stage of the algorithm uses a point alignment techniques to identify matching control points and to erect the registering transformations. The algorithm is tested on two quite different scenes: a portrait, representing an uncomplicated scene, and a Landsat TM image of the Pacific Northwest. In both cases, images are tested which differ by a rotation and which differ by a rigid transformation. The algorithm allows a choice of different metrics in which to do the compression and selection of control points.

KEYWORDS: nonlinear wavelet compression, registration, complexity index

1 INTRODUCTION

The purpose of this paper is to show how high level wavelet based image compression can be effectively utilized for image registration. As noted in the abstract, our method involves three steps: (i) an analysis of the complexity of the images, (ii) high level wavelet based compression to extract a few control points in each of the images to be registered, (iii) registration of the images through the alignment of the control points.

The first step of the algorithm determines the complexity of any image. The "complexity index" is based on that rate at which the compressed image approximates the original image. This steps allows a choice of error metrics which can be used to measure the error between the compressed and original images. The purpose of the complexity index is to indicate how many control points will need to be retained for the purpose of image registration. Highly complex images will need many more control points vis a vis simpler images.

The second stage of the algorithm uses high level wavelet compression to extract a relatively small number of control points in each of the images to be registered. This is accomplished by choosing an error metric suitable for registration and then utilizing nonlinear wavelet based compression relative to that metric. The result is that each image is compressed to images which have very few nonzero pixel values. These pixels are the control points which are utilized in the registration step that follows. The third stage of the algorithm uses a point alignment technique to identify matching control points and to erect the registering transformations.

We tested the method empirically to see if (1) the computationally determined index would group images into classes over which we could measure the effectiveness of compression algorithms for the purpose of registration; (2) whether for a given class of images it would determine from the index an optimal level of compression for the purpose of registration; and if so, (3) whether the thresholding strategy built on this value would optimally balance the level of compression with the need to preserve those elements in the image necessary for registration.

We applied the method to two quite different scenes: a portrait, representing a composed scene, and a Landsat TM image of the Pacific Northwest. In both cases we registered images which differed by a rotation and by a rigid transformation. We used biorthogonal wavelet bases for our analysis and compression.

In Section 2 we introduce the analytical tools on which we base the complexity analysis, decide the level of compression, and produce a nonlinear compression of the input images for the purpose of registration. In Section 3 we describe the statistical point matching approach we use to register the compressed images. In Section 4 we present the results of empirical tests done on two distinct classes of images. In Section 5 we present our results. In Section 6 we present our conclusions.

2 AN INDEX FOR COMPLEXITY

Our algorithm for image registration has at its core compression of the images to be registered. We would like to utilize as high a level of compression as possible but still retain some essential features of the image. We are confronted, therefore, with the question of how to autonomously decide on the compression level of the image. To answer this question, we introduce a quantity which we call the "complexity index" of the image. Its purpose is to classify images according to their complexity. For simple images, we can then use very high compression; for more complex images, we shall have to retain more information and use lower levels of compression.

Our complexity index is based on the work of R. DeVore, B. Jawerth, B. Lucier, and V Popov, and X. Yu.[1-4] These authors develop a mathematical framework by which one can objectively compare the effectiveness of compression. Their framework is based upon concepts from approximation and function space theory. We outline the elements of this subject needed to describe the complexity index.

We can regard an image as a function I defined over a two-dimensional continuum, which we shall take to be the unit square. We regard the set of pixel values associated with the image as a sampling of the function obtained by averaging it over small squares, the dimensions of which are determined by the pixel size. Thus a digitized image is viewed as a set of pixel values obtained by sampling a continuous image I. We can use the pixel values to form a wavelet representation of I. For this, let ϕ be a univariate scaling function and ψ its associated wavelet. From the pixel values of I, we form the discrete wavelet representation I_0 of I given by

$$I \approx I_0 := \sum_j c_j \phi_{j,m}(x, y), \tag{1}$$

where $j \in \mathbb{Z}^2$ and the $\phi_{j,m}$ are the usual shifted-dilates of the bivariate scaling function $\phi(x)\phi(y)$. Most of coefficients in the expansion are computed by a fast wavelet transform on the pixel values. Coefficients of elements whose support intersect the boundary of the image are obtained in a slightly more sophisticated way.[1]

We transform to the wavelet basis by Fast Wavelet Transform,

$$\sum_j c_j \phi_{j,m}(x, y) = \sum_j c_{j,0}\phi_{j,0} + \sum_{k=0}^{m-1} \sum_j \sum_{\psi \in \Psi} d_{j,k,\psi}\psi_{j,k}, \tag{2}$$

where $\Psi := \{\phi(x)\psi(y), \psi(x)\phi(y), \psi(x)\psi(y)\}$ is the set of bivariate wavelet functions generated by ϕ. The coefficients $d_{j,k,\psi}$ of the functions $\phi(x)\psi(y), \psi(x)\phi(y)$, emphasize features in the vertical and horizontal directions, while the coefficients of the functions $\psi(x)\psi(y)$ emphasize point singularities.

Wavelet methods of compression can be viewed as replacing the coefficients $c_{j,0}$, $d_{j,k}$ in (2) by new coefficients $c'_{j,0}$, $d'_{j,k}$. Most of these new coefficients are zero (thresholding) and the others are an approximation using fewer bits (quantizing).[1] The starting point for determining the thresholding and quantization strategies is to choose an error metric for measuring the difference between two functions which is commensurate with the image processing task. This metric is then applied to compute the "error" between the exact and compressed image.

Once a compression strategy is in place, we can compute a complexity index for each image. We do this by computing a compression error for different levels of compression and then examining how fast the error tends to zero as the compression level is lessened. For simple images, the error will tend to zero relatively fast when compared to complex images.

To make this more precise, let $E_N(I)$ denote the approximation error between I_0 and the compressed image I_N given by the wavelet compression algorithm. Here N represents the number of coefficients (or bits) needed to represent the compressed image. We then define the complexity index of I as the largest number $\alpha > 0$ such that

$$E_N(I) < CN^{-\alpha/2}. \tag{3}$$

If we graph this metric measure of the error from the compression against the reciprocal of the number N on a log-log scale, equation (3) asserts that an estimate of the slope of the graph over a range of N will give a lower bound for α. We use this number is our empirical estimate of the complexity index. The mathematical theory says that the choice of wavelet basis does not affect the value of the index α computed.[1] Rather, it affects the constant C in the bound for the error.

The usual candidates for error metric in image processing are the L_p metrics: most notably L_1, L_2, and L_∞. For determining the complexity index of an image, we have used the L_2 metric since it is very easy to compute and

Figure 1: Hannah with Hat

gave a good classification of the complexity of the image in practice. However, for the actual compression (which is the preprocessor to the registration), we would like the error metric to emphasize edges and point singularities, since these are precisely the elements of the image which we would like to retain. We examined several metrics arising from L_p function space norms to determine which one was most sensitive in detecting a point-by-point deviation of two edges. We found the metric based upon the $Y := W^1_\infty$ norm as our best candidate. Basically, it measures the maximum of the differences between the first derivatives of an image and its compression at each point in the domain. So, small error in this metric means gradients (points, corner, and edges) in the image and its compression line up point-by-point.

Once a compression metric has been chosen, the mathematical theory can be used to determine an optimal strategy for thresholding and quantization; moreover, the thresholding strategy for doing compression in Y is exactly the same as that for doing compression in L_2.[1] This further justifies the use of the L_2 metric for indexing the complexity of an image.

Our empirical efforts consisted of: (1) numerically computing the complexity indices for images; (2) using the complexity index together with the mathematical theory to determine a level of compression for each image for the purposes of registration; (3) carrying out the compression and analyzing the accuracy of registration, based upon these predictions.

The two images we used for our experiments were: a composed scene, called "Hannah with Hat", shown as Figure 1, and a Landsat/Thematic Mapper image of the Pacific Northwest, shown as Figure 2. We chose these images, because, at least to the human eye, the first image appears uncomplicated, while the second appears to be much more complicated. We wished to determine whether or not the computation of the complexity index and its analysis would characterize the two images as representatives of two distinctly different classes of images.

Theorem 6 in the principal reference specifies how we can use this index to group images into classes which characterize the effectiveness of compression for the intended purpose.[1] We demonstrate this classification in the results Section 3. Theorem 5 and Algorithm 2 in the principal reference specify how the complexity index determines a thresholding strategy which can be used to produce a compression algorithm which will converge at an optimal rate.[1] So the algorithm is a *constructive* procedure. This strategy was set forth in general terms. We

Figure 2: Thematic Mapper, Pacific NW

adapted it here to the purpose of maintaining those elements of the image valuable for registration.

3 REGISTERING THE COMPRESSED IMAGES

In this section, we shall describe our image registration algorithm. Suppose, we are given two images I and $\tilde{I}$ which represent different views of the same scene which we wish to register. Our algorithm first determines the complexity index of these two images. Using this complexity index, we then apply high level compression in the metric of Y to the two images to produce data which serve as our control points. We then register the images using these control points.

The compression step is nonlinear, in that it retains wavelet coefficients at various levels of resolution dictated by the thresholding strategy. This compression stands in contrast to the linear compression, which retains all the wavelet coefficients up to a specified level of resolution. When the complexity index α of the image was large (eg. about 0.5 in the L_2 norm), we attempted registration at very high levels of compression, using few pixels. When the computed index of the image was low (eg. about 0.2 in the L_2 norm), we attempted registration at moderately high levels of compression, again retaining relatively few pixels.

How we registered the compressed images differed fundamentally from approaches using linear wavelet representations.[5] Since the nonlinear compression uses coefficients from all levels of the decomposition, a successive approximation based upon increasing the level of decomposition would not apply. Because we have very few pixels in the compressed images, relatively speaking, a measure based upon correlation likewise would be inappropriate.

Rather, to register the two compressed images which differ by a a rigid transformation, we used a point alignment procedure. Specifically, when the images differed by a pure rotation, we sampled each compressed image using a family of annuli of fixed thickness δR concentric about the point of rotation. We then chose an increment of angle $\delta\theta$ and partitioned each annulus to create a vector $\nu = \{\nu_j\}$ whose index j determined a relative position on the circle, and whose value at that index , ν_j was set to 1 or 0, depending upon whether or not the annulus intersected a pixel in the image at that location.

The registration proceeded in the following way. For corresponding annuli from the reference and the input compressed image, we recorded the number of "hits", where $\nu_j = 1$ at the same location j on both annuli. On a histogram along the horizontal axis at the rotation angle value 0 we plotted along the vertical axis the total number of hits computed this way over the family of annuli.

Then we "rotated" the compressed reference image through an angle $\delta\theta$ in this sense: we transformed the vectors associated with each annulus on the reference image by the transformation $\nu_j \to \nu_{j-1}$. We recorded the number of hits between this family of vectors associated with the "rotated", compressed reference image and the corresponding members of the vectors in the compressed input image. On the histogram along the horizontal axis at the rotation angle value $\delta\theta$ we plotted along the vertical axis the total number of hits. We proceeded to build the histogram by "advancing" the rotation of the compressed reference image by iterating the transformation of its family of vectors stated above. For each advance, we plotted the total number of hits over all annuli at the corresponding value for the rotation angle.

We would achieve accurate, reliable registration by rotation of the compressed images by this method if the histogram produced peaks sharply and in a unimodal fashion. We could then compare the value for the rotation of these compressed images to the actual value of the angle of rotation between the two original images.

An analogous procedure produces an algorithm for registering images which differ by a pure one-dimensional translation. Here, the partitions of the images into strips with a prescribed widths are made perpendicular to the axis of translation of interest. Hits are recorded and a histogram formed in the manner which parallels that described for the case of rotation.

To register to images which differ by a rigid transformation, specifically a rotation followed by a translation, we proceed in this way.[8] First, we "mod out" the translation by creating "images" (datasets) of differences of reference control points and image control points, respectively. These "reference" and "input" control point sets differ by a pure rotation. We then register the rotation, rotate the original (compressed) reference image, and consider a new problem of registering it against the (compressed) input image by a pure translation, possibly two-dimensional.

To register the two-dimensional translation, we "mod out" one of the translational directions by creating "images" (datasets) of differences of the reference control points and image control points in one coordinate direction. These new "reference" and "input" control point sets differ by a one-dimensional translation. We register this translation, as outlined above, translate the original (compressed) reference image, then consider the final problem of registering it against the (compressed) input image by a pure one-dimensional translation.

4 THE EXPERIMENTS

We examined the problem of registrating scenes from two distinct classes. We took as our reference images the composed scene Hannah with Hat, and the involved scene, a Thematic Mapper image of a segment of the Pacific Northwest region of the U. S. Both were 512×512 images.

We computed the complexity index for both images by the following procedure. We chose a Haar wavelet basis with which to represent the two images. We produced compressed images for a number of levels of compression. We then produced the log-log graph described in Section 2 and estimated the index. We show the graph for Hannah with Hat, and the estimates for the indices for both images in Section 5.

We performed two registration experiments on these scenes. In the first, we took the input images to differ from the reference images by a rotation of fifteen degrees about the center pixel. In the second, we took the input images to differ from the input images by the fifteen degree rotation, followed by horizontal and vertical translations of 40 and 30 pixels, respectively.

Based upon the complexity index computed for each reference image, we chose a level of compression at which to attempt registration. We compressed each image using the nonlinear strategy described above, and retained a sampling pixels which had the highest grey-level values in each reconstructed, compressed image. We retained approximately 50 pixels for the case of the composed scene, Hannah with Hat, and approximately 100 pixels in the case of the involved scene, the TM image of the Pacific Northwest. These choices were approximate; we simply wished to produce a high level of compression and to retain relatively few, but a statistically significant number of pixels to produce a registration. We can set forth more precise criteria for the number and selection of control points, and the mechanisms to automatically achieve them.[8]

We then produced the histograms for the angle of rotation and for the translations which would register the compressed images in the two experiments. We analyzed the resulting histograms to determine if each produced a reliable prediction for the registration for the uncompressed images, and whether the prediction varied as we varied the decrements used in producing the histograms.

5 THE RESULTS

To produce the complexity indices for the two scenes Hannah with Hat and the TM image of the Pacific Northwest, we produced plots of the natural log of the L_2 differcnce between the image and its compression versus the natural log of the reciprocal of the number of coefficients in the compressed image. Figure 3 shows the graph for Hannah with Hat. The slope of the best linear approximation of the graph over a midrange of values for the number of coefficients gives a lower bound for the index of the image. These estimates are 0.5 for Hannah with Hat, and 0.2 for the Thematic Mapper image.

In the description of the smoothness of images in terms of this index the value of the index for Hannah with Hat is high.[1] In terms of picture quality, it indicates that the image could be compressed using wavelet to a high degree without appreciable loss in perceived picture quality. In contrast, the relatively low value for the index associated with the TM image indicates that the scene can be compressed using wavelets, but that the visual quality of the image will degrade more rapidly as the level of compression increases. Consequently, the index predicts that we can retain visual quality of the TM image by compressing the image at low levels.

Therefore, we compressed the reference and input images for Hannah with Hat to a level of 1000:1, using the nonlinear strategy described above, and retained the hottest 50 pixels. Figures 4 and 5 display the reconstructed, compressed reference and input images. We compressed the reference and input images for the TM scene to a level of 200:1, using the nonlinear strategy described above, and retained the hottest 100 pixels. Figures 6 and 7 show these reconstructed, compressed reference and input images.

For the first registration experiment, we applied the control point alignment algorithm described in Section 2 to the compressed reference and input images of Hannah with Hat. We obtained the histogram displayed in Figure 8. The decrement in the radius used for building the annuli was 5. The estimated value for the rotation angle was 15 degrees. When we applied the point matching algorithm to register the compressed images for the TM scenes, we obtained the histogram shown in Figure 9. The estimated value for the angle of rotation was 15 degrees.

For the composed scene, Hannah with Hat, and for the involved scene, the TM image, we subsequently applied the point matching algorithm to the reference and input images differing by the rigid transformation described in Section 3. We present the histograms for the angle, and for horizontal, and vertical translations in Figure 10. Estimated values for the rotation, horizontal, and vertical translation parameters were 15 degrees, 40 pixels, and 28 pixels. Details and discussions of these developments are available.[8]

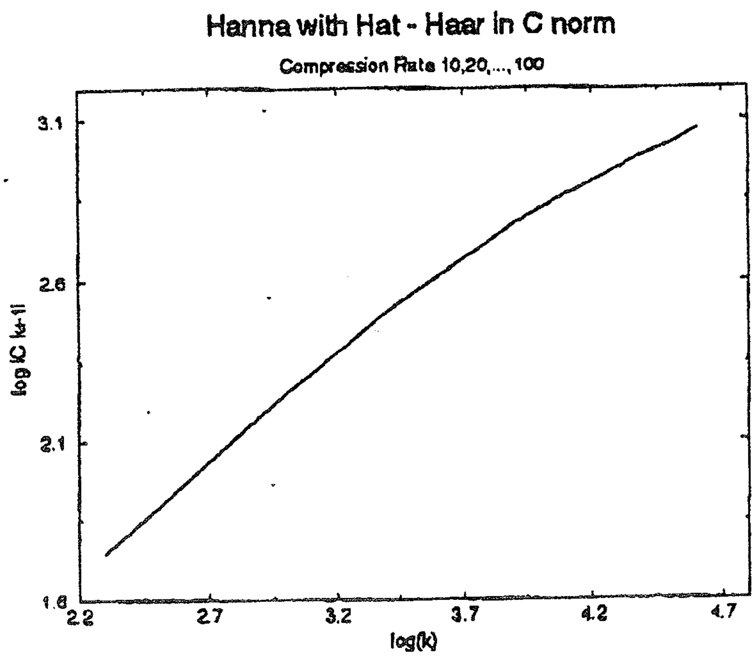

Figure 3: Complexity Index, Hannah with Hat

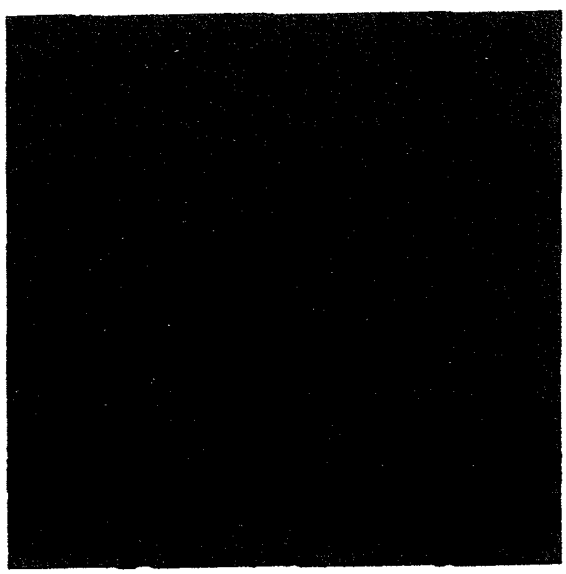

Figure 4: Hannah, 1000:1, 50 Pixels

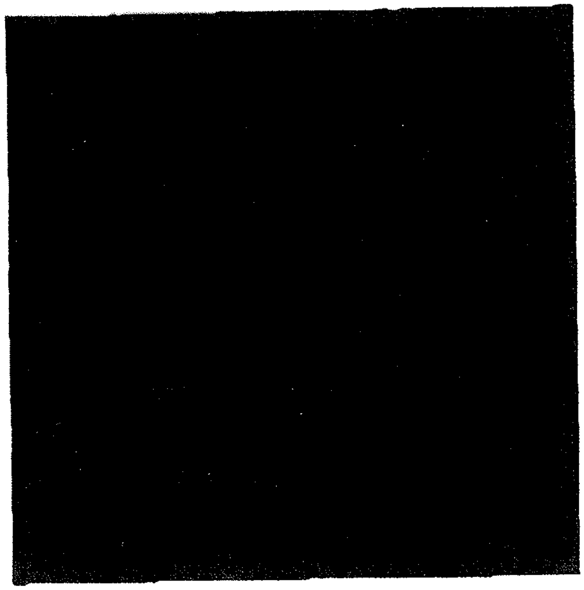

Figure 5: Hannah, Rotate 15, 1000:1, 50 Pixels

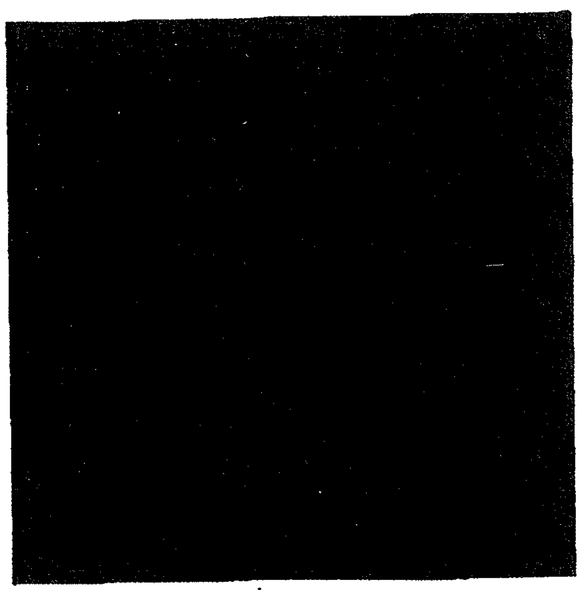

Figure 6: TM Pacific NW, 200:1, 100 Pixels

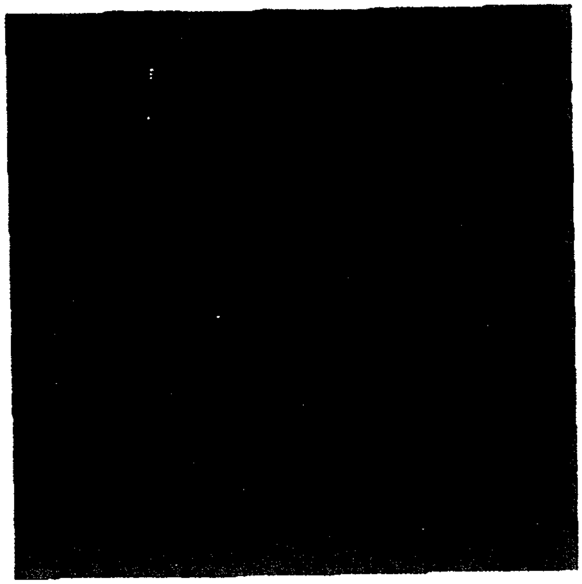

Figure 7: TM Pacific NW, Rotate 15, 200:1, 100 Pixels

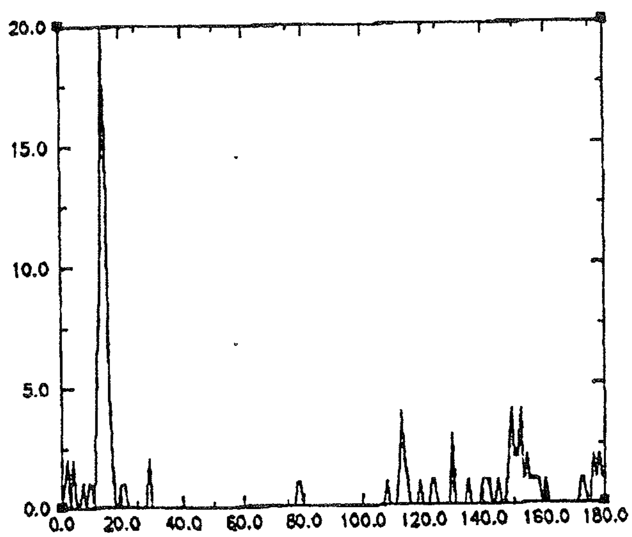

Figure 8: Angle Histogram, Hannah, Rotate 15, $\delta R = 5$

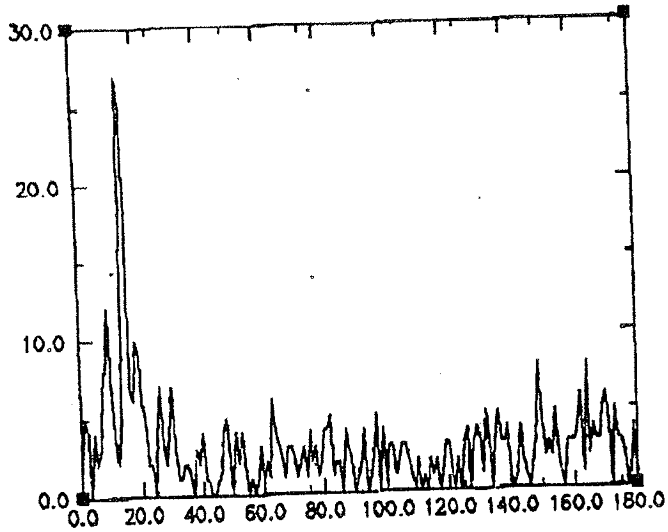

Figure 9: Angle Histogram TM Pacific, Rotate 15, $\delta R = 5$

6 CONCLUSIONS

Figure 8 demonstrates that we can register a composed scene with rotations of it accurately and with high reliability using very high compression and few pixels. This result confirms the prediction made from the computed value of the complexity index for the composed scene. Moreover, experiments using annuli of various sizes demonstrate that the rotation conclusion is insensitive to the particular choice of dimensions for the annuli used in the point matching algorithm.

Figure 9 demonstrates that even for more involved scenes, such as the TM image, we can obtain accurate registration by rotation using the moderately compressed images, and that the registration will be reliable, though not as reliable as with the composed scene. These results again confirm the prediction made from the computed value of the complexity index for the TM image. Figure 10 demonstrates that the method continues to function accurately when we press to register involved scenes which differ by a rigid body transformation.

Finally, when we compare the graphs of correlation versus angle for the registration of the composed scene and the TM images which were obtained using another technique to the histograms in Figures 8 and 9, the comparisons demonstrate that the nonlinear compression and registration algorithm provides a far more accurate and reliable registration.[6]

In summary, our results demonstrate that it is possible to build a method for automatically registering images through rigid body transformation by using compressed versions of them, and which takes into account the complexity of the image. The method consists of an image complexity analysis stage, a nonlinear wavelet compression stage, and a registration stage.

The method is directed, because all computations about the complexity of an image, and all decisions about the level and strategy of compression for an image are directed by the *purpose* for the compression, in this case, registration. The purpose manifests itself in the choice of metrics used in building the complexity index and in thresholding and quantizing the wavelet coefficients.

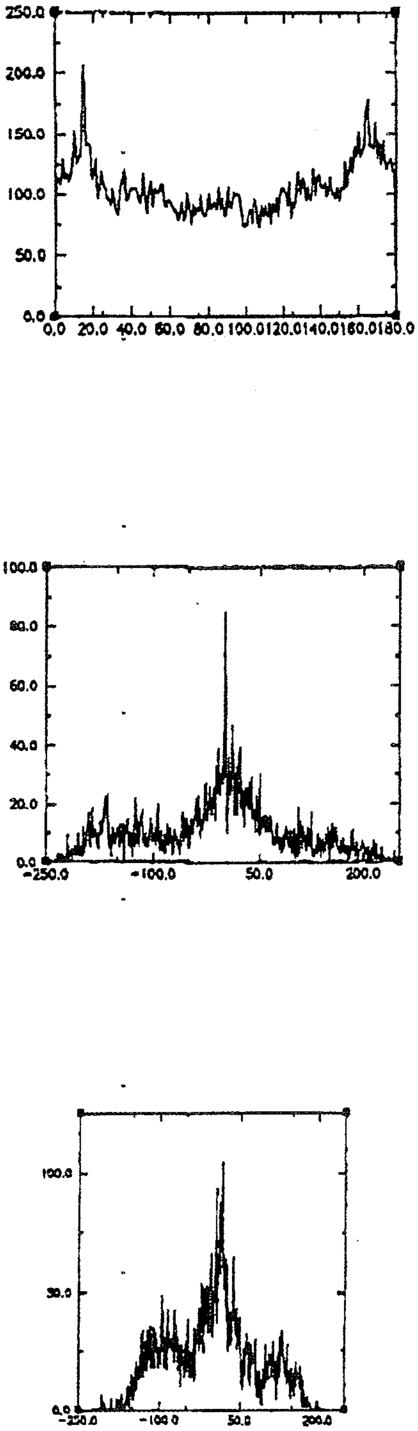

Figure 10: θ, x, and y Histograms, TM Pacific NW

The method is stable, because the complexity index doesn't simply characterize a single image; rather, it *classifies* images by grouping them into collections (Besov spaces).[1,2] Decisions about the level and strategy of compression for a given purpose hold over an entire *class* of images. Thus, if an second image is a slightly blurred version of the original image, though strictly speaking a different image, it generally would continue to belong to the same complexity class. So, the level and strategy for compression for the purpose of registration would vary only slightly from that associated with the original image. We refer to this phenomenon as the "continuous selection" property.[3]

The method is robust, again because the index classifies the images. The level and the way an image will be compressed is determined by *one* algorithm, the instantiation of which will *vary* depending upon in which class the image resides. If one image differs significantly in nature from another, the one procedure for computing its index produces two different values, associating the two images with different Besov spaces. Because their classes differ, the one procedure which determines the level and strategy of compression will produce different conclusions for each image.

7 REFERENCES

[1] DeVore R., Jawerth, B., Lucier, B., "Image compression through wavelet transformation coding," *IEEE Transactions on Information Theory*, 38, 1992, 719-746.

[2] DeVore, R., Popov, V., "Interpolation of Besov spaces," *Trans. Amer. Math. Soc.*, 305, 1988, 397-414.

[3] DeVore, R., and Yu, X., "Nonlinear n-widths in Besov spaces," *Approximation Theory VI: 1*, C. K. Chui, L. L. Schumaker and J. D. Ward, editors. New York: Academic Press 1989, 203-206.

[4] Jawerth, B., Hilton, M., and Huntsberger, J., "Local enhancement of compressed images," *J. Math. Imag. Vision*, 3, 1993, 39-49.

[5] LeMoigne, J., "Parallel registration of multi-sensor remotely sensed imagery and wavelet coefficients," *SPIE Wavelet Applications Conference*, Orlando, April 1994.

[6] Kaymaz, E. Lerner, B-T, Campbell, W, Le Moigne, J., Pierce, J., "Registration of Satellite Imagery utilizing the low-low components of the wavelet transform," *25th ARPR Workshop Proceedings*, Vol. 2962, October 1996.

[7] Mallat, S., "Multiresolution approximation and wavelets," *Trans. Amer. Math. Soc*, 315, 1989, 69-88.

[8] Wei-Shao, "Image Processing and Neural Networks," PhD Thesis, Mathematics Dept. Univ. of South Carolina, 1996.