

Fast multi-modality image matching

Anthony Apicella, J. Shane Kippenhan and Joachim H. Nagel

Dept. of Biomedical Engineering
University of Miami
P.O. Box 248294, Coral Gables, Florida 33124

ABSTRACT

Automated image matching has important applications, not only in the fields of machine vision and general pattern recognition, but also in modern diagnostic and therapeutic medical imaging. Image matching, including the recognition of objects within images as well as the combination of images that represent the same object or process using different descriptive parameters, is particularly important when complementary physiological and anatomical images, obtained with different imaging modalities, are to be combined. Correlation analysis offers a powerful technique for the computation of translational, rotational and scaling differences between the image data sets, and for the detection of objects or patterns within an image. Current correlation-based approaches do not efficiently deal with the coupling of the registration variables, and thus yield iterative and computationally-expensive algorithms. A new approach is presented which improves on previous solutions. In this new approach, the registration variables are de-coupled, resulting in a much less computationally expensive algorithm. The performance of the new technique is demonstrated in the matching of MRI and PET scans, and in an application of pattern recognition in linear accelerator images.

1. INTRODUCTION

The problem of matching images or data sets that contain similar or complementary information arises in a wide variety of applications in signal and image processing. Current interest is focused on the fields of pattern recognition, computer vision and medical imaging, where image matching is relevant to both diagnostic and therapeutic applications. Current medical imaging technology enables precise visualization of internal metabolic function as well as anatomy and/or pathology. Nuclear medicine imaging, such as Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT), provides functional information, whereas Magnetic Resonance Imaging (MRI), ultrasound and x-ray imaging, including Computerized Axial Tomography (CAT) provide anatomical information. The problem with available functional images is their low resolution and lack of anatomical landmarks, preventing the physician from identifying the actual position of displayed source areas within the human body. If the functional and anatomical images are correctly registered, scaled and combined, then the anatomical information enables precise location of the metabolic processes in the corresponding functional images. Efforts to detect tumors, and clinical diagnoses in general, that are based on the two complementary images will be more effective than those based on either of the two types alone. The problem of overlaying or comparing functional and anatomical images is not a simple one, however. The two modalities have different display parameters; anatomical boundaries may not exactly correspond to regions of metabolic activity. Variations in scale factors, viewing angles, and the fact that images may be acquired at different times with different equipment lend additional complications.

Object recognition problems arise in image-controlled radiotherapeutic cancer treatment¹. High-energy radiation, generated by a linear accelerator, needs to be directed such that it strikes the tumor, but not the surrounding healthy tissue. Since projection images obtained from the penetrating radiation are of extremely poor quality and contain almost no anatomical detail (the image frame should not exceed the area of treatment), control of accelerator alignment represents a serious problem. Image matching offers a solution by detecting the pattern of the accelerator image in a normal x-ray picture taken prior to treatment. This enables the physician to locate the site of radiation incidence with reference to the high-quality x-ray image.

Due to the absence of distinctive biological landmarks in many medical images, the simple observation of temporal developments in anatomy or function can be difficult. This problem, important in medical research and control of therapy, requires the detection of three-dimensional translational and rotational shifts and scale factor adjustments between successive images. A volume representation needs to be constructed from a set of slice images, and the volume images must then be transformed into a common frame of reference.

Image matching has proven to be a difficult task for a human operator, and automated registration systems would certainly improve the reliability and precision of many diagnostic evaluations. Previously-proposed techniques for automated image matching have suffered from problems of reliability, precision and computational expense. The method to be presented here is less computationally expensive than previous methods, yet allows for precise, objective, automated matching of images collected at different times or with different modalities. Additionally, the algorithm tolerates changes of imaging parameters, a very important feature for the combination of PET and MRI images, in which each type of image is based on parameters which can vary widely and thus produce images of various qualities and characteristics. The appearance of a PET image, for instance, will vary with the type of radiotracer administered to the patient, the length of data acquisition, and the metabolic condition of the patient (which, for brain images, depends on the patient's mental state), while MRI image appearance will vary according to the pulse sequence used during acquisition. Figure 1 shows three typical PET images (in the top row) and six typical MRI images obtained with various acquisition parameters.

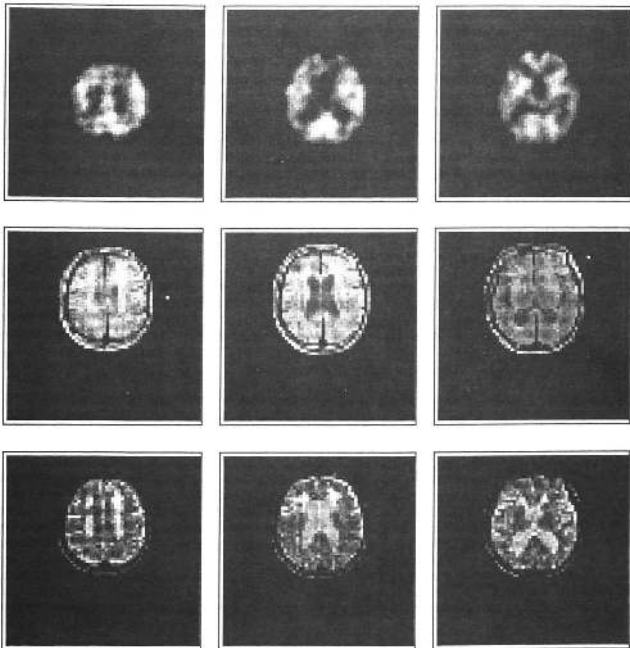


Figure 1 : Three typical PET images (top row) and six typical MRI images

2.BACKGROUND

The widely-used "manual" approach to combining multi-modality information, in which an expert human observer attempts to visually match similar structures within the image, is clearly prone to subjectivity. The precision, reliability, and reproducibility of this method all greatly depend on the structural similarities between images, and they all suffer if patient positioning differences are not limited to 2D (two-dimensional) translations. If the positioning differences are more complex, and if structural similarity is low, computer support is required to give reliable results.

In the past, rigid head fixation devices have been used² to facilitate multi-modality image combination. Such devices are used to maintain a constant spatial position of the patient within the tomographic field of view, and they contain distinctive markers that enable the geometric transformation between images to be easily calculated. This straightforward approach has been successfully used in neurosurgery, but it has several disadvantages. The fixation devices are inconvenient when used on an outpatient basis, they often cause problems with patient compliance (especially claustrophobic patients), and they cannot be reapplied in exactly the same way for repeat examinations.

The problem of overlaying images from different modalities would seem to be best approached with automated digital image processing techniques. The authors have investigated the accuracy and convenience of three such methods for superimposing images. Each will be described below, and the strengths and weaknesses of each approach will be considered, with special emphasis on the newly-developed fast correlation technique.

2.1 The Method of Moments

The method of moments is an early approach³ to image matching that is based on geometrical analyses of image structures. The images are reduced to binary representations through threshold-based edge detection⁴. Both images are thus segmented to indicate the boundary of the brain. Pixels inside the brain boundary are assigned a value of one; all others are given a value of zero. The center of mass of the binary object is calculated from its zeroth- and first-order moments, and the angles of the principle axes of the object are obtained from the second-order moments. The differences in the centers of mass and principal axis angles between PET and MRI images give the parameters for registration (for this method, scale factors for both images must be identical). The method of moments is satisfactory for normal anatomy and physiology, but its results deteriorate when applied to abnormal cases that most require reliable diagnosis. For cases involving metabolic deficiencies near the brain periphery, edge detection can falter, in which case the registration parameters are calculated incorrectly. The method also fails when applied to brain images that are circular, since the concept of a principle axis is not meaningful in this case. Additionally, the anatomy of the sinuses near the base of the brain can be complicated enough to frustrate attempts to find its contours.

2.2 Relaxation Labelling

In relaxation approaches⁵, image matching is attempted by matching corresponding individual segments of the images. Relaxation techniques are much less sensitive to edge-detection errors than is the method of moments, and thus tend to be more reliable. As in the method of moments, calculations are based on a segmented binary image. Polygonal approximations⁶ of the boundaries of the two objects to be matched are computed, and features of each line segment are stored. Based on feature similarities, probabilities of matching individual line segments from the two object boundaries are calculated. An optimizing technique⁷ then uses these initial probabilities, or labelling, to determine the final matching probabilities. The final matching parameters are obtained by averaging the transfer functions of matching line segment pairs. The main difficulty of relaxation labelling in its application to medical image registration lies in the determination of valuable line-segment features. The brain images from both PET and MRI are rather blob-like, and segments from the boundaries all have similar properties, leading to ambiguities in the matching. Additional drawbacks to this approach are the sometimes lengthy convergence times of the relaxation algorithms, and the fact that reduction of the image to a simple boundary prior to the matching process discards a considerable amount of valuable information.

Neither of the above two methods permits the detection of matching inner structures of images or the detection of a partial pattern within an image. Altogether, the practical value of these methods is limited to a few special applications with minor medical importance.

2.3 Correlation Analysis

Correlation analysis is primarily used for the detection of similarities between two non-uniform data sets. This technique is sensitive enough to detect the presence of corresponding structures that are barely noticeable by a human observer, and thus is ideally suited for the image matching problem. Realization of the correlation analysis is impeded, however, by the requirement that several dependent parameters for matching (translation, rotation and scale factors) be determined at the same time. Since these parameters are coupled, they cannot be separately calculated from the spatial-domain representations of the images. Changing the orientation of one image will affect the relative translation between the two images, a translation of one image will affect the rotational differences, and scaling can influence both translational and rotational parameters. Previous correlation-based approaches have not dealt with this coupling adequately, with the result that the algorithms have been inefficient and computationally expensive^{8,9}.

A solution for finding 2D translational and rotational differences has been proposed in which translation and rotation differences are corrected in an iterative manner⁸. This algorithm first employs a 2D cross-correlation, in Cartesian coordinates, to find the relative shift between the two images to be matched. This translational difference is accounted for, a second cross-correlation is performed, this time in polar coordinates, to determine the rotational difference, and this difference is accounted for. This process continues in an iterative fashion, alternately compensating for translational and rotational differences, until the compensations are smaller than a certain threshold value.

The major drawback of this approach is that it does not deal with the coupling of the translational and rotational variables. Calculation of translation in the presence of rotational differences results in an incorrect translation, and vice versa. These errors could impede or even prevent convergence of the algorithm, or cause convergence to a wrong result. In addition, the inherent inefficiency in this approach produces unnecessarily high computational expense. Each step in the iteration requires a total of six 2D-FFT's (two-dimensional fast Fourier transforms), that is, four forward and two inverse transforms. Thus, a registration that needed ten iterations would require sixty 2D-FFT's in addition to the numerous image manipulations and coordinate transformations.

A simple modification of this technique has been proposed⁹ which correctly deals with the coupling of the translational and rotational variables. In this procedure, one image is rotated with respect to the other in small increments over some range, and the angle at which the correlation function reaches its maximum is considered to be the best estimate of the rotation for proper registration, with the location of the maximum specifying the shift parameters. Though the algorithm correctly deals with the coupling of translation and rotation, it is very expensive computationally, and does not include scale factor parameters. The inclusion of these parameters would increase the computational expense even more, as would expansion of the method to three dimensions.

3. FAST CORRELATION MATCHING

Significant reductions in computational expense can be gained by taking advantage of some invariance properties of the Fourier Transform. The fast correlation approach to be described here was developed by reformulating previous correlation approaches, taking these invariance properties into consideration. The algorithm efficiently de-couples the matching variables, yielding a fairly simple algorithm that is relatively fast and computationally inexpensive. Each variable is determined individually and in sequence, without requiring iterations that involve multiple 2D or 3D FFT's. For the sake of simplicity, relevant concepts are described below in terms of two dimensions. The principles apply equally well in three dimensions.

3.1 Cross-Correlation

If $I_1(x,y)$ and $I_2(x,y)$ are two spatial domain images, and $F_1(u,v)$ and $F_2(u,v)$ are the corresponding frequency domain representations, then the cross-correlation $R_{1,2}(x,y)$ between these two images is given by

$$R_{1,2}(x,y) = \text{FFT}^{-1} | F_1(u,v) \cdot \text{CONJ}(F_2(u,v)) |,$$

where CONJ refers to the complex conjugate operation, and FFT^{-1} refers to the inverse fast Fourier transform operation. If $I_1(x,y)$ and $I_2(x,y)$ represent similar images which differ only by a translation, the amount of shift or translation can be easily determined from the cross-correlation function. The coordinates x_m, y_m for which $R_{1,2}(x_m, y_m)$ is the maximum value of the correlation function indicate the relative shift or translation between the two images.

3.2 Fourier Transform Relations and Principles of Invariance

The basic principles of invariance are the Fourier relations that apply to the operations of translation, rotation and scaling. These principles are utilized initially to decouple the three matching variables, and then to compute the differences in translation, rotation and scaling between the two data sets.

3.2.1 Translation Translation in the spatial domain corresponds to a phase shift in the Fourier domain. That is, if the image $I(x,y)$ is translated by some distance (x_1,y_1) , then

$$\text{FFT}[I(x-x_1,y-y_1)] = F(u,v)\exp((-j2\pi/N)(ux_1 + vy_1)).$$

Computing the magnitude of both sides,

$$\begin{aligned} | \text{FFT}[I(x-x_1,y-y_1)] | &= | F(u,v)\exp((-j2\pi/N)(ux_1 + vy_1)) | \\ &= | F(u,v) | \cdot | \exp((-j2\pi/N)(ux_1 + vy_1)) | \\ &= | F(u,v) |. \end{aligned}$$

This demonstrates that the magnitude spectrum, and thus the power spectrum of an image is invariant with respect to translation. That is, translation of an image in the spatial domain does not affect the image's power spectrum.

3.2.2 Rotation Rotation of an image in the spatial domain corresponds to a rotation of the spectrum by the same angle, as shown in Figure 2. That is, if an image $I(x,y)$ and its Fourier transform $F(u,v)$ can be transformed into the polar coordinate representations $I(r,T)$ and $F(w,P)$, respectively, where

$$\begin{aligned} x &= r \cos(T), \\ y &= r \sin(T), \\ u &= w \cos(P), \\ v &= w \sin(P), \end{aligned}$$

then the relationship between an image that is rotated by an angle D and its corresponding Fourier transform is¹⁰:

$$\text{FFT}[I(r,T+D)] = F(w,P+D).$$

Rotation in polar coordinates can be thought of as a simple shift or translation of the angular coordinate. The power spectrum in polar coordinates will be invariant to such a "shift", and thus will be invariant to rotation. The polar-coordinate representation can not only be used to eliminate rotational considerations by using the power spectrum of an image, it also enables one to easily determine rotational differences between a pair of similar images from the complex Fourier spectrum, since these differences manifest themselves as simple "shifts" of the angular coordinate. To determine a difference in rotation, one should cross-correlate the two polar-coordinate representations of images that have already been made translation-invariant, and note the rotation angle at which the cross-correlation function $R_{1,2}(r,T)$ is a maximum. This angle will indicate the orientation difference between the two images.

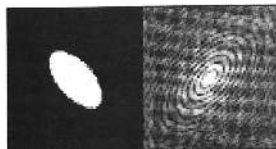
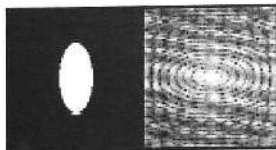


Figure 2 : Demonstration that rotation of images in the spatial domain (left panels) corresponds to the same rotation of the power spectra (right panels).

3.2.3 Scaling Scaling in the spatial domain has two effects in the spatial-frequency domain: an inverse scaling of the frequency axes, and an adjustment in the amplitudes of the Fourier coefficients, as shown in Figure 3. If the x-direction scale factor is given by A, and the y-direction scale factor is given by B, then the Fourier relationship pertaining to scaling is:

$$\text{FFT}[I(Ax,By)] = (1/AB) F(u/A,v/B).$$

A logarithmic transformation of the spatial domain coordinate system converts the multiplicative scaling process to a shift¹¹, making it possible to create a scale-invariant spectrum. This conversion of scale to a shift also enables the detection of scaling factors for image matching by cross-correlation, since the "shifts" (representing scale factors) can be detected in a manner similar to that for detecting rotational differences.

3.3 New Algorithm

The new algorithm, employing the Fourier Transform invariance principles, is summarized in the flowchart shown in Figure 4. The short description given here is followed by more detailed information regarding individual stages of the algorithm.

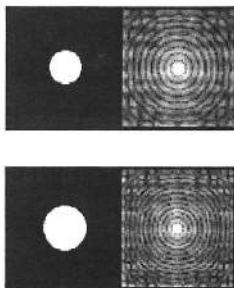


Figure 3 : Demonstration of the effect of scaling of spatial domain images (left) on the image spectra (right).

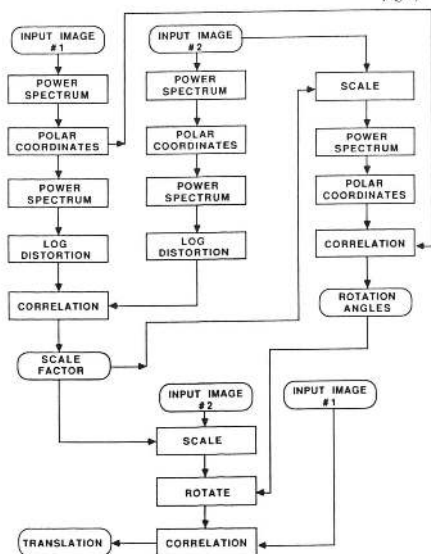


Figure 4 : A flowchart illustrating the various steps in the fast correlation image matching algorithm.

The first step in the image-matching process is to create translation-invariant information from the input images by computing their power spectra. These power spectra are then treated as new images to be matched; they are converted to polar coordinates, and the power spectra of these polar coordinate representations are computed. These last power spectra are thus invariant to both rotation and translation. The relative scaling factors are conveniently determined from these rotation- and translation-invariant power spectra, using the logarithmic transformation. The algorithm currently determines a global scale factor. Solutions for the case in which scale factors differ with coordinate axes are currently being investigated. The polar coordinate versions of the power spectra of the equivalently-scaled images are then cross-correlated in order to determine the relative rotation between the two input sets of image data. After correcting for this rotation, the final step is to determine the relative translation from a correlation analysis of the orientation-equivalent image data sets.

As described above, the first step in the algorithm is the computation of the power spectra of the input images. If $I_1(x,y)$ and $I_2(x,y)$ represent the input images, then the respective Fourier spectra $S_1(u,v)$ and $S_2(u,v)$ are given by

$$\begin{aligned} S_1(u,v) &= \text{FFT} [I_1(x,y) \exp(-j2\pi/N)(ux_0 + vy_0)] \\ S_2(u,v) &= \text{FFT} [I_2(x,y) \exp(-j2\pi/N)(ux_0 + vy_0)] \end{aligned}$$

where x_0 and y_0 represent shifts that are equal to half of matrix size, or $N/2$. The phase shift resulting from this exponential term thus centers the resulting spectra so that the DC component is located at the center of the matrix. The power spectra $P_1(u,v)$ and $P_2(u,v)$ are given by

$$\begin{aligned} P_1(u,v) &= S_1(u,v) \cdot \text{CONJ} [S_1(u,v)] \\ P_2(u,v) &= S_2(u,v) \cdot \text{CONJ} [S_2(u,v)] \end{aligned}$$

As previously noted, these power spectra are translation-invariant, meaning that, at this stage, translation has been de-coupled from rotation.

After converting each power spectrum to polar coordinates, integrations are performed along equally-distributed radial vectors. These integrations need only be performed in one half-plane, because of the inherent symmetry in the power spectrum. This series of integrations reduces the information contained in each 2D power spectrum to a 1D series. The rotational difference between the two input image data sets are then determined from a 1D cross-correlation of these two series.

The intrinsic rotational resolution of the algorithm, apart from interpolation, which will be considered later, is specified by the number of line integrals that are performed. If thirty-two line integrals are performed, for example, the resulting rotational resolution is 180° divided by 32, or 5.6° . In the authors' tests, one hundred and twenty-eight line integrals were performed, yielding a rotational resolution of 1.4° . This simple description of the rotational resolution, however does not consider the actual resolution of the input image data as specified by the number of projections and image matrix size used in reconstruction. Additionally, it should be noted that, because of symmetry in the power spectrum, rotations greater than 180° are detected as rotations modulo 180° .

The line integrals through the power spectra need not begin at the DC component, nor do they need to continue to the Nyquist frequency. In fact, better results can be achieved by ignoring the very low frequency components (which are mainly related to the image background) during the radial integration. Since the high-frequency information content of images from current imaging modalities is very limited (depending on the image resolution), computations can be reduced by integrating only partway to the Nyquist frequency. For optimum performance, the limits of integration should be determined from corner frequencies of the transfer functions of the imaging modalities involved. If two different modalities are involved, the integration should be performed only over those frequency bands which the two modalities have in common, since spectral components outside these bands do not contribute to the correlational matching process.

The rotational difference between two images is sometimes considered to be the difference in angular orientation when the two objects in the images are translationally aligned at their respective centers of mass. This object-oriented approach to rotation, however, is not always clinically relevant. In addition, it is easier, from a computational viewpoint, to eliminate the need for center-of-mass calculations by choosing the image frame itself as the rotational reference. For these reasons, the image-matching method described here determines the angle of rotation required to orient the images with respect to the center of the image frame.

The new algorithm has been implemented within the Asyst¹² software package running on an IBM PS/2 Model 60 computer. The algorithm was validated by computing the relative orientation and shift between two 2D scale-equivalent images. Data size limitations imposed by the Asyst software required that the data be reduced to 64x64 matrices. With this matrix size, the algorithm required three minutes of computation: ninety seconds to compute the rotational difference, sixty seconds to apply this rotation (a total of 2 1/2 minutes to compensate for rotation), and thirty seconds to compute and compensate for translational differences. The use of special-purpose software and hardware would certainly reduce this calculation time. Integration of this algorithm with image reconstruction software typically available on tomographs would allow direct access to the projection data and thus greatly increase the speed and precision of the rotation computations. Regardless of implementation-specific computation times, the algorithm is inherently faster than previous correlation approaches^{6,9}. As previously noted, one approach required as many as sixty 2-D FFT's and ten image rotations to solve the 2-D matching problem, whereas the algorithm presented here requires three 2-D FFT's and only one image rotation.

3.4 Windows

While the fast correlation matching algorithm produced good results with synthetic data, initial tests with clinical single- and multi-modality images resulted in unsatisfactorily large errors in the estimation of the rotation angle. To eliminate these errors, and achieve the theoretically-predicted results, some processing, prior to the matching process, is required.

The basic problem stems from the "window effect" of the rectangular image frame. The spectrum of the image is effectively convolved with the spectrum of the rectangular image frame, and the spectrum of the rectangular shape, which is shown in Figure 5, does not rotate along with the spectrum of the image within the rectangle. If not removed, this stationary artifact will obviously have an adverse influence on the rotational cross-correlation.

The above concepts are illustrated in Figures 6-8. Figure 6 shows two PET images which differ in rotation by 20°. Figure 7 represents contour views of the corresponding spectra. For purposes of illustration, only the low-frequency range of the spectra is shown. It can be seen that, although there is some difference between the two spectra, the difference does not represent a real rotation. The obvious distortion is a direct result of the un-rotated power spectrum of the rectangular image frame. One can see the inappropriate similarity in features between the power spectra shown in Figure 7 and that shown in Figure 5. The application of a Hanning window effectively removes the unwanted spectral components, allowing a true expression of the rotation, as illustrated in Figure 8. These spectra are easily identifiable as rotated versions of one another.

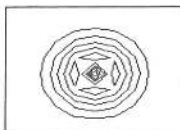
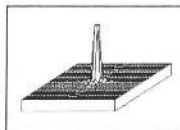


Figure 5 : The power spectrum of a rectangular image frame, in 3D form (above) and in contour view.

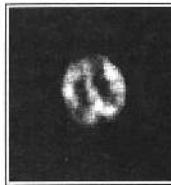


Figure 6 : Two PET images which differ in rotation by 20°.



Figure 7 : The corresponding power spectra of the two PET images shown in Figure 6, illustrating the detrimental influence of the rectangular image frame on rotation.



Figure 8 : The corresponding power spectra of the two PET images shown in Figure 6 after removal of the unwanted spectral components of the rectangular image frame.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the new image-matching algorithm was evaluated for two general classes of problems. The first was the single-modality case, in which an image was matched against a rotated, translated, and noisy version of itself. Single modality experiments were performed with synthetic data and with images from either PET or MRI tomographs. The second test category was the multi-modality case, in which images from different modalities, or images obtained with the same modality with different acquisition parameters, were matched against each other.

Each single-modality experiment required the choice of a reference image and the creation of a test image by rotating and translating the reference image. Images were intentionally corrupted by adding Gaussian noise at three different levels, as specified by the variance of the noise. The three different values of variance used were: 0 (no noise), 5 and 10. Since the mean value of the pixels in the 8-bit images was approximately 40, this last noise level (variance = 10) represented approximately 25% of the mean value of the image. The nine images shown in Figure 1 were chosen as reference images for these experiments. Each reference image was rotated by amounts of -6° , -3° , 0° , $+3^\circ$ and $+6^\circ$, and, for each case of rotation, Gaussian noise with variances of either 0, 5 or 10 was added, yielding a total of 15 test images for each of the nine reference images. A bi-linear interpolation algorithm¹³ was used to perform rotations.

Test images, paired with the proper reference images, were submitted to the matching algorithm, and the results are shown in Table I. Results are presented as a function of rotation angle and Gaussian noise level. Each entry corresponds to a single rotation angle and a single variance value, and represents the average result of all nine reference images as each was matched against its particular class of test image.

Table I: Computed rotation angles for single-modality test images

Noise Variance	Actual Rotation (°)	-6.00	-3.00	0.00	3.00	6.00
0		-5.28	-2.79	0.00	2.64	4.97
5		-5.28	-2.79	0.15	2.79	5.13
10		-5.44	-2.33	0.17	2.64	4.97

It can be seen that the results in Table I show some errors in the computed values of rotation. One source of these errors was the quantization due to the 1.4° rotational resolution of the algorithm. Since the test images were not rotated by angles that were exact multiples of 1.4° , some quantization error occurred. Another source of error was the bi-linear interpolation algorithm used in the rotation, which tended to degrade an image in the process of rotating it. It should also be noted that the fairly coarse (64×64) image matrix used for these tests further limited the rotational precision. The effective resolution of the algorithm under these circumstances is probably no better than 2.8° . Smaller pixels and better interpolation schemes would certainly have reduced the errors shown in Table I.

For multi-modality testing, pairs of PET and MRI images that had been determined (by an expert observer) to represent the same image plane were given to the matching algorithm. The matching parameters were computed and used to re-orient the MRI image so as to match the PET image. Figure 9 illustrates the result of the algorithm's application to a particular PET-MRI image pair. The un-registered image pair is shown in Figure 9a. Figure 9b shows the two 1D series resulting from the radial integrations (at bottom), and the resulting 1D cross-correlation. Figure 9c shows the 2D cross-correlation which indicates translational shifts, and Figure 9d shows the properly registered PET-MRI image pair.

The results of the matching were evaluated by an expert observer, using an interactive graphics device. After re-orientation of the image, the two images were displayed side by side. Expert medical observers then positioned boxes within the PET images to measure the metabolism within the corresponding rectangular regions. The boxes were automatically positioned within the registered MRI image, allowing the expert observer to determine how well the metabolism measurements corresponded to the anatomy, thus giving a qualitative evaluation of the results. The results shown in Figure 9 were judged to be satisfactory.

A second multi-modality test involved a test object that was scanned by both tomographs. In both the PET and MRI scans, the object was positioned so that the rotation angle was 0° . The MRI image of the test object was then rotated to -6° , -3° , $+3^\circ$ and $+6^\circ$, and each rotated image was submitted, along with the original PET image, to the matching algorithm. The results of these tests are shown in Table II.

Table II: Computed rotation angles for dual-modality phantom images

Actual Rotation	;	-6.00°	-3.00°	0.00°	3.00°	6.00°
Computed Rotation	:	-5.60°	-2.75°	0.00°	4.25°	7.00°

These results again show that the algorithm is correctly computing the rotation angle, with the expected precision. Primary sources of error

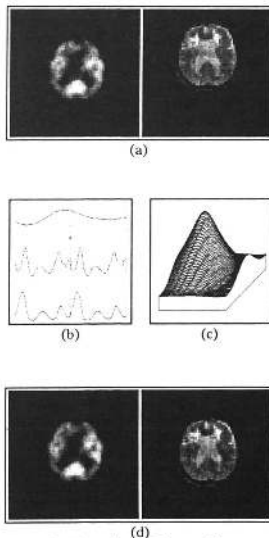


Figure 9: Illustration of the matching process for a particular PET-MRI image pair.

are presumably the same as those in the first multi-modality test case, namely, the low resolution of the image matrix and the weakness of the rotation algorithm. It should be noted that the rotation algorithm's detrimental effect is relevant only because the algorithm was used to generate test images. The computation of rotational registration parameters in a real application will be unaffected by rotation algorithms.

Additional matching difficulties arise in those applications where, in contrast to the above described MRI/PET image combinations, one of the images represents only a partial area of the second image. In such cases, the algorithm must be able to perform object recognition, that is, those body parts in the image with the smaller view area must be identified in the image with the 'whole-body' view. The recognition process is impeded, however, by the fact that differences in the imaging parameters can lead to considerable differences in the appearance of the object. The performance of the fast correlation algorithm on such an image combination is demonstrated in Figure 10. The translational parameters required to position the partial pattern shown in Figure 10b within the whole-view image shown in Figure 10a are determined from the correlation function shown in Figure 10c.

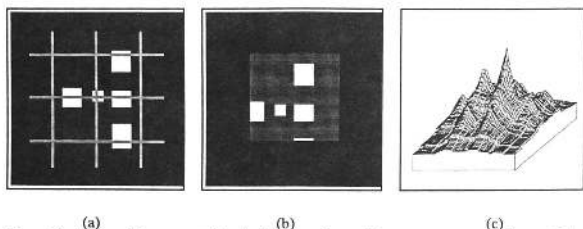


Figure 10 : A case of image matching in which one image (b) represents a partial area of the other image (a). The 2D cross-correlation of the two images is shown in (c). These images represent synthetic data.

Real images, of course, are normally corrupted by some amount of noise, artifacts and imperfections of the imaging modalities, so that the detection process often requires some pre-processing (e.g., contrast enhancement) of the images prior to the correlation analysis. Figure 11 shows an image combination in which a phantom is shown in two projection photographs, a normal projection radiograph (Figure 11a), and an image produced by a linear accelerator used for cancer treatment (Figure 11b). In spite of the poor image quality, the described matching algorithm determines the registration parameters with the predicted precision. Figure 12, which, for purposes of demonstration, illustrates only the method for determining translational differences, shows the 2D cross-correlation of the two images. The cross-correlation function has a well-defined peak at the x - and y -coordinates that are the required registration parameters.

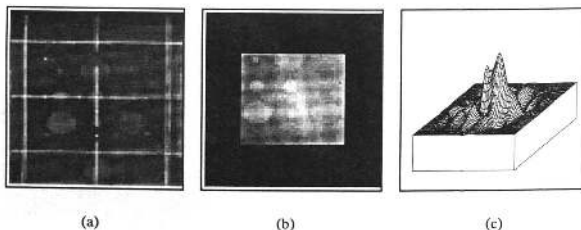


Figure 11 : An example of a clinical case of image matching in which one image represents a partial area of the other. In this case, a linear accelerator image (b) is to be matched against an x-ray image. The 2D cross-correlation of the two images is shown in (c).

5. CONCLUSIONS AND FUTURE WORK

The authors have developed a general-purpose matching algorithm which provides for precise comparison of medical images obtained at different times or with different modalities. This algorithm is faster and more reliable than previous techniques, and it can be practically implemented on low-cost computer systems, giving it the potential for wide clinical acceptance. Experiments to date indicate that the algorithm can be successfully applied to a wide variety of applications of image matching.

With this algorithm, the reliability and quality of image matching are completely determined by the properties of the input data. Limitations in the final result are imposed only by the quality, information content and comparability of the two sets of data. Low-resolution images, for example, do not allow high-resolution matching. In situations involving images of different resolutions, the image with the lowest resolution will determine the precision of the matching. Again, this restriction results, not from the algorithm, but from the quality of the data. The results presented here were certainly adversely affected by the reduction of images to 64x64 matrices. Implementation of the algorithm with higher-resolution matrices would be expected to improve the resolution of the matching. Similarly, results that involved only high-resolution modalities would be expected to be better than results with low-resolution PET images.

Future work will include the determination of the algorithm's reliability and precision as a function of image parameters such as resolution, noise level and bandwidth. Formal guidelines for application of the correlation algorithm to multi-modality images should also be established. These should include criteria for image comparability, so that users do not attempt to match incomparable data (apples and oranges) and rely on meaningless, formal correlations. Further work will also include the extension of the algorithm to three dimensions, and computation of the scale factor for all three coordinate axes.

ACKNOWLEDGEMENTS

This project was supported by BRSG S07 RR07022-20 awarded by the Biomedical Research Support Program, Division of Research Resources, National Institutes of Health, and also by a student fellowship to Anthony Apicella from the Education and Research Foundation, Society of Nuclear Medicine.

REFERENCES

- [1] S. Shalev. Personal communication.
- [2] M. Bergstrom and T. Grietz, "Stereotaxic Computed Tomography", American Journal of Roentgenology, 127:pp167-170, 1976.
- [3] M. K. Hu, "Visual Pattern Recognition by Moment Invariants", IEEE Transactions on Information Theory, Volume IT- 8, pp179-187, February 1962.
- [4] J. Winter, "Discrete Contour Representation of Image Matrices", IEEE Transactions on Medical Imaging, Volume MI-3, pp149-154, 1984.
- [5] B.Bhanu and O.D. Faugeras, "Shape Matching of Two- Dimensional Objects", IEEE PAMI-6, No.2, pp137-156, March 1984.
- [6] T. Pavlidis and S.L. Horowitz, "Segmentation of Plane Curves", IEEE Transactions on Computers, Volume C-23, No. 8, pp860-870, August 1974.
- [7] J.L.Mohammed, R.A.Hummel, and S.W.Zucker, "A Gradient Projection Algorithm for Relaxation Methods", IEEE PAMI-5, No.3, May 1983.
- [8] L.S. Hibbard, J.S.McGlone, D.W.Davis, and R.A.Hawkins, "Three Dimensional Representation and Analysis of Brain Energy Metabolism", Science, Vol.236, June 26,1987, pp1641-1646.
- [9] E. De Castro and C. Morandi, "Registration of Translated and Rotated Images Using Finite Fourier Transforms", IEEE PAMI-9, No. 5, pp700-703, September 1987.
- [10] R.C. Gonzalez, Digital Image Processing, Addison-Wesley, 1986.
- [11] J. Altmann and H.J. Reitbock, "A Fast Correlation Method for Scale- and Translation-Invariant Pattern Recognition", IEEE- PAMI-6, No.1, pp46-57, January 1984.
- [12] Asyst User's Manual. MacMillan Software Company, 866 Third Avenue, New York, New York, 10022.
- [13] T. Pavlidis, Algorithms for Graphics and Image Processing, Rockville, Maryland:Computer Science Press, 1982, pp299-316.