Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

SaBP2.14

Image Registration based on Neural Network and Fourier Transform

Antoine B. Abche [1], Fadi Yaacoub [1] , Aldo Maalouf [1] and Elie Karam [1]

[1] Electrical Eng. Dept., University of Balamand, P.O. Box 100, Tripoli, Lebanon

Abstract – **An image registration technique based on feed forward neural network and Fourier Transform is developed and presented. In the proposed scheme, the spectrums of the acquired images are computed, the Fourier coefficients within a selected central window of each spectrum are extracted and fed as inputs to the neural network. The feed forward neural network is implemented to estimate the transformation, defined in terms of the translation, rotation and magnification parameters, to align the corresponding images. This approach does not estimate the various registration parameters separately. They are estimated simultaneously leading to a better-optimized set of registration parameters. The approach is successful and yields better results than another Fourier based registration technique. The approach is validated on 2D images. However, it can be easily extended to 3-D application.**

Keywords – **Fourier Transform, Image Registration, Neural Network, back propagation, MRI,**

I. INTRODUCTION

Image registration is the process of determining the geometrical transformation that brings two sets of data into coincidence in the same coordinate system. Thus, the registration techniques can be implemented to integrate complementary information acquired by different imaging modalities. That is, anatomical structures (in modalities such as MRI and x-Ray CT) would be of great importance in the analysis of functional images that lack recognizable structural information (in modalities such as PET and SPECT). Different techniques were developed and presented in the literature [1-3]. Similarly, the registration of images acquired by the same imaging modality brings the corresponding information into coincidence so that the same features visible in both images overlap and at the same time the differences (for example a tumor that has grown or shrunk) are readily apparent. For example, MRI is currently one of the fastest developing medical imaging modality, and is applied to an increasingly number of different medical diagnostic situations. Acquisition of the same anatomical region offers opportunities to monitor changes in a particular Region Of Interest (ROI) over time. That is, it makes possible to investigate, for instance, how a patient responds to a particular treatment and allow medical practitioners to monitor the progress of a tumor over time. Furthermore, since the positioning of the patient in the same position during different acquisition sessions is practically impossible, the geometric alignment of the acquired images in the same coordinate system is required so that similar features are overlaid on each other. Consequently, the analysis of the corresponding images is facilitated for better diagnostic purposes.

Various registration techniques based on Fourier Transform have been proposed [5-8]. They are based on the following three properties: shifting (translation), rotation and scaling properties [4]. While some researches propose registration techniques that estimate the rotation and translation parameters using an iterative approach and cross correlation analysis [5] or by estimating the cross correlation at various increments of the registration parameters [6], other researches extend these approaches to include a magnification in the overall transformation using the correlation as a similarity metric [7] or its variant the phase correlation [8]. These properties were implemented to decouple the various registration parameters, estimate each parameter separately and form the overall transformation to align the corresponding spatial images.

In this work, a registration technique based on Fourier transform and a feed forward neural network is developed to estimate the registration parameters to align medical images acquired by the same modality. The proposed approach is validated by registering 2-D images. However, it can be easily extended to align 3-D images.

II. METHOD

The proposed approach does not decouple all the registration parameters from each other. They are estimated simultaneously. Consequently, that would lead to a better optimum value than if they are estimated separately. The overall transformation between the two images is given by:

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} Mx & 0 \\ 0 & My \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \qquad (1)$$

Where $\vec{x}_2$ (x_2, y_2) and $\vec{x}_1(x_1, y_1)$ are the spatial coordinates of a corresponding pixel in the transformed and the original images, respectively. The orthogonal 2 by 2 rotation matrix is defined in terms of an angle θ in 2-D image registration. The 2 by 2 magnification matrix M is defined in terms of M_x and M_y along the x and y-axes, respectively. However, warping along other directions (such as xy direction) can be taken into account by filling the appropriate zero values. The vector $\vec{t}_x$ (t_x, t_y) defines the translation parameters along the x and y axes. However, this approach can be easily extended to 3-D image registration. That is, the transformation is parameterized in terms of a rotation matrix defined by the three Euler angles (α,β,δ), a 3 by 3 magnification matrix (Mx, My and Mz) and a translation vector (t_x, t_y, t_z) [10,11].

The approach is described as follows: the spectrums of the images to be registered are computed using Fourier Transform. Then, a square window is selected and is centered at the origin of each spectrum (zero frequency). The corresponding Fourier coefficients within the selected window are extracted and fed as input to a feed forward neural network. The outputs of the neural network are the estimated parameters of the overall transformation required to align the two images. A block diagram of the registration approach is illustrated in Figure 1.

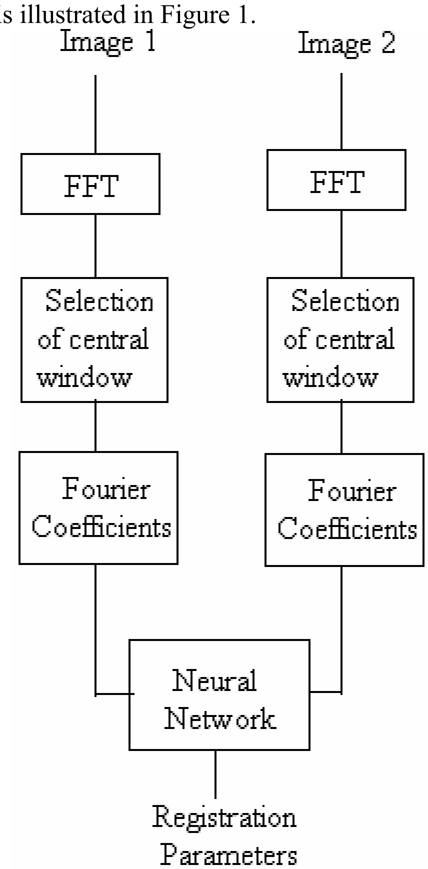


Figure 1: block diagram of the registration approach.

Figure 2 illustrates the neural network implemented in this work. It consists of an input layer, a hidden layer and an output layer. The input layer depends on the selected window. For example, it consists of 64 neurons for an 8 by 8 window. Each neuron in the input layer corresponds to one Fourier coefficient extracted from the central window of the FFT image. The hidden layer consists of 40 neurons since the tests have shown that this number is more appropriate. Sigmoid transfer functions were deployed in the hidden layer neurons i.e. the output is calculated according:

$$output = z_i = \frac{2}{1+e^{-2\,net1}} - 1 \qquad (2)$$

where "net1" is the "net input". The number of neurons in the output layer is equal to number of parameters to be estimated. Linear functions characterize the output layer neurons i.e. they are given in matrix form as:

$$output = \vec{y} = f(net2) = W_k\ \vec{z} \qquad (3)$$

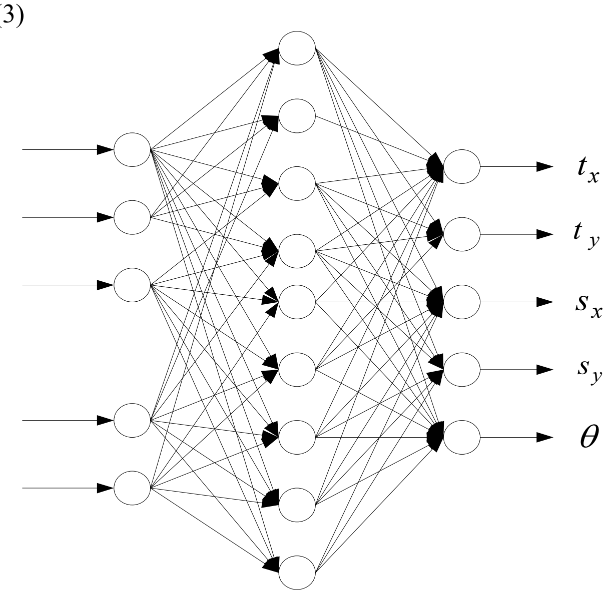


Figure 2 Structure of the Neural Network

The neural network is trained to estimate the corresponding weights between the different layers. In this regard, an image dataset consisting of 27 MR images is selected. Each image is transformed to the coordinate system of the second modality using a set of pre-specified transformations covering the span or range of possible transformations for the particular application. This set depends on the range and the increment (fine or coarse) associated with each parameter. The generated sets of pair of images are used to train the network using Levenberg-Marquardt method with gradient descent learning functions [12]. The training phase begins by initializing randomly the weights of the network (W_0). Then, the weights (w) and biases are updated according to the Levenberg-Marquardt optimization:

$$\Delta\vec{w} = -[\nabla^2 \varepsilon]^{-1} \nabla\varepsilon \qquad (4)$$

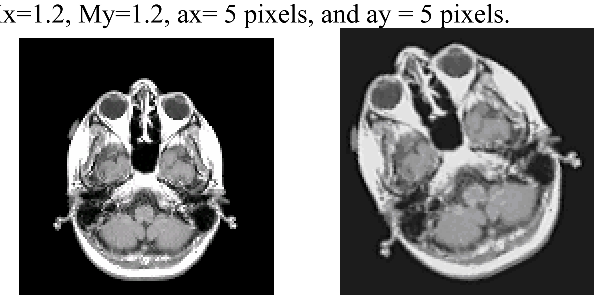
$$\Delta\vec{w} = (J^T J + \mu I)\, J^T\ \vec{\varepsilon} \qquad (5)$$

where $\nabla^2\varepsilon$ is the Hessian matrix, $\nabla\varepsilon$ is the gradient, J is the Jacobian matrix with respect to the weights, μ is the parameter that controls the stability and the rate of convergence (adaptive value), I is the identity matrix and ε is defined as the errors between the desired output and the corresponding computed values. The parameter μ is adjusted according to the result obtained each iteration i.e. it is multiplied by a value β if the result leads to an increased ε. On the other hand, it is divided by β if the results lead to a reduced value of ε. Thus, the weight vector $\vec{W}$ is updated in the direction opposite to the estimated gradient. This procedure is repeated until the MSE function is minimal and the weight vector approaches an optimal value.

The time to accomplish the training depends on the window size and the increment of each parameter. Even though, the training phase is lengthy, the registration phase consists merely of computing the FFT, extracting the central window and feeding it to the neural network.

III. RESULTS

Having trained the neural network and estimated the corresponding weights between the different layers, the next step is to evaluate the performance of the proposed registration approach and study the precision of the registration. In this regard, MRI images were selected from a group of images that do not belong to the training data set. Each image is transformed from its coordinate system to the coordinate system of the second modality using a predefined transformation. Figure 3 shows a typical original (3-a) and transformed images (3-b). The pre-specified parameters of the transformation were θ= 25°, Mx=1.2, My=1.2, ax= 5 pixels, and ay = 5 pixels.



(a) (b)

Figure 3: a- Original Image, b- Transformed Image

Then, an 8 by 8 window is defined and the corresponding 64 Fourier coefficients were fed to the neural network to recover the original parameters. Using the Fourier-Neural network based approach, the optimized parameters were: $t_x = t_y = 4.99$, $M_x = M_y = 1.199$, $\theta = 24.999°$. It is clear that this approach is successful in estimating the transformation parameters. Similar results are obtained with the other experiments.

A. Effect of the window's size

Having the spectrums of the original and transformed images, adequate amount of spectral coefficients should be taken to provide the necessary information required to estimate the registration parameters. This is dependent on the size of the central window overlaid on the corresponding Fourier spectrums. In this context, a number of experiments were performed for various windows' sizes. In each experiment, the previous image dataset (27 MRI images) is used to train the neural network. Each experiment corresponds to a different size of the central window. The outlined training and registration procedures in the previous section were performed. Having estimated the registration parameters, the precision of the registration was evaluated by comparing the recovered parameters with the pre-specified transformation. In this regard, the corresponding residuals were computed and plotted versus the number of Fourier coefficients corresponding to the size of the window. Figure 4 shows the dependence of the residuals on the window's size. It is evident that the error decreases as the size of the window increases until it reaches a plateau. The plateau could be related to the increments of the parameters during the training procedure.

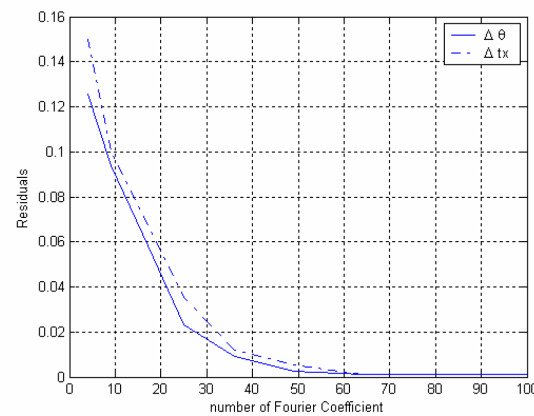


Figure 4: dependence of the precision of registration on the window's size.

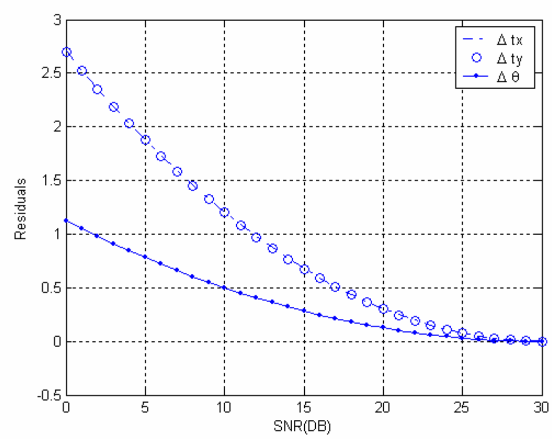


Figure 5: the effect of noise on the precision of the registration

B. Effect of noise

In this regard, each pixel in the original and transformed images is perturbed randomly and independently by Gaussian measurement error. After the neural network is trained on the noisy images and the corresponding weights are estimated, the precision of the registration is evaluated using the proposed approach on images other than the training data set. Figure 5 depicts the dependence of the residuals Δθ (°), Δtx (pixels) and Δty (pixels) on the noise. It is evident that as the Signal-to-Noise-Ratio (SNR) increases, the precision of registration is improved. It yields acceptable results with Low SNR.

C. Residuals of point landmarks

The proposed approach can be evaluated quantitatively by studying the residuals of corresponding point landmarks in both images along x (Δx), y (Δy) directions as well as the residuals in distance (Δd). In a similar fashion, a comparison with another FFT-based approach can be performed. A description of the letter approach is found in [8]. It can be summarized as follows: First, the power spectrums of the two images are computed. According to the shift property, the resulting spectrum images are related only by rotation and magnification parameters. Then, the Fourier spectrums of the original spatial images are converted from rectangular coordinates to polar coordinates (ρ,θ). Thus, the rotation and magnification are reduced into shifting along their corresponding axes. These two parameters are estimated using the phase correlation technique and couple derived relations. Having rotated and scaled one image with respect to the other, the translation can be estimated using the phase correlation technique.

Having identified the corresponding pairs of point landmarks in both images, the Neural-FFT and FFT-based approaches were performed to register the images shown in Figure 3. Table 1 and 2 show the residuals Δx, Δy and Δd of the corresponding point landmarks after registration was performed. Clearly, the precision of the registration is much better using the proposed Neural-FFT based approach than the FFT-based approach. That is clearly observed by the smaller residuals in distance for each of the four point landmarks. Thus, the developed approach leads to the estimation of the registration parameters with a very good precision.

TABLE 1
RESIDUALS USING THE NEURAL NETWORK_-FFT APPROCH

Marker ID	Δx (Pixel)	Δy (Pixel)	Δd (Pixel)
1	0.058	0.056	0.081
2	0.024	0.036	0.043
3	0.101	0.129	0.164
4	0.032	0.056	0.065

TABLE 2
RESIDUALS USING AN FFT- BASED APPROACH

Marker ID	Δx (Pixel)	Δy (Pixel)	Δd (Pixel)
1	1.46	1.38	2.009
2	0.61	0.89	1.079
3	2.56	3.17	4.075
4	0.82	1.38	1.605

Unlike the previous Fourier based image registration technique, The FFT- Neural network based approach assumes that the magnification could be different along different directions. The magnification matrix indicates that a total of 4 parameters could be defined. Similarly, 9 parameters can be defined to perform a 3-D registration of two data sets [9-10]. Furthermore, the developed approach allows all parameters to be varied simultaneously to estimate the registration parameters. That could lead to a better optimum set of the parameters of the overall transformation. Having trained the network, the developed approach requires to perform FFT to align the two images involved.

IV. CONCLUSION

In this work, an approach based on Fourier Transform and feed forward neural network to align medical images is developed and introduced. This approach is compared with another based Fourier transform approach by studying the residuals of the corresponding pair of point landmarks in both images. The results show that the registration is very successful with a very good precision. Its performance is better than the other FFT-based approach. The precision could be related to the increments (of the registration parameters) associated with the training procedure of the neural network.

REFERENCES

[1] Maintz, J. B. A. and Viergever M. A.,"A Survey of Medical Image Registration", Med. Image Anal., Vol 2, pp. 1-36, 1998.

[2] Fitzpatrick J., Hill D. and Maurer C., "Image Registration", M.sonka and J. Fitzpatrick eds, Bellingham, WA, Medical image Processing and Analysis, SPIE, Vol. 2, pp 447-513.2000.

[3] Pluim J.P.W., Maintz J.B.A. and Viergever M.A.," Mutual Information-based Registration of medical Images A Survey", IEEE Transaction On Medical Imaging, Vol. 22, No 8, 2003, pp 986-1004.

[4] Raphael Gonzalez & Robert woods," Digital Image processing", Prentice Hall, 2002.

[5] 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 1987, PP 1641-1646.

[6] E. De Castro and C. Morandi, " Registration of Translated and Rotated Images Using Finite Fourier Transforms,", IEEE Trans. On Pattern Anal. & Mach,. Intel., Vol. 9, No. 5, September 1987, PP 700-703.

[7] A. Apicella, J. Shane Kippenhan and J. H. Nagel " Fast Multi-modality Image Matching", Proceedings of SPIE, Vol. 1092, 1989, pp. 252-263.

[8] Srinivasa B. R. and Chatterji B. N., An FFT-Based Technique, and Scale Invariant Image Registration. IEEE Transactions on Image Processing, Vol 5, No 8, 1996.

[9]J. Altmann and H.J. Reitbock," A fast Correlation Method, for Scale and Translation Invariant Pattern Recognition" IEEE Trans. On Patt. Anal. & mach. Intel., Vol 6, No. 1, Jan. 1984, pp. 46-57.

[10] Abche A.B., Tzanakos G.S. and Micheli-Tzanakou E., "A New Method for Multimodal 3-D image Registration with External Markers", Nucl. Sci. symp.On Med. Imag. Conf., Oct 30-Nov5, 1822-1826, 1994, Norfolk, Virginia, USA.

[11] George S. Tzanakos, Antoine B. Abche, Evangelia Micheli -Tzanakou, Theodore Stahl, Xiaoming Wan and Tony Yudd, "Multimodal 3-D Image Registration of MRI - SPECT Volume Images", World Congress 2000, IEEE – EMBS , paper # 5886-26335, Chicago, 23-28 July 2000.

[12] Hagan M. T. and Menhaj M., Training Feed Forward Networks with the Marquardt Algorithm, IEEE Transactions in Neural Networks, Vol 5, No 6, pp 989-993, 1994 .