Comparison of Recent Least Square Approaches for Fusion of Multimodal Medical Images

Mohammad EsmaeilZadeh [a] **Hamid Soltanian-Zadeh** [a,b] **Gholam-Ali Hossein-Zadeh** [a]

m.esmaiel@ece.ut.ac.ir hszadeh@ut.ac.ir ghzadeh@ut.ac.ir

[a] Control and Intelligent Processing Center of Excellence, School of ECE, Faculty of Eng., University of Tehran, Tehran 14395-515, Iran.
[b] Radiology Image Analysis Laboratory, Henry Ford Health System, Detroit, MI 48202 USA.

Abstract: *Image fusion is a useful technique toward the better analysis of multimodal medical images. In this paper some methods are presented for image registration and fusion with focus on CT and MRI images. All proposed methods are Least squares based which are modified by using FFT for faster performance. To deal with intensity mismatches between multimodal images, some previous methods are reviewed and a new modified method is also provided. The new method proposed, utilizes the combination of two famous structural characteristics of images, instead of one, that are gradient and features obtained from phase congruency model. We applied the proposed method to CT and MRI images of head and thorax. Results suggest the priority of the new proposed method over old ones, due to higher "correctly aligned percent" and smaller "mean alignment error", specially for head images. Visualization of the registration is done via color-fusion.*

Keywords: Multimodal image registration, medical image fusion, analysis of CT and MRI images, structural characteristic.

1 Introduction

Nowadays, medical imaging is a powerful tool for diagnosis and analysis of disease in medicine. There are many imaging methods such as computerized tomography (CT), X-Ray radiography, magnetic resonance imaging (MRI) and positron emission tomography (PET), which magnify different aspects of human body, and consequently help physicians in dealing with various diseases. Accordingly the analysis of medical images can provide a deeper insight into the human body.

One of the widely-used techniques in medical image analysis is "Medical Image Fusion", which combines the primary characteristics of a set of multimodal images (e.g. CT and MRI) to provide a fused image [1].

To illustrate the importance of medical image fusion, consider a patient who has a brain tumor and must undergo a surgical operation for this reason. Although a CT scan of his head, and a head MRI each provides useful information for a surgical plan, but a CT-MRI fused-image provides a better visualization of the proximity of the tumor to bony structures [2]. As it is obvious, for an efficient image fusion, a precise image registration is of great importance. To achieve a precise registration, it is possible to exclude the tumour region in the image registration. We will return to this method of ROI-selection (region of interest) later in section 2.

An example of CT-MRI fusion is illustrated in Fig. 1. In the fused image, bone structures are shown in red and tissue structures are in blue.
In the registration procedure, various similarity metrics may be used including [1]:

1. Cross correlation,
2. Sum of squared differences (SSD) [2,3],
3. Mutual information,
4. Phase correction.

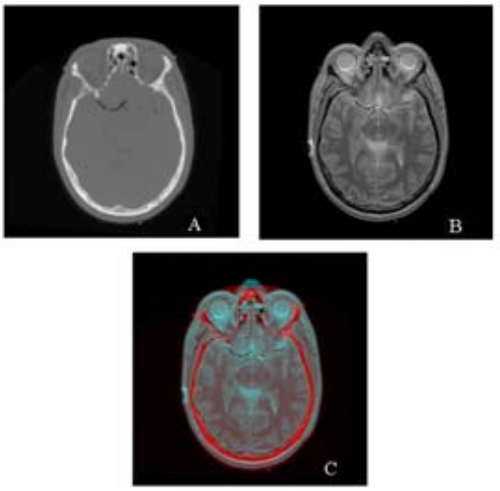

Fig. 1: (A) CT image; (B) proton density-weighted MRI image, (C) fused-image

Generally, in the procedure of image registration, one of the images undergoes a series of geometric transformations, whereas the other one is kept unchanged. For medical images, "scaling" can be handled easily, due to the known parameters of the imaging systems, and is ignored in this context.
Finding the optimal "rotation" for alignment of medical images is the most challenging one. To deal with it, first the SSD method for finding the optimal "shift" is introduced, and then by considering rotations of few degrees and also using some approximations, method is used for handling the "rotation" [1, 2].
Another issue in the multimodal image registration is the intensity difference between images. One of previous approaches to this problem is the intensity remapping method. In this paper we introduce a modified method by utilizing the combined structural features of images produced by gradient and phase congruency model [4,5].
After the image registration, image fusion is performed as the final step. There are lots of methods for fusion. Some of them are introducedin [1] and [6]. Here we use color-fusion.
So in brief, the proposed method in this paper is an SSD one which is modified for increasing speed by applying FFT. Moreover, since the structural features of images are used, the method is independent of pixel intensity, which makes it suitable for multimodal images.
All the underlying theories of this paper and also the data specifications are presented in section 2. Section 3 contains the experimental results, and section 4 is dedicated to discussion and conclusion.

2 Materials and Methods

2.1 Modified SSD Method by Using FFT

In the simple SSD method, the optimal shift for alignment of two signals (images), namely f and g, is achieved through minimizing the cost function defined in Eq. (1) over different shift values of x.

$$\begin{cases} E(x;f,g,w)=\sum_{n=1}^{N}[f_{n-x}-g_n]^2 w_n \\ \hat{x}=\arg\min_x\left[\sum_{n=1}^{N}[f_{n-x}-g_n]^2 w_n\right] \end{cases} \quad (1)$$

As signals (images) f and g may have just partial overlap, or we may want to exclude some parts intentionally (as hinted in the scenario of section 1), a weighting function w is imposed on the cost function. w is '1' in the region of interest (ROI) and '0' outside of it.
The above problem of least square optimization is computationally expensive. However Eq. (1) can be expanded to:

$$\sum_n f_{n-x}^2 w_n - 2\sum_n f_{n-x} g_n w_n + \sum_n g_n^2 w_n \,. \quad (2)$$

Since the last term in Eq. (2) is constant (with respect to x), it will be ignored in the optimization. Using the definition of convolution, the above relation is simplified to a new error function $\tilde{E}$ as below:

$$\begin{cases} \tilde{E}(x;f,g,w)=\left\{\bar{f}^2 * (gw)\right\}_x - 2\left\{\bar{f} * (gw)\right\}_x \\ \bar{f}_x = f_{-x} \end{cases} \quad (3)$$

In order to reduce the computation complexity in time domain (and spatial domain for images), Eq. (3) is revised in the frequency domain.

$$\tilde{E}(x;f,g,w)=\Im^{-1}\left\{\Im(\bar{f}^2).\Im(gw)\right\}_x - 2\Im^{-1}\left\{\Im(\bar{f}).\Im(gw)\right\}_x \quad (4)$$

In the above Eq., $\Im$ denotes the Fourier transform. By using Eq. (4), alignment error is calculated efficiently (50-600 times faster than the simple SSD method [3]) for every possible shift, and the optimal one is easily obtained.

2.2 Intensity Remapping Method

For aligning signals with different levels (brightness for an image) and amplitudes (contrast for an image), Eq. (1) fails to provide a desirable result. Intensity remapping (IR) is a solution, which applies a linear transformation on the signal amplitude (pixel intensity) to compensate the differences. Thus f will be replaced by $s_0+s_1 f$, in which s_0 is the brightness adjustment and s_1 is the contrast adjustment.
To take advantage of the matrix algebra, Eq. (1) is modified after applying intensity remapping transformation as Eq. (5).

$$E(x,s;h,g,w)=\sum_n \left[h^T(f_{n-x})s-g_n\right]^2 w_n \qquad (5)$$

In the above equation, the term $h^T(f_x)s$ is an arbitrary transformation on the intensities of image f. In the IR method:

$$h^T(f_x)=\begin{bmatrix}1 & f_x\end{bmatrix} \quad , \quad s=\begin{bmatrix}s_0\\ s_1\end{bmatrix}. \qquad (6)$$

By minimizing the error function with respect to x and s, registration is done by including the brightness and contrast adjustment.
The overall method for error minimizing is discussed in the next section, after describing image rotation.

2.3 Image Rotation

So far, the method for finding the optimal shift to register two images is discussed, now we extend the method to comply with image rotation. Since the image rotation is usually very small (few degrees), one can approximate the rotated version of f through Eq. (7).

$$R(f,\theta)\approx R(f,0)+\frac{\partial f}{\partial \theta}\theta = f+\frac{\partial f}{\partial \theta}\theta \qquad (7)$$

We can replace f in IR method by $R(f,\theta)$, which leads to:

$$\begin{aligned} &s_0+s_1R(f,\theta)= \\ &=s_0+s_1\left(f+\frac{\partial f}{\partial \theta}\theta\right)=s_0+s_1 f+s_2\frac{\partial f}{\partial \theta}, \\ &s_2=s_1\theta \end{aligned} \qquad (8)$$

So in order to solve the rotation problem, Eqs (5) and (9) should be considered together.

$$h^T(f)=\begin{bmatrix}1 & f & (\partial f/\partial\theta)\end{bmatrix} \quad , \quad s=\begin{bmatrix}s_0\\ s_1\\ s_2\end{bmatrix} \qquad (9)$$

Now by expanding Eq. (5), omitting uncorrelated terms to s and x and using $h(f_x)=h_x$ for simplicity, Eq. (10) will be obtained:

$$\begin{aligned} &\tilde{E}(x,s;h,g,w)= \\ &=s^T\left[\sum_n h_{n-x}h_{n-x}^T w_n\right]s-2\left[\sum_n h_{n-x}^T g_n w_n\right]s= \\ &=s^T A_x s-2b_x^T s \end{aligned} \qquad (10)$$

For minimizing the error function with respect to s, matrix algebra is used and the answer is summarized in the linear system of Eq. (11).

$$\begin{aligned} &A_x s=b_x \\ &A_x=\Im^{-1}\left[\overline{\Im\{hh^T\}}\Im\{w\}\right]_x \\ &b_x=\Im^{-1}\left[\overline{\Im\{h^T\}}\Im\{gw\}\right]_x \end{aligned} \qquad (11)$$

So the method in short is [2]:

1) set $\Theta=0$, this is total amount of rotation,
2) $g=R(g,\Theta)$,
3) calculate A_x and b_x for every possible shift x by using FFT,
4) solve Eq. (11), and the optimal coefficients of s will be obtained,
5) next to do is to calculate the cost function of Eq. (10) for every shift x; take the minimum one and record he corresponding x and s,
6) as s is known, due to Eq. (8), θ is calculated by s_2/s_1 and $\Theta=\Theta+\theta$,
7) goto step 2 if $\theta>0.5^\circ$, or if 6 iterations are not completed,
8) the solution is x and Θ.

2.4 Structural Characteristics of Images

The methods which are discussed so far are only applicable when the brightness difference between images of interest is small. However for medical images of different modalities, due to the intensity discrepancies of corresponding tissues (e.g. in CT images bone structure are bright, whereas they are dark in MRI), this method fails to respond in a desired manner. One solution is to use structural characteristics of images as a feature space, instead of pixel intensities. Two of them are utilized here: magnitude of the gradient, which is discussed in section of "proposed method", 2.5; and the characteristics extracted from phase congruency (PC) model (this method is taken from [1]).
The PC model, which is based on local energy of the signal (image), claims that important structural characteristics of a signal (image) occur at the point where the fourier components are maximally in phase. Due to the simple definition of the model in Eq. (12), if all fourier components are completely in phase, the ratio of Eq. (12) would be 1, which is the highest possible. Otherwise, local energy (numerator) would be less than the sum of fourier amplitude (denominator) and the ratio will fall beneath 1.

$$PC(x) = \frac{|E(x)|}{\sum_n A_n(x)} \tag{12}$$

In the case of 2D signals, Eq. (12) is calculated in several orientations, and the structural features are obtained by using the principal moments of PC (for a complete overview of phase congruency model see [4] and [5]).
If we denote the structural image of A by m_A, all necessary is to replace f and g in Eq. (5) with m_f and m_g, and the remaining procedure is intact.

2.5 Proposed Method

Based on the discussed topics till now, two new modified methods will be proposed in this section. First we used magnitude of gradient as a structural feature, instead of PC-model feature. The formula is presented in Eq. (13) and it postulates that significant features are found where magnitude of the gradient is maximum.

$$f_\nabla = |\nabla f| = \sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2} \tag{13}$$

In the second approach, we tried to take advantage of both structural features and so we defined a combined cost function as in Eq. (14). Here k_{PC} and k_∇ are considered equal.

$$C(x,y,s,h,w) = k_{PC}\tilde{E}(x,y,s,h(m_f),m_g,w) + k_\nabla \tilde{E}(x,y,s,h(f_\nabla),g_\nabla,w) \tag{14}$$

In the above Eq., m_f and m_g are PC-structural features of f and g, as mentioned in 2.4, f_∇ and g_∇ are magnitude of gradient as in Eq. (13), and $h()$ is the vector function introduced in Eq. (9). The same as the method explained in 2.3, this cost function should be minimized and optimal shift values (x and y), and optimum rotation will be obtained.

2.6 Data

To measure the efficiency of the discussed methods, two set of images are used. Each set of images contains one CT image plus three MRI images (PD, T1 & T2-weighted). One set includes the images of an axial slice of the brain and the other contains the images of a coronal slice of the thorax. All images are 8-bit grayscale having 256×256 pixels. They are selected from the datasets of “Visible Human Project” from the “National Library of Medicine”.

2.7 Testing Method

For each image set, there exists a “gold standard optimal shift” that perfectly aligns CT and MRI images and is achieved by hand (trial and error). These optimal shifts are used for quantitative evaluation of the algorithms. The threshold of 3 pixels for alignment error (distance between the obtained optimal shift and gold standard one) is used [1].
Testing procedure is the same as in [1]; for each image pair, 100 random ROIs (w function in Eq. (1)[1]) are produced, and “mean alignment error” and “correct aligned percent” are calculated for the below methods:

1) Intensity remapping (IR)
2) Intensity remapping with phase congruency model as a structural feature (PC)
3) Intensity remapping with gradient as a structural feature (GRAD)

[1] For better comparison of the methods' efficiency, a unique series of 100 ROIs is used for all the different cases in each image set.

Table 1: The performance of CT-MRI alignment for head and thorax images. Mean alignment error is the Euclidian distance between gold standard optimal shift and the obtained shift. Correctly aligned percent are those registered with error of less than 3 pixels Results of [1] are also provided in ( ).

	IR		GRAD(+IR)		PC(+IR)		PC+GRAD(+IR)	
	Mean error in terms of pixels	Correctly aligned %	Mean error in terms of pixels	Correctly aligned %	Mean error in terms of pixels	Correctly aligned %	Mean error in terms of pixels	Correctly aligned %
Head								
CT to pd	**11.4(80.7)**	**13(30)**	**1.48**	**100**	**1.4(1.4)**	**100(100)**	**1.47**	**100**
CT to t1	**9.8(111)**	**12(0)**	**17.99**	**80**	**2.93(10.7)**	**76(90)**	**3.68**	**80**
CT to t2	**22.8(92.4)**	**1(15)**	**2.75**	**98**	**1.7(9.7)**	**100(85)**	**1.71**	**100**
Thorax								
CT to pd	**12.0(23.2)**	**74(37.5)**	**5.9**	**60**	**2.4(4.5)**	**86(57.5)**	**2.6**	**72**
CT to t1	**31.4(47.6)**	**44(35)**	**24.8**	**66**	**15.3(20.7)**	**66(62.5)**	**19.7**	**66**
CT to t2	**45.9(73.45)**	**32(17.5)**	**35.3**	**54**	**24.7(50.8)**	**64(30)**	**22.5**	**60**

4) Combined PC-GRAD cost function.

All the implementations are performed in MATLAB.

3 Results

Numerical results for the efficiency of the registration methods are shown in Table 1. Numbers in parenthesis are the ones obtained in [1] and are provided for comparison. The first column of each method corresponds to the mean alignment error and the second column is the correctly aligned percent. Greater correctly aligned rate means method is more "ROI-independent". Besides Table 1, results of color-fusion are also available in Figs 1and 2.

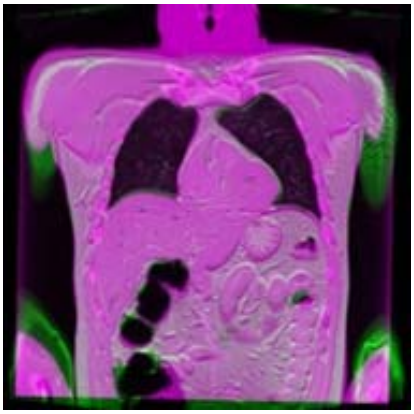

Fig. 2: Fused-image of a proton density weighted MRI image and a CT image of thorax.

Fig. 3 is also an example where all the methods failed to register the images correctly, this condition happens when the ROI lacks enough structural features, which is obvious in Fig 3.

4 Discussion and Conclusions

In this paper a comparison of different SSD-based methods for multimodal medical images is performed. After reviewing the required theories, previous methods and also the modified proposed method are applied on two sets of data and the results are provided in Table 1.

In a view of Table 1 again, some differences between obtained results and previous results from [1] is clearly seen. This difference is due to the rotation-management that we used [2]. Thus some improvements in results of "IR" and "PC" is made, especially in the field of "mean alignment error".

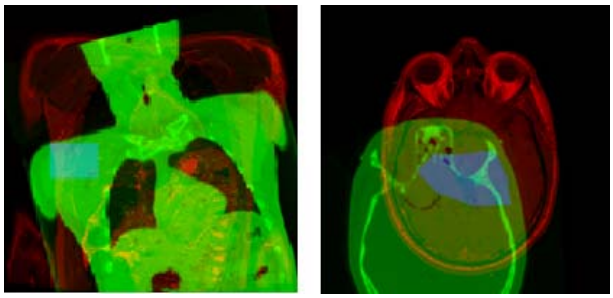

Fig. 3: Examples of conditions where registration methods failed, the corresponding ROI are depicted too.

For head images, we see that the "GRAD" results are comparable to "PC" results, and since "PC" is more time-consuming, "GRAD" may be preferred. However if computational load of registration is of less importance, then the new proposed method ("PC+GRAD") is completely promising and it contains both the advantages of the PC and GRAD methods (high correctly aligned rate percent and low mean alignment error). For the thorax datasets, it seems that the combined cost function does not work properly and PC alone works better. It is also

observable that "IR" somehow works well for alignment of CT and pd-weighted MRI, which is due to similarity of intensity in this image pair.
For image fusion, in this paper we only visualized the results with color fusion; other method such as using mathematical or logical operation (+,-, AND, OR,…) could also be used to amplify the desired feature.
Another suggestion for future trends in image fusion is to wavelet-decompose the registered images into "approximation" and "detail" coefficients, perform a weighted summation on approximation terms and detail terms, and then reconstruct the image by inverse wavelet transform. This way there is direct control on the contribution of each image to approximation and detail.

Acknowledgements

The authors would like to thank the National Library of Medicine for the Visual Human Project data, and also Mr. A. Wong for his useful comments.

References

[1] A. Wong, and W. Bishop, “Efficient least squares fusion of MRI and CT images using a phase congruency model,” *Pattern Recognition Letters*, vol. 29, pp. 173–180, 2008.
[2] J. Orchard, "Efficient least squares multimodal registration with a globally exhaustive alignment search," *IEEE Trans. Image Process.* vol. 16, No. 10, pp. 2526–2434, Oct. 2007.
[3] J. Orchard, "Efficient global weighted least-squares translation registration in the frequency domain," *Proc. ICIAR Lecture Notes in Computer science*, vol. 3656, pp. 116–124, Sep. 2005.
[4] P. Kovesi, "Phase congruency detects corners and edges," *Proc. Australian Pattern Recognition Soc. Conf.*, pp. 309–318, 2003.
[5] P. Kovesi, "Image features from phase congruency," *Videre: A J. Comput. Vision Res.* Vol. 1, no. 3, pp. 2–26, 1999.
[6] Y. Kirankumar, and S. Shenbaga Devi, "Transform-based medical image fusion," *Int. J. Biomedical Engineering and Technology,* Vol. 1, No. 1, pp.101–110, 2007.