Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 29 (2008) 173–180
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Efficient least squares fusion of MRI and CT images using a phase congruency model

Alexander Wong *, William Bishop

Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada

Received 5 February 2007; received in revised form 7 August 2007
Available online 11 October 2007

Communicated by F.Y. Shih

Abstract

This paper presents an efficient and robust approach for MRI–CT image fusion using a phase congruency model. The fast fourier transform (FFT) is used to efficiently evaluate the similarity cost function. This approach is largely invariant to pixel intensity mappings.

Keywords: Least squares fusion; Multi-modal; MRI; CT; Medical imaging; Phase congruency

1. Introduction

Medical imaging has become increasingly important in medical analysis and diagnosis. Different medical imaging techniques such as X-rays, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) provide different perspectives on the human body that are important in the diagnosis of diseases or physical disorders. For example, CT scans provide high-resolution information on bone structure while MRI scans provide detailed information on tissue types within the body. Therefore, an improved understanding of a patient's condition can be achieved through the use of different imaging modalities.

A powerful technique used in medical imaging analysis is medical image fusion, where streams of information from medical images of different modalities are combined into a single fused image. The effectiveness of medical image fusion can be illustrated by the fused image of an axial slice through the head using a PD-weighted MRI scan and a CT scan shown in Fig. 1. Both the bone structure (shown in magenta) and tissue structure (shown in green) can be clearly identified in the single image. Therefore, image fusion allows a physician to obtain a better visualization of the patient's overall condition.

To perform fusion between medical images, it is first necessary to align the images with each other. A number of techniques have been proposed for the purpose of image alignment. These methods are often categorized based on the similarity metric used to perform the comparison between images. These techniques include the following:

1. Cross-correlation (Capel and Zisserman, 1998; Cideciyan, 1995; Solaiyappan and Gupta, 2000),
2. Sum of squared differences (SSD) (Fitch et al., 2005; Lucas and Kanade, 1981; Orchard, 2005),
3. Mutual information (Maes et al., 1997; Shekhar and Zagrodsky, 2002), and
4. Phase correlation (Averbuch and Keller, 2002; Heng et al., 2000; Reddy and Chatterji, 1996).

In the general case, the successful alignment of two images may require one of the images to undergo a series of geometric transformations. Scaling, rotating, and shifting one of the images is likely to be necessary to bring the images into alignment. In the context of medical

* Corresponding author. Tel.: +1 647 280 1947; fax: +1 519 746 3077.
E-mail addresses: a28wong@engmail.uwaterloo.ca (A. Wong), wdbishop@uwaterloo.ca (W. Bishop).

0167-8655/$ - see front matter
doi:10.1016/j.patrec.2007.08.018

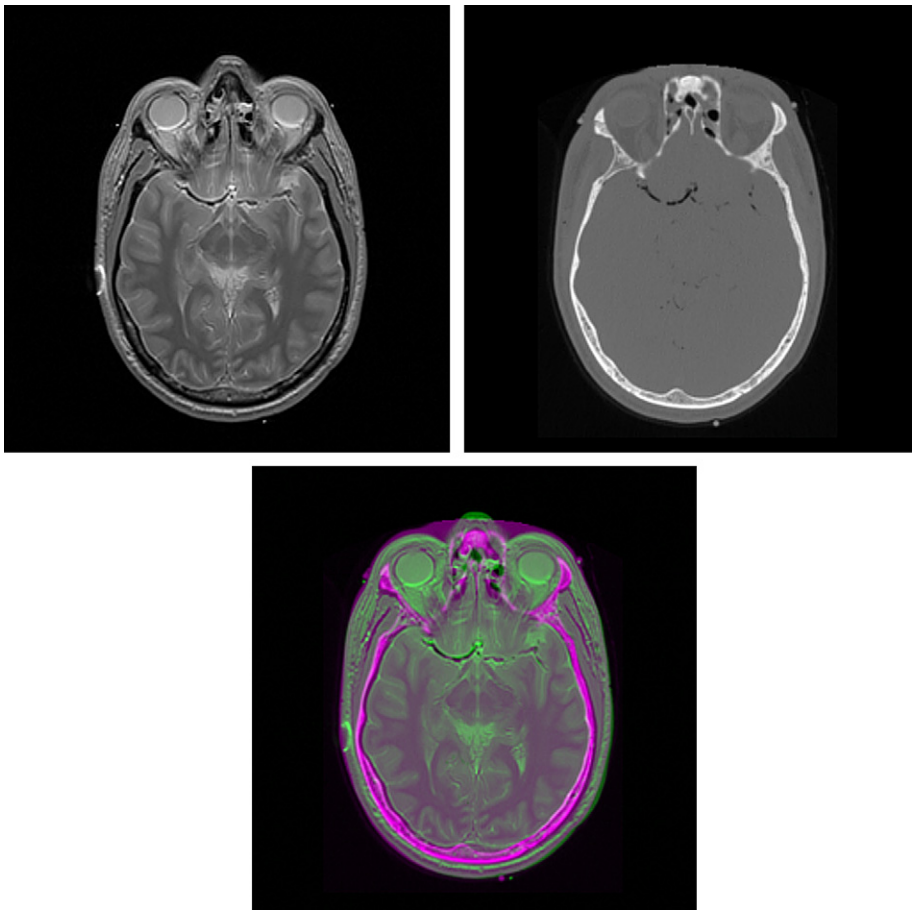

Fig. 1. An axial cranial slice: (a) Top-left image: PD-weighted MRI scan, (b) Top-right image: CT scan, and (c) Bottom image: fused image.

imaging, the scale of an image is typically known for a particular medical imaging device. Given two medical images of different scales and known sources, it is possible to perform scaling compensation on one of the images to create a set of images with consistent scale. The problem of scaling medical images is substantially easier to solve than the problem of scaling arbitrary images since the scale is known a priori.

The task of determining the near optimal rotation required to align two images is a challenging one. Given a small range of possible rotations, it is possible to exhaustively search the range for the rotation that delivers the best possible alignment. In the context of medical imaging, a tractable approach involves searching the very small range of possible rotations. This exhaustive approach for determining the near optimal rotation relies upon the existence of an efficient technique for determining the near optimal shift between two medical images. Therefore, this paper focuses on the determination of a near optimal shift between two medical images of different modalities.

Recently, it has been shown that image alignment using a weighted SSD cost function can be performed efficiently through the use of the fast fourier transform (FFT) (Fitch et al., 2005; Orchard, 2005). It has been demonstrated that the performance of this approach is 60–500 times faster than directly evaluating the SSD cost function (Orchard, 2005). One major issue with these approaches for image fusion is that they are unsuitable for aligning images with very different intensity mappings. The algorithm proposed by Orchard (2005) attempts to reduce the effects of varying intensity mappings through the use of linear intensity remapping. However, this approach fails in the case of aligning MRI images with CT images because there is often no appropriate intensity mapping that can make them appear similar in terms of pixel intensities. One approach to this problem is to discard the use of pixel intensities as the feature space and instead exploit the structural characteristics of the image content. The goal of this paper is to extend this efficient alignment method to the problem of MRI–CT image fusion by utilizing a phase congruency model that takes advantage of the important structural characteristics of image content.

The main contribution of this paper is an efficient image fusion algorithm based on a phase congruency model. The proposed image fusion algorithm attempts to find a near optimal shift that aligns two medical images of different modalities. This approach is invariant to pixel intensity mappings and it is suitable for the fusion of MRI and CT images. In this paper, the theory underlying weighted SSD cost function evaluation using the FFT and the phase

congruency model is presented in Section 2. The proposed image fusion algorithm is described in Section 3. The testing methods and test data are presented in Section 4. Experimental results are discussed in Section 5 and conclusions are drawn in Section 6.

2. Proposed fusion method using phase congruency

Before outlining the proposed fusion algorithm, it is important to examine the theory behind the key concepts of the algorithm. First, weighted cost function evaluation using the FFT framework is described in Section 2.1. Then, the concept underlying the phase congruency model is described and explained to illustrate how an intensity–invariant feature space can be derived for the purposes of aligning MRI scans with CT scans in Section 2.2. For the purpose of determining the near optimal shift, it is assumed that all medical images have comparable scale and orientation. A brief discussion of how to determine the near optimal rotation necessary to produce medical images with a comparable orientation is presented in Section 2.3.

2.1. Weighted cost evaluation using the FFT

For two images f and g, the weighted SSD cost function between image g and image f shifted by a is given by:

$$E(a) = \sum_x (f(x-a) - g(x))^2 w(x) \tag{1}$$

where $w(x)$ is a weighting function over $g(x)$, x is the coordinate of a pixel in the image, and a is the shift. For the proposed algorithm, the weighting function is used to identify the region of interest in g (0 if outside the ROI and 1 if inside the ROI). A low value of E signifies a high level of image similarity. Hence, the optimal shift of f that brings it into alignment with g is the value of a that results in the minimum value of E. This can be expressed in the form of a least squares optimization problem as the following:

$$a = \arg\min_a \left[\sum_x (f(x-a) - g(x))^2 w(x) \right] \tag{2}$$

This least squares optimization problem is computationally expensive to solve directly. This is because the SSD cost function must be evaluated for every possible shift of f. A more efficient approach to solving this problem is to formulate the dual of the problem such that it can be solved using the FFT in a computationally efficient manner. As shown in previous research (Fitch et al., 2005; Orchard, 2005), (1) can be rewritten as follows:

$$E(a) = \sum_x f^2(x-a)w(x) - 2\sum_x f(x-a)g(x)w(x) + \sum_x g^2(x)w(x). \tag{3}$$

Note that the last term in (3) does not depend on the value of a so this value can be removed from the equation. Reformulating the first two terms into convolutions results in the following cost function:

$$E(a) = \{\bar{f}^2(x) * w(x)\}_a - 2\{\bar{f}(x) * (g(x)w(x))\}_a \tag{4}$$

where the $*$ operator denotes convolution and $\bar{f}(x) = f(-x)$. Typically, convolutions are computationally expensive if performed in a direct fashion. However, it is important to realize that convolutions in the spatial domain become multiplications in the frequency domain. Multiplications are significantly faster to evaluate. Therefore, (4) can be rewritten as the following:

$$E(a) = F^{-1}\{F(\bar{f}^2(x))F(w(x))\}(a) - 2F^{-1}\{F(\bar{f}(x))F(g(x)w(x))\}(a) \tag{5}$$

where F and F^{-1} represent FFT and inverse FFT (IFFT), respectively. The use of FFT allows the value of E for all possible values of a to be evaluated simultaneously in a very efficient manner.

2.2. Phase congruency model

The main problem with previous approaches is that they make use of pixel intensities for determining the similarity between images. While the use of pixel intensities works well for images acquired using the same device or modality, it is unsuitable for comparing images captured with different modalities. The reason is that images captured with different modalities can possess significantly different pixel intensity mappings, even if the target content is the same. This is especially true when comparing MRI scans and CT scans, which respond differently to tissue and bone structures. Therefore, algorithms must utilize features that do not rely on pixel intensities when comparing MRI scans with CT scans.

Using the perception characteristics of the human vision system, an alternative approach to the task of fusing MRI scans and CT scans is apparent. When comparing the similarity between two objects, the human vision system relies on the significant structural characteristics of the objects. Therefore, two identical objects with very different color patterns can be identified as similar objects based on structural characteristics alone. The efficient alignment framework described in Section 2.1 can be extended to make use of significant structural characteristics as a feature space for comparing MRI scans with CT scans. To extract these significant structural characteristics as a feature space, phase congruency can be utilized.

The phase congruency approach to feature perception is based on the Local Energy Model (Morrone and Owens, 1987), which postulates that significant features can be found at points in an image where the Fourier components are maximally in phase. Furthermore, the angle at which phase congruency occurs signifies the feature type. The phase congruency approach to feature perception has been

used for feature detection (Kovesi, 1999, 2003) and perceptual blur estimation (Wang and Simoncelli, 2004). The phase congruency model used in the proposed algorithm is based on the one presented by Kovesi (2003) which was designed to provide good feature localization and noise compensation. First, logarithmic Gabor filter banks at different discrete orientations are applied to the 2D image signal in the spatial frequency domain and the local amplitude and phase at a point (x, y) in a signal are obtained. The phase congruency, $P(x, y, \theta)$, is then calculated for each orientation θ using the following measure:

$$P(x,y,\theta) = \frac{\sum_n W(x,y,\theta)\lfloor A_n(x,y,\theta)(\cos(\phi_n(x,y,\theta) - \bar{\phi}(x,y,\theta)) - |\sin(\phi_n(x,y,\theta) - \bar{\phi}(x,y,\theta))|) - T\rfloor}{\sum_n A_n(x,y,\theta) + \varepsilon} \tag{6}$$

where (x,y) is the point in the image, θ is the orientation, $W(x, y, \theta)$ is the weighting factor based on frequency spread, $A_n(x, y, \theta)$ and $\phi_n(x, y, \theta)$ are the amplitude and phase for wavelet scale n, respectively, $\bar{\phi}(x, y, \theta)$ is the weighted mean phase, T is a noise threshold constant and ε is a small constant value to avoid division by zero. A measure of feature significance can then be calculated using the principle moments of phase congruency. For the purpose of the algorithm, the minimum moments, $M(x, y)$, are calculated using the following equation:

$$M(x,y) = \frac{1}{2}\left(\sum(P_s(\theta))^2 + \sum(P_c(\theta))^2 - \sqrt{4\left(\sum(P_s(\theta))(P_c(\theta))\right)^2 + \left(\sum(P_c(\theta))^2 - \sum(P_s(\theta))^2\right)^2}\right) \tag{7}$$

where $P(x, y, \theta)$ is the phase congruency at point (x, y) for a given orientation θ, $P_s(\theta)$ is $P(x, y, \theta)\sin(\theta)$, and $P_c(\theta)$ is $P(x, y\,\theta)\cos(\theta)$. The minimum moments are a good indicator of significant structural characteristics, where moments with higher values indicate higher structural feature significance. Two examples of matching significant features with high minimum moment values are illustrated in Fig. 2. Note that the feature indicated by the dotted line in the 1st image is clearly discernable from its surrounding region by the change in contrast. Although the same feature indicated by the dotted line in the 2nd image lacks a significant contrast with the surrounding region, it still possesses a relatively high minimum moment value. As such, it is deemed to be a significant feature. This example demonstrates that the moments are invariant to intensity mappings. Therefore, the minimum moment map of an image can be used as an effective feature space for comparing similarity between images with very different pixel intensity mappings.

Based on the above model, the new weighted SSD cost function becomes:

$$E(\vec{a}) = \sum_{\vec{x}} (M_f(\vec{x} - \vec{a}) - M_g(\vec{x}))^2 w(\vec{x}) \tag{8}$$

where M_f and M_g are the minimum moments of images f and g, respectively. Since the minimum moments are invariant to the pixel intensity mappings, linear remapping

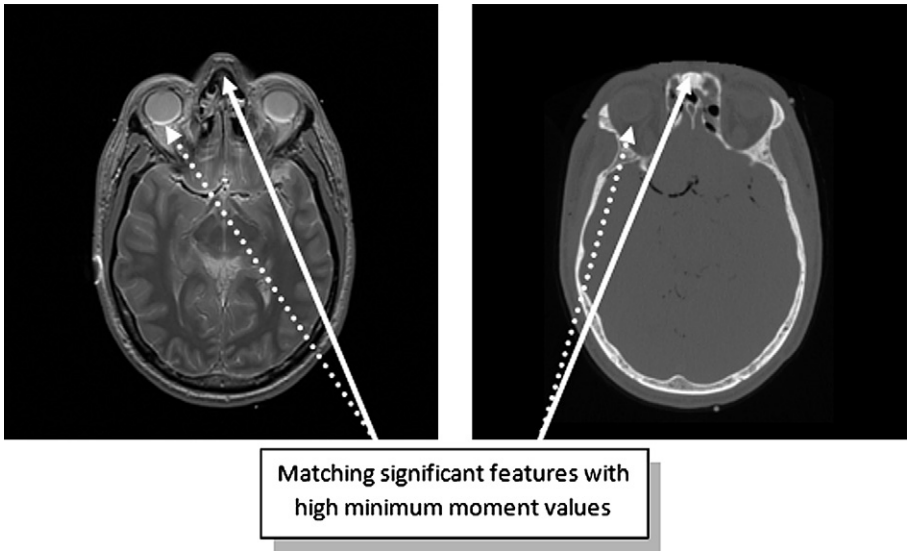

Fig. 2. Matching significant features with high minimum moment values.

can be performed to make the corresponding portions of the minimum moment maps of the two images as similar as possible. This is accomplished by substituting M_f with $sM_f + t$ to form the following least squares optimization problem:

$$\vec{a}, s, t = \arg\min_{\vec{a},s,t} \left[\sum_{\vec{x}} (sM_f(\vec{x} - \vec{a}) + t - M_g(\vec{x}))^2 w(\vec{x}) \right] \tag{9}$$

Table 1
Image alignment performance results

Test Set	MI		Linear		PC	
	Mean alignment error[a]	Images aligned correctly (%)[b]	Mean alignment error[a]	Images aligned correctly (%)[b]	Mean alignment error[a]	Images aligned correctly[b] (%)
THORAX						
PD to CT	54.43	45.00	23.20	37.50	4.52	57.50
T1 to CT	66.22	35.00	47.60	35.00	20.7	62.50
T2 to CT	64.62	30.00	73.45	17.50	50.85	30.00
HEAD						
PD to CT	32.34	50.00	80.73	30.00	1.40	100.00
T1 to CT	26.27	60.00	111.64	0.00	10.75	90.00
T2 to CT	27.92	60.00	92.36	15.00	9.67	85.00
PELVIS						
T1 to CT	26.10	65.00	17.36	65.00	3.24	70.00
Overall	42.56	49.29	63.76	28.57	14.45	70.71

[a] The mean alignment errors are computed over 40 test alignments for each alignment pair.
[b] Images are aligned correctly if the alignment error is less than or equal to 3 pixels. The alignment error is computed over 40 test alignments for each alignment pair.

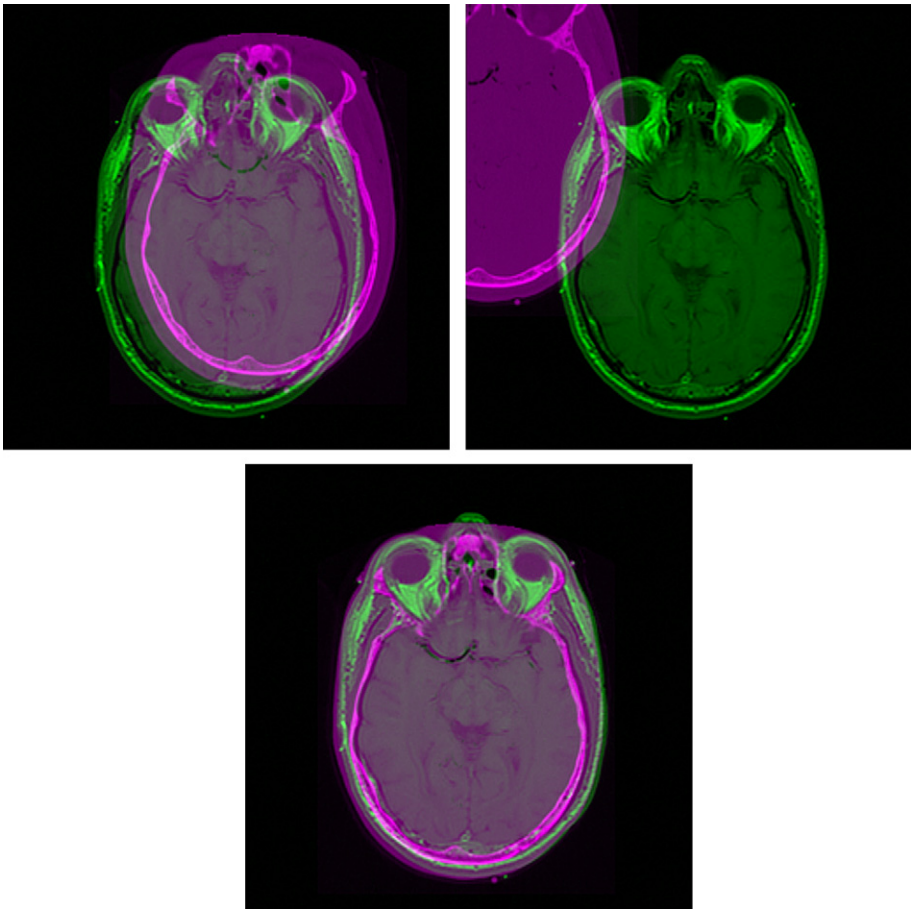

Fig. 3. An axial cranial slice: (a) Top-left image: fused image using MI, (b) Top-right image: fused image using LINEAR, and (c) Bottom image: fused image using PC.

The final cost function evaluation using the FFT becomes the following:

$$E(\vec{a},s,t) = s^2F^{-1}\{F(\overline{M}_f^2(\vec{x}))F(w(\vec{x}))\}(\vec{a}) - 2sF^{-1}\{F(\overline{M}_f(\vec{x}))F(M_g(\vec{x})w(\vec{x}))\}(\vec{a}) + 2stF^{-1}\{F(\overline{M}_f(\vec{x}))F(w(\vec{x}))\}(\vec{a}) \quad (10)$$

2.3. Determining the near optimal rotation for medical image alignment

The near optimal rotation for medical image alignment can be determined using the proposed cost function for evaluating the near optimal shift. Due to the nature of medical imaging equipment, medical images typically do not exhibit large rotations. Therefore, it is reasonable to conduct an exhaustive search within a small range of rotations to find the near optimal rotation. The problem of finding the near optimal rotation (θ_{no}) can be expressed mathematically as follows:

$$\theta_{no} = \arg\min_{\theta}[E_{\min}(\theta)], \quad \text{where } \theta \in [\theta_{\min}, \theta_{\max}] \quad (11)$$

where $E_{\min}(\theta)$ represents the lowest cost shift for each value of θ in the acceptable range. An exhaustive search can be conducted in linear time using the proposed cost function. If the range of θ is well constrained by $\theta_{\min}$ and $\theta_{\max}$, the time required by an exhaustive search is reasonable for most practical applications.

3. Algorithm outline

Utilizing the theory presented, the fusion algorithm consists of the following steps:

1. Given images f and g, compute the minimum moment maps M_f and M_g using Eqs. (6) and (7) as described in Section 2.2.
2. Perform weighted SSD cost evaluation using the FFT using Eq. (10) as described in Section 2.2 to determine the optimal shift between f and g.
3. Align f and g using the optimal shift and fuse the two images into a single image using techniques such as color fusion (where each image is assigned to a specific color band to create a fused color image) or fusion using mathematical operators (e.g. Addition, Subtraction, Logical AND, Logical OR, etc.).

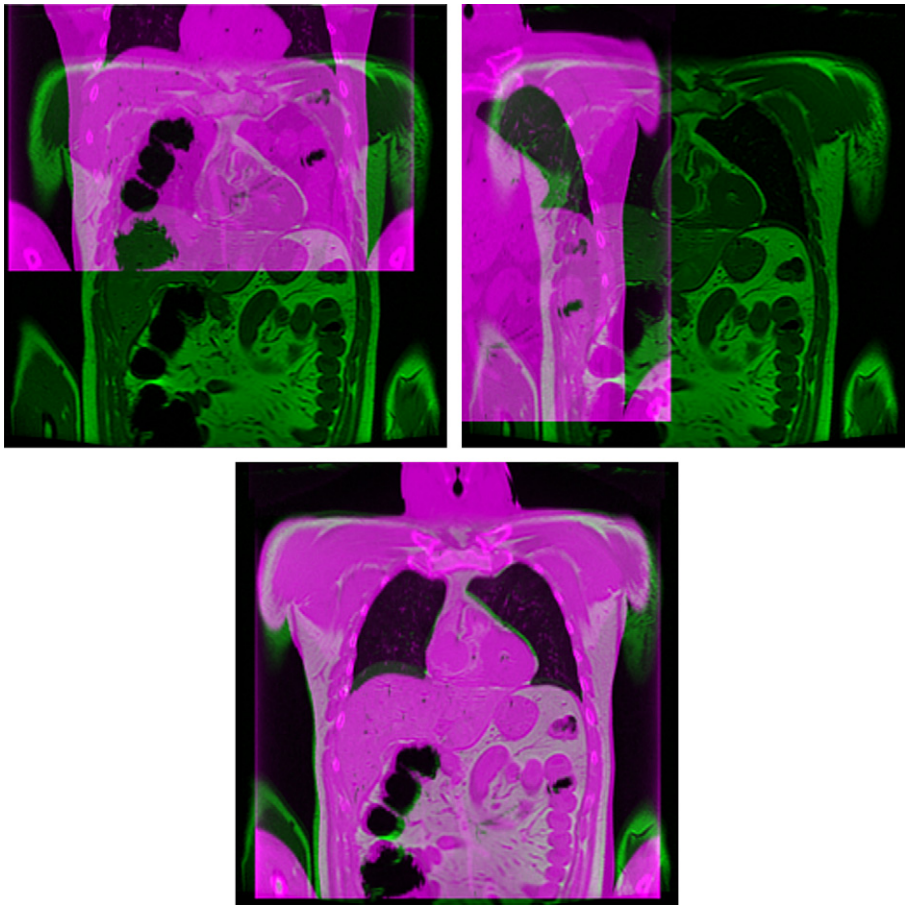

Fig. 4. A coronal thorax slice: (a) Top-left image: fused image using MI, (b) Top-right image: fused image using LINEAR, and (c) Bottom image: fused image using PC.

4. Testing methods

The fusion algorithm was implemented in MATLAB using a color fusion technique. The algorithm was tested using two sets of images derived from the Visible Male and Female datasets of the National Library of Medicine's Visible Human Project. All of the test images used were 8-bit grayscale images of size 256×256. The THORAX and HEAD test sets consisted of T1-weighted MRI scans, T2-weighted MRI scans, PD-weighted MRI scans, and CT scans. The PELVIS test set consisted of T1-weighted MRI scans and CT scans. A description of each test set is given below:

- ***THORAX:*** A coronal slice through the thorax.
- ***HEAD:*** An axial slice through the head.
- ***PELVIS:*** An axial slice through the pelvis.

To judge the performance of the fusion methods, each MRI image was fused with the corresponding CT image and vice-versa in the same test set over a set of 40 randomly generated regions of interest (ROI) using three methods:

- ***MI:*** mutual information using pixel intensities.
- ***LINEAR:*** weighted SSD evaluation with linear remapping using pixel intensities (Orchard, 2005).
- ***PC:*** the proposed algorithm using the phase congruency model with linear remapping.

The mean alignment error and the percentage of image alignments with alignment errors less than or equal to 3 pixels (1.17% of the image dimension) were calculated. The alignment error between two images was calculated as the Euclidean distance from the estimated optimal shift to the gold standard optimal shift.

5. Experimental results

The experimental results are shown in Table 1. It can be observed that the proposed method achieved noticeably better alignment performance over the MI and LINEAR methods in all test cases. Examples of the fused images achieved using the three methods are shown in Figs. 3–5. The proposed algorithm outperformed the other two methods due to the fact that the intensity mappings of the MRI and CT scans are significantly different. The

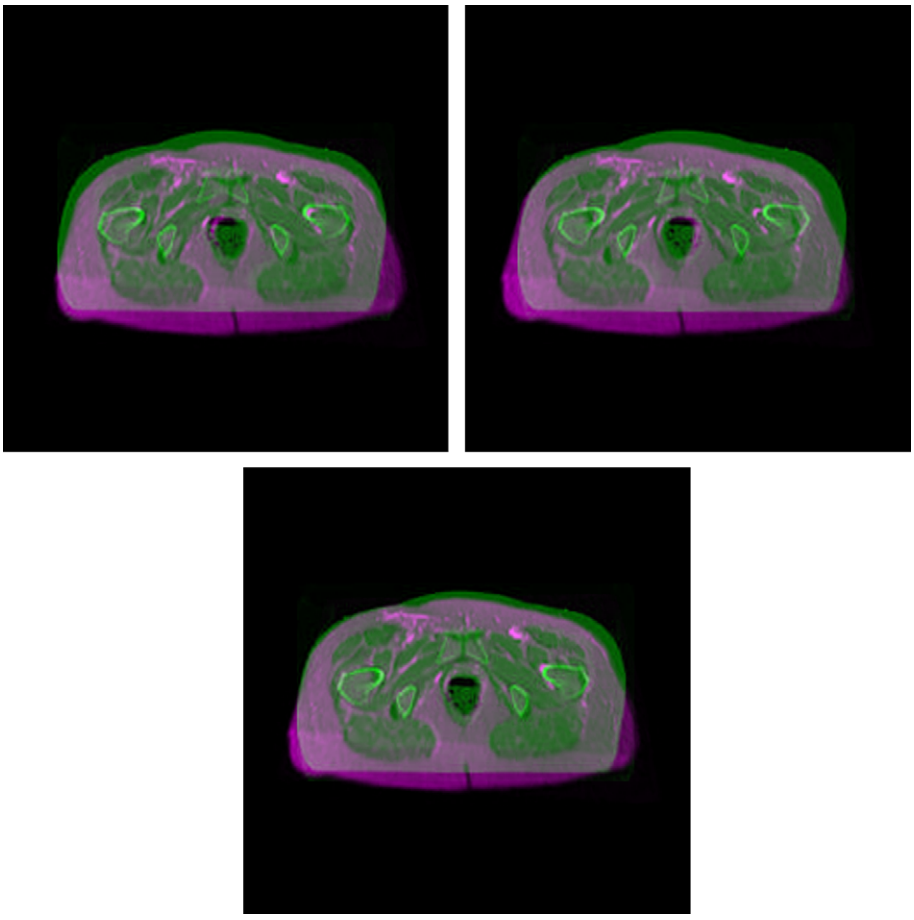

Fig. 5. An axial pelvis slice: (a) Top-left image: fused image using MI, (b) Top-right image: fused image using LINEAR, and (c) Bottom image: fused image using PC.

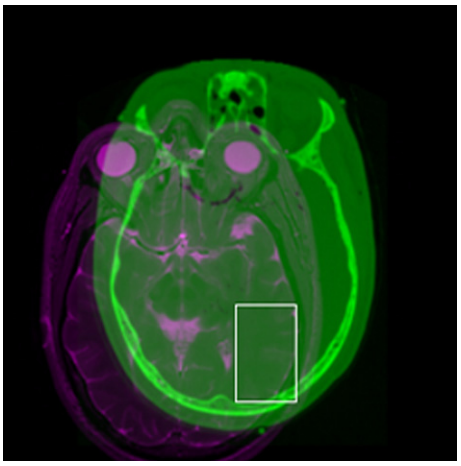

Fig. 6. An example of a misalignment produced using the proposed fusion algorithm.

two intensity-based methods (MI and LINEAR) are unable to find a good correspondence. However, the proposed algorithm relies on the significant structural characteristics that are mostly present and consistent between the MRI scans and the CT scans so the proposed algorithm is able to find a good correspondence. The experimental results demonstrate the effectiveness of the proposed algorithm for fusing MRI images with CT images.

An example of a misalignment using the proposed method is shown in Fig. 6. In this particular case, the randomly selected region of interest for alignment was a region that did not exhibit significant structural characteristics in one of the two modalities. In the absence of significant structural characteristics, the proposed algorithm has difficulty finding a suitable alignment. It should be noted that if a human expert were given this particular alignment task, a suitable alignment would be difficult to find between the two medical images in the region of interest. To help reduce misalignments, the entire image should be used for alignment since full medical images often exhibit significant structural characteristics that may not be present in a particular region of interest.

6. Conclusions

In this paper, we have introduced a new method for efficient image fusion based on phase congruency. The method uses structural characteristics of images that are invariant to pixel intensity mappings. The method determines the near optimal shift required to align medical images of different modalities. Experimental results show that the proposed method improves alignment accuracy when compared to existing methods. This method is well suited for multi-modal medical image fusion.

Acknowledgements

The authors would like to thank the National Library of Medicine for the Visual Human Project data. This research has been sponsored in part by Epson Canada and the Natural Sciences and Engineering Research Council of Canada.

References

Averbuch, A., Keller, Y., 2002. FFT based image registration. In: Proc. IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing, 4, pp. 3608–3611.

Capel, D., Zisserman, A., 1998. Automated mosaicing with super-resolution zoom. In: Proc. IEEE Comput. Soc. Conf. on Computer Vision and Pattern Recognition. pp. 885–891.

Cideciyan, A., 1995. Registration of ocular fundus images: an algorithm using cross-correlation of triple invariant image descriptors. IEEE Eng. Med. Biol. Magazine 14 (1), 52–58.

Fitch, A., Kadyrov, A., Christmas, W., Kittler, J., 2005. Fast robust correlation. IEEE Trans. Image Process. 14 (8), 1063–1073.

Heng, L., Hwa, E., Kok, H., 2000. High accuracy registration of translated and rotated images using hierarchical method. In: Proc. IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing, 6, pp. 2211–2214.

Kovesi, P., 1999. Image features from phase congruency. Videre: A J. Comput. Vision Res. 1 (3), 2–26.

Kovesi, P., 2003. Phase congruency detects corners and edges. In: Proc. Australian Pattern Recognition Soc. Conf. pp. 309–318.

Lucas, B., Kanade, T., 1981. An interactive image registration technique with an application to stereo vision. In: Proc. Imaging Understanding Workshop. pp. 121–130.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multi-modality medical image registration by maximization of mutual information. IEEE Trans. Med. Imaging 16 (2), 187–198.

Morrone, M., Owens, R., 1987. Feature detection from local energy. Pattern Recognition Lett. 6, 303–313.

Orchard, J., 2005. Efficient global weighted least-squares translation registration in the frequency domain. In: Proc. Internat. Conf. on Image Analysis and Recognition. pp. 116–124.

Reddy, B., Chatterji, B., 1996. An FFT-based technique for translation, rotation and scale invariant image registration. IEEE Trans. Image Process. 5 (8), 1266–1271.

Shekhar, R., Zagrodsky, V., 2002. Mutual information-based rigid and nonrigid registration of ultrasound volumes. IEEE Trans. Med. Imaging 21 (1), 9–22.

Solaiyappan, M., Gupta, S., 2000. Predictive registration of cardiac MR perfusion images Using geometric invariants. In: Proc. Internat. Soc. Magnetic Resonance in Medicine, 1, pp. 37.

Wang, Z., Simoncelli, E., 2004. Local phase coherence and the perception of blur. Advances in Neural Information Processing Systems, 16.