Robust Multimodal Registration Using Local Phase-Coherence Representations

Alexander Wong · Jeff Orchard

Abstract Automatic registration of multimodal images has proven to be a difficult task. Most existing techniques have difficulty dealing with situations involving highly non-homogeneous image contrast and a small initial overlapping region between the images. This paper presents a robust multi-resolution method for registering multimodal images using local phase-coherence representations. The proposed method finds the transformation that minimizes the error residual between the local phase-coherence representations of the two multimodal images. The error residual can be minimized using a combination of efficient globally exhaustive optimization techniques and subpixel-level local optimization techniques to further improve robustness in situations with small initial overlap. The proposed method has been tested on various medical images acquired using different modalities and evaluated based on its registration accuracy. The results show that the proposed method is capable of achieving better accuracy than existing multimodal registration techniques when handling situations where image non-homogeneity and small overlapping regions exist.

Keywords Multimodal registration · Non-rigid · Local phase-coherence · Non-homogeneity

A. Wong (✉) · J. Orchard
University of Waterloo, Waterloo,
Ontario, N2L 3G1, Canada
e-mail: a28wong@uwaterloo.ca

J. Orchard
e-mail: jorchard@cs.uwaterloo.ca

1 Introduction

Image registration refers to the process of aligning images acquired of the same scene under different intrinsic or extrinsic conditions. Of particular interest is the registration of images acquired using different modalities such as magnetic resonance (MR), computed tomography (CT), and positron emission tomography (PET). This process is known as multimodal registration. The registered images can be used for improved data analysis in applications such as disease diagnostics and computer-assisted surgery. In the past, the task of multimodal registration was performed manually by placing markers on the human body during the image acquisition process and then finding the geometric transformation that brings the markers into alignment. This approach to multimodal registration is very laborious and time consuming. As such, methods for registering images acquired using different modalities in an automated fashion is highly desired.

There are many important issues that make automated multimodal registration a very challenging problem to solve. First, images being acquired using different modalities (i.e., MR and CT) are captured using different imaging devices at different times. As such, the images acquired using different modalities often contain different geometric distortions that make it difficult to compare image content. Second, images acquired using different modalities are mapped to different intensity values. This makes it difficult to compare images based on their intensity values since the same content within the images may be represented by different intensity values. The problem is further complicated by the fact that various intrinsic (for MRI, static field and RF non-homogeneity [1, 2]) and extrinsic

(patient motion) sensing conditions can lead to image non-homogeneity. Such image non-homogeneities result in the same content within a single image to be represented by different intensity values. Finally, the disparity between the intensity values of multimodal images can lead to coincidental local intensity matches between non-corresponding content. These matches can result in local minima along the convergence plane that make it difficult for local optimization schemes such as conjugate gradient, Nelder-Mead simplex [3], and sequential quadratic programming (SQP) methods to converge to the correct solution, particularly when the initial overlap is small and the distance traveled during the optimization process is large. The goal of the proposed method is to address all of the these issues to deliver robust registration of multimodal images.

The main contribution of this paper is a novel approach to the task of automated registration of multimodal images. The proposed method constructs local phase-coherence representations of the images that are then aligned automatically in a coarse manner using efficient globally exhaustive optimization techniques. The resulting alignment is then refined on a sub-pixel level using a multi-resolution local optimization approach. The paper is organized as follows. Previous work in multimodal registration for medical images is described in Section 2. The proposed multimodal registration method is described in detail in Section 3. The testing methods and test data are described in Section 4. The experimental results are discussed in Section 5. Finally, conclusions are drawn and future work is discussed in Section 6.

2 Previous Work

One of the most widely-used groups of multimodal registration methods is mutual information (MI) and its entropy-based alternatives [4–7]. Such techniques are very powerful and considered to be state-of-the-art. An extensive survey of MI-based methods can be found in [8]. The main advantage of MI is that it allows for the direct intensity-based comparison of multimodal images, making no underlying assumptions regarding the intensity relationships between the images under evaluation. However, since the intensity relationship between the images is relatively unconstrained by MI, this cost function can be highly non-monotonic with many local maxima. Some causes that can lead to high non-monotonicity include sampling (i.e., number of histogram bins and image resolution) and interpolation effects. Since local optimization schemes are dependent on the monotonicity of the underlying cost function,

such schemes can often get trapped in local maxima. This is particularly problematic in situations with small initial overlaps, where the optimization scheme must travel a long distance to the correct solution. This need for careful initialization of the system to achieve proper registration is one of the drawbacks to the use of MI.

A related approach was proposed by Mellor et al. [9], which utilizes local monogenic phase [10] directly as the basis for MI. The claim of the authors is that the use of local monogenic phase reduces the effects of signal non-homogeneity and therefore improves the performance of MI. However, there are several major drawbacks to this approach. While local monogenic phase provides structural information regarding the image, it provides no indication on the significance of these characteristics. This makes it sensitive to image artifacts and background degradation. It does not account for image noise and so is highly susceptible to misalignment due to noise. Finally, since local phase does not provide features that are directly comparable between images acquired using different modalities, it is highly reliant on MI to establish feature correspondences and thus shares all of the problems faced by other MI techniques. It is very important to make the distinction that the proposed method is based on local phase-coherence characteristics, which are features derived from local phase and thus does use local phase directly during the registration process. Therefore, it is not as affected by the same problems faced by local phase and Mellor's method. These issues will be demonstrated later in the paper to demonstrate the benefits of the proposed method over Mellor's method.

A popular alternative to entropy-based methods are feature-based methods that utilize intensity gradients [11–13]. These methods attempt to capture structural information based on spatial variations in image intensity. There are several advantages to this approach. First, the cost functions for such methods are more constrained and generally monotonic in nature. In fact, it is common to employ monomodal cost functions to compare intensity gradient information. Local optimization schemes for monomodal cost functions are less likely to get trapped in local minima. Furthermore, methods exist to decouple the motion parameters and allow for efficient exhaustive evaluation of some cost functions [13]. The main problem with these methods (particularly those based on gradient magnitude) is that they are highly sensitive to image non-homogeneity. Methods based on gradient fields [11, 12] are less sensitive to image non-homogeneity (though they can still be greatly affected), but are highly sensitive to noise. Furthermore, gradient field methods treat all content equally and are therefore highly sensitive to struc-

tural disparities caused by insignificant structures. Related approaches were proposed by Liu et al. and Hemmendorff et al., which are based on local phase gradients [14, 15]. There are several major drawbacks to these approaches. These methods do not make use of local frequency information from multiple scales or orientations. Also, the computation of local phase gradients is dependent on maximum local magnitude, making it sensitive to image non-homogeneity. Finally, it does not account for image noise.

We recently proposed the use of local phase-coherence (described later) as a possible cost function for multimodal registration [16]. However, issues regarding image contrast non-homogeneity and non-rigid transformations were not addressed in our earlier work. Our earlier work also did not provide registration performance at a sub-pixel level as our new method does. Furthermore, the monotonicity of the local phase-coherence cost function was not analyzed or compared to other popular multimodal registration techniques used for medical images. In particular, we demonstrate in this paper that an extension of our earlier work greatly improves its monotonicity, making it much more suitable for optimization than MI based techniques. We report on a set of experiments that demonstrate the capabilities of our new method.

3 Proposed Method

The proposed method is a multi-stage technique that utilizes various concepts to address the issues associated with multimodal registration of medical images. First, a local phase-coherence representation is extracted for each of the images under evaluation. The local phase-coherence representations can then be used to align the images on a pixel level using an efficient globally exhaustive registration algorithm. Finally, the alignment of the images can then be refined on a sub-pixel level using a local optimization algorithm.

3.1 Image Representation Using Local Phase-Coherence

Images acquired using different modalities have different intensity mappings for the same content, making them difficult to compare in a direct manner. One possible solution to this problem is to construct representations of the multimodal images that allow for comparison independent of the underlying image intensity. A common approach to an intensity-independent representation of the images is based on its structural characteristics. Two images obtained of the same scene using different modalities can have significantly different intensity characteristics but should have very similar structural characteristics. Therefore, the structural characteristics of the disparate images can be compared in a direct fashion. While previous attempts have been made to develop structural representations using gradient fields, these methods are sensitive to image non-homogeneity and can lead to incorrect alignments. Therefore, a structural representation that is relatively independent of image non-homogeneity is desired.

In recent years, algorithms such as that proposed by Mellor et al. [9] have used local phase directly as the representation basis for comparison. The key advantage to the use of local phase representations over other representations is the fact that it is largely independent of intensity. Thus, it is highly robust to the presence of signal non-homogeneities. Despite this very important advantage, there are a number of issues associated with using local phase directly in multimodal registration. First, the local phase representations used in current techniques are based on a single scale. Important local frequency information from other scales are not truly utilized. Secondly, these local phase representations provide no information on structural significance, thus placing no importance on significant structural characteristics. The local phase representations used in current methods do not account for image noise. Finally, local phase representations cannot be compared directly using simple cost functions such as sum of squared distances and must rely on more complex techniques such as MI. Given these issues, we propose an alternative approach to local phase representations by extracting features based on local phase-coherence instead of local phase. This alternative representation is able to alleviate most of the problems associated with the use of local phase in a direct manner.

The approach used in the proposed method for constructing a structural representation of the images is to use local phase relationships to identify structurally significant characteristics within an image. Local phase-coherence was recently proposed as an effective method for determining structural significance in an image [17–19]. This approach to measuring structural significance is based on the theory that the local phase-coherence peaks at locations of high perceptual significance [17]. More importantly, these points of high perceptual significance coincide with points of high structural significance within an image. Furthermore, since only phase information is used, local phase-coherence is largely independent of intensity. These properties make local phase-coherence an effective method for creating a structural representation of the images that can be evaluated in a direct fashion.

The proposed method measures local phase-coherence based on the method proposed by Kovesi [18]. Localized frequency information is obtained using complex-valued wavelets (i.e. Gabor wavelets and Dual-tree complex wavelets [20]) across multiple scales and orientations. The local phase-coherence at point $\underline{x}$ and orientation θ is computed as follows:

$$P(\underline{x}, \theta) = \frac{\sum_n W(\underline{x}, \theta) \lfloor A_n(\underline{x}, \theta) \Delta\Phi(\underline{x}, \theta) - T \rfloor}{\sum_n A_n(\underline{x}, \theta) + \varepsilon} \quad (1)$$

$$\Delta\Phi(\underline{x}, \theta) = \cos\left(\phi_n(\underline{x}, \theta) - \bar{\phi}(\underline{x}, \theta)\right)$$
$$- \left|\sin\left(\phi_n(\underline{x}, \theta) - \bar{\phi}(\underline{x}, \theta)\right)\right| \quad (2)$$

where W represents the frequency spread weighting factor, A_n and ϕ_n represent the amplitude and phase at wavelet scale n respectively, $\bar{\phi}$ represents the weighted mean phase, T represents the noise threshold and ε is a small constant used to avoid division by zero. The values of T, ε, and n used in the proposed method are 2.0, 0.01, and 4 respectively. If all the complex-valued wavelet components are in phase, the phase deviation terms ($\Delta\Phi$) go to zero and the phase-coherence goes to approximately one (if the amplitudes of the wavelet components are non-zero).

It is helpful to visualize the concept of phase-coherence graphically, as illustrated in Fig. 1. Suppose that the complex-valued wavelet components at a particular point are plotted on a polar plot in the form of a vector sum. When the wavelet components are
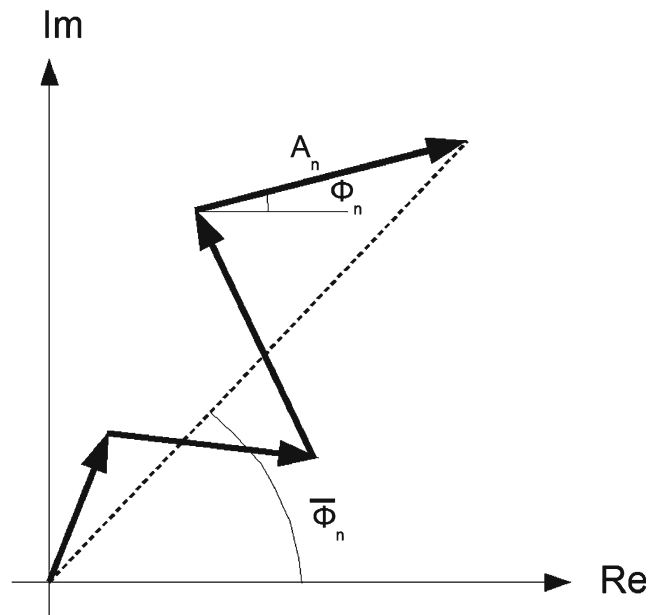
Figure 1 Simple visualization of local phase-coherence. A_n and ϕ_n represents the amplitude and phase of the nth component respectively. $\bar{\phi}$ represents the mean phase.

maximally in phase, the magnitude of the vector sum is equal to the sum of the individual amplitudes and the local phase-coherence goes to one. When the wavelet components are maximally out of phase, the local phase-coherence goes to zero.

To account for variations of local phase-coherence due to orientation, the structural significance of a point in the image can be computed as the maximum moment of phase-coherence $\mu(\underline{x})$ given by:

$$\mu(\underline{x}) = \frac{1}{2}\sum_\theta P(\underline{x}, \theta)^2 + \frac{1}{2}\sqrt{4\left(\sum_\theta \left(P(\underline{x}, \theta)\sin(\theta)\right)\left(P(\underline{x}, \theta)\cos(\theta)\right)\right)^2 + \left(\sum_\theta \left[\left(P(\underline{x}, \theta)\cos(\theta)\right)^2 - \left(P(\underline{x}, \theta)\sin(\theta)\right)^2\right]\right)^2} \quad (3)$$

where $P(\underline{x}, \theta)$ is the local phase-coherence at orientation θ. A high value of μ represents high structural significance within the image.

An issue with using maximum moments of phase-coherence directly as a structural representation is that it provides a narrow capture range, making it difficult for local optimization on a sub-pixel basis. Therefore, a mechanism is needed to propagate the influence of high moment points to nearby low moment points. In doing so, the value associated with each point would implicitly represent its spatial relationship to structurally significant image characteristics. This creates a wider capture range for local optimization on a sub-pixel basis. The proposed method uses nonlinear anisotropic diffusion

[21] to propagate the influence of high moment points to neighboring low moment points to create the final structural relationship representation used in the proposed cost function. This propagation process can be expressed as follows:

$$M(x, y, 0) = \mu(x, y) \quad (4)$$
$$M(x, y, t) = M(x, y, t-1)$$
$$+ \lambda\left(\sum_i \sum_j c(x+i, y+j, t)\nabla M(x+i, y+j, t)\right) \quad (5)$$
$$c(x, y, t) = \frac{1}{1 + \left(\frac{\nabla M(x,y,t)}{\kappa}\right)^2} \quad (6)$$

where t represents the iteration step, κ is the diffusion constant, and $\nabla M(x, y, t)$ is the moment gradient in the (x, y) direction. For the proposed method, the values t and κ are set to 5 and 20 respectively based on testing. This effectively propagates the influence of structures over a larger spatial area to create a wider capture range. This representation (as represented by M) will be referred to as the Local Phase-Coherence Relationship (LPCR) representation.

There are several benefits to using an LPCR representation for image registration. Like the local phase representation used in current methods such as that proposed by Mellor et al., the LPCR representation is robust to the effects of image non-homogeneity. This is illustrated in Fig. 2, where the LPCR representation remains consistent despite high image non-homogeneity while the intensity gradient representation is significantly degraded. The robustness of LPCR to image non-homogeneity will be further demonstrated in the experimental results.

The key benefit of the LPCR representation over the local phase representation used in current methods is that LPCR representations can be compared in a direct manner while local phase representations rely on techniques such as MI to determine feature correspondence relationships. The LPCR representation allows for a more constrained cost function than mutual-information based cost functions and therefore is less prone to being trapped in local optima. This key benefit can be demonstrated using the example shown in Fig. 3, which illustrates that the local phase representations are very different and cannot be compared in a direct manner. It can also clearly be seen that the LPCR representations of the two images are very similar and can be evaluated in a direction manner. Figure 4 shows the cost function cross sections of the proposed LPCR cost function and the neuromagnetic imaging (NMI) cost function [6] using local phase representations as proposed by Mellor et al. [9]. The cost function of NMI using local phase representation contains many local minima, while the proposed LPCR cost function remains largely smooth. Thus, the method proposed by Mellor et al. requires careful pre-alignment to achieve reasonable registration performance. The proposed method, on the other hand, does not require this type of preprocessing to achieve good results. The benefits of the proposed method will be further demonstrated in the experimental results.

3.2 Coarse-grained Globally Exhaustive Registration

Given the proposed LPCR representation, the transformation T that aligns two multimodal images f and g can be estimated using the simple optimization formulation

$$\hat{T} = \arg\min_{T} \left(\sum \left\| M_g(\underline{x}) - M_f \left(T\left(\underline{x}\right)\right)\right\| \right), \tag{7}$$

Figure 2 Example of effect of image non-homogeneity on structural representation: **a** PD-weighted MR image, **b** intensity gradient magnitude representation, **c** LPCR representation.
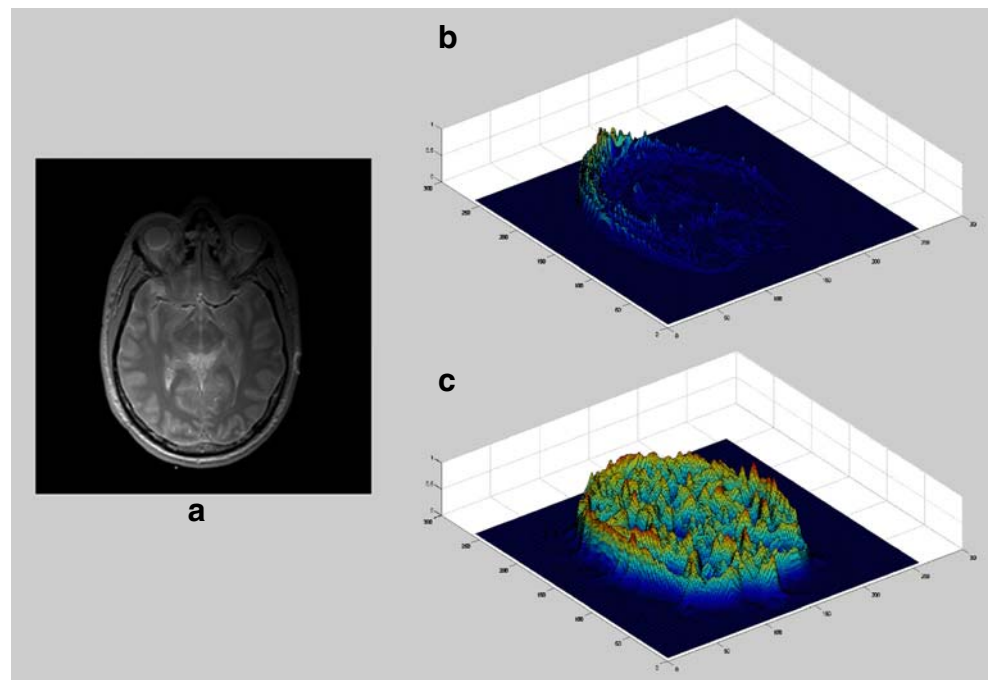
Figure 3 Example of local phase representation and LCPR representation (Courtesy of BrainWeb [26]): **a** T1-weighted MR image, **b** local phase representation of T1-weighted MR image, **c** LPCR representation of T1-weighted MR image, **d** T2-weighted MR image, **e** local phase representation of T2-weighted MR image, **f** LPCR representation of T2-weighted MR image.
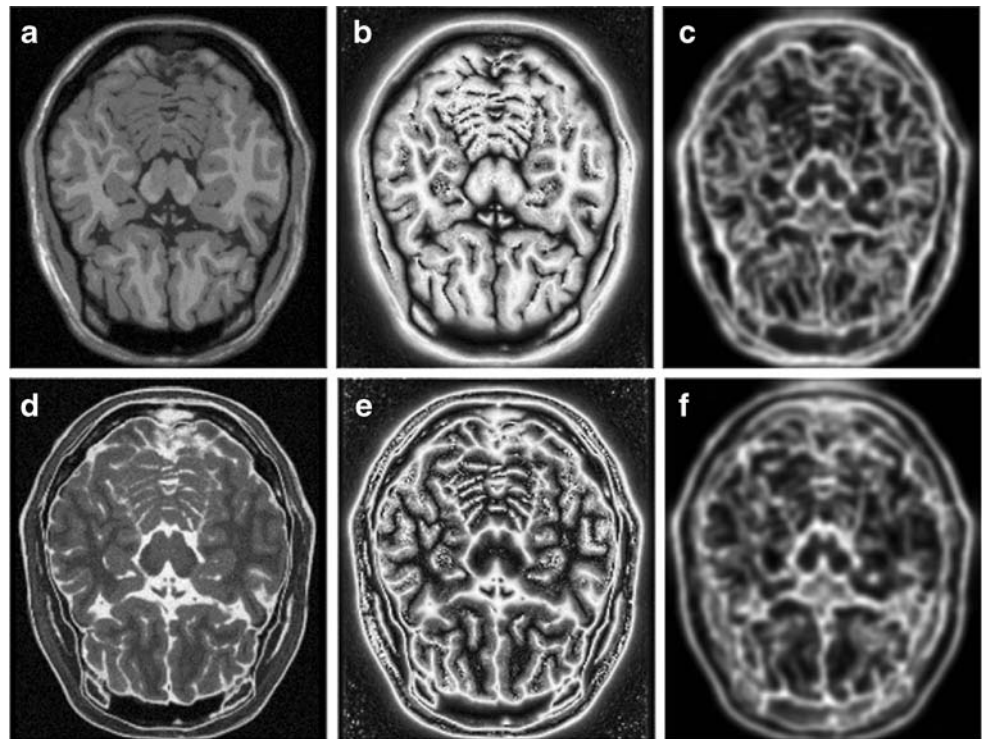

where M_f and M_g are the LPCR representations of images f and g respectively. Therefore, the correct alignment of multimodal images should be achieved when the error residual is minimized between the LPCR representations. In the current implementation of the proposed method, the sum of squared distances metric

Figure 4 Plots **a**–**b** show NMI (negated) of local phase for different translations, while plots **c**–**d** show the LPCR cost for the same translation ranges.
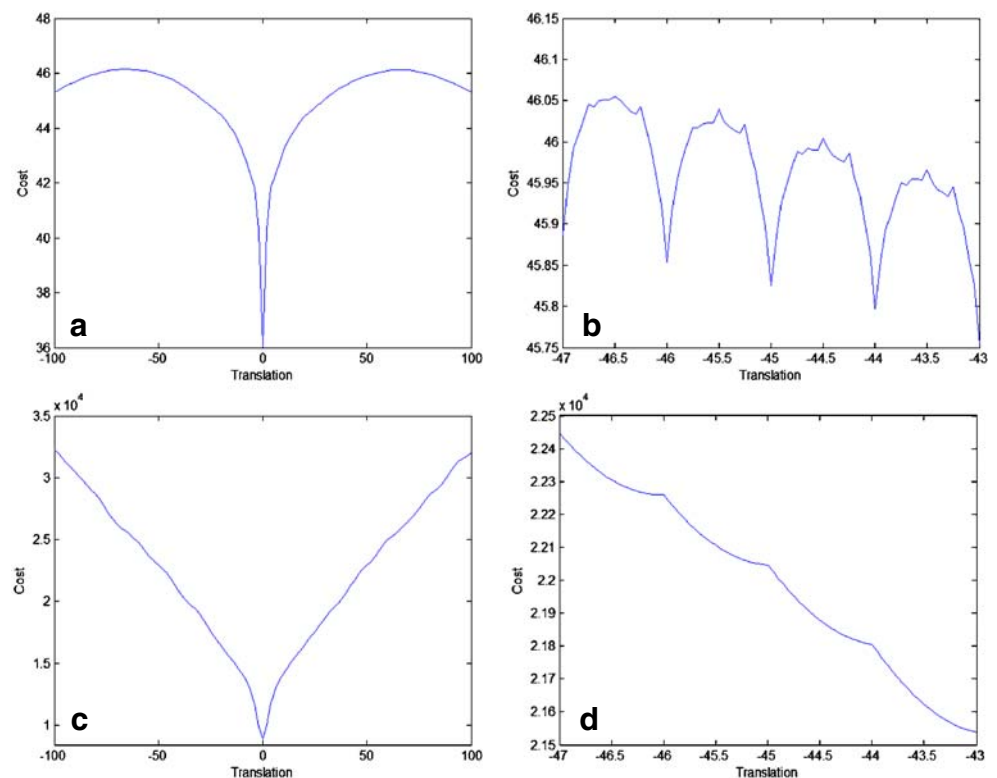
Figure 5 Test 1: Evaluation of cost over rotation: **a** PD-weighted MR image, **b** T2-weighted MR image, **c** LPCR, **d** NMI (negated), **e** correlation of gradient magnitude.
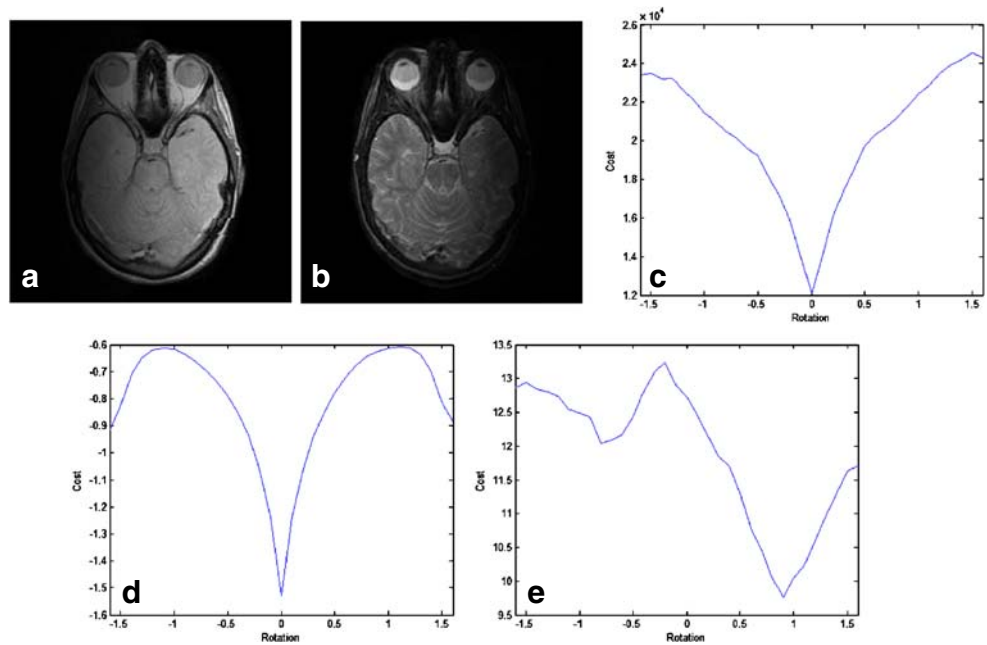
was used to compute the error residual between the LPCR representations,

$$e\,(\text{T}) = \sum_{\underline{x}} \left(M_g(\underline{x}) - M_f(\text{T}(\underline{x})) \right)^2. \tag{8}$$

One major advantage to the above formulation is that there are known globally exhaustive optimization schemes that can solve the globally optimal solution of this problem on a pixel level in a very efficient manner for rigid transformations [11, 13, 16, 22]. These global optimization schemes can be used to provide a good initial transformation estimate without worrying about local minima. This is particularly important for situations characterized by small initial overlaps, where

the distance between the initial position and correct alignment is large. A high performance globally exhaustive registration algorithm is used in the current implementation of the proposed algorithm based on that described in [16, 22]. In this approach, the sum of squared distances metric is reformulated and reduced as a series of Fourier transforms, which can be greatly accelerated through the use of Fast Fourier Transform (FFT) algorithms:

$$e(\delta, \theta) = F^{-1} \left[F\left(\overline{M_f}^2 \right) \right] (\delta)$$
$$-2F^{-1} \left[F\left(\overline{M_f} \right) F\left(\text{T}_\theta(M_g) \right) \right] (\delta) \tag{9}$$

Figure 6 Test 2: Plots **a**–**c** show the LPCR cost for different translations, while plots **d**–**f** show NMI (negated) for the same translation ranges.

Figure 6 Test 2: Plots **a**–**c** show the LPCR cost for different translations, while plots **d**–**f** show NMI (negated) for the same translation ranges.
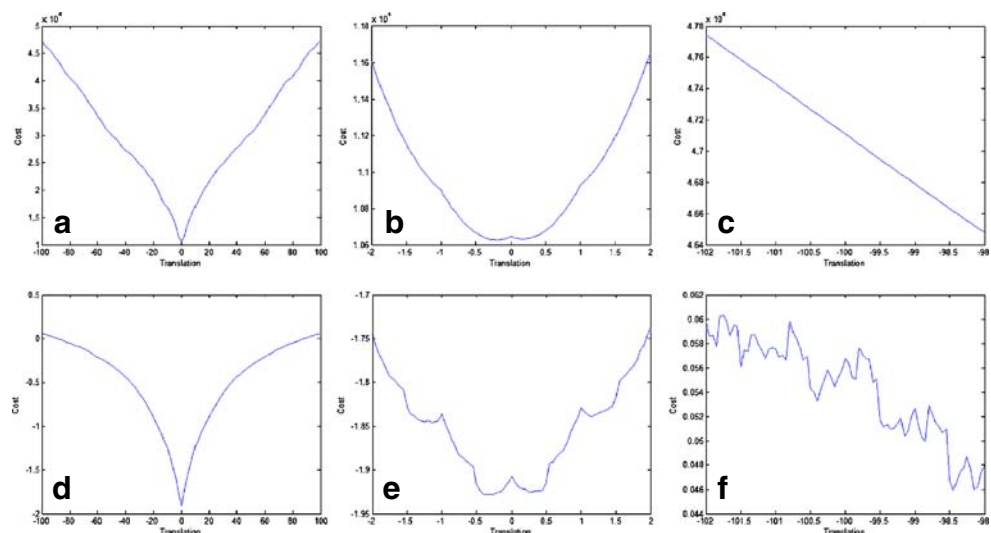
Table 1 Registration accuracy for Test 3.

Test set	RMSE[a]	
	NMI	Proposed method
TEST1	0.3552 / 15.6746 / 57.0678	0.3131 / 0.5265 / 1.0301
TEST2	0.4033 / 27.9273 / 76.4422	0.2634 / 0.3560 / 0.4649
TEST3	0.2028 / 11.2110 / 51.8502	0.1058 / 0.1615 / 0.2290
TEST4	1.0501 / 22.3580 / 78.9831	0.6031 / 1.8968 / 3.5440
TEST5	0.4475 / 12.4048 / 47.1655	0.3618 / 0.6477 / 1.0913

[a]The RMSE is computed over 10 random distortions. The table reports the min/mean/max of the 10 trials. The RMSE is computed in pixels relative to the reference image.

where δ represents the translation, $\bar{f} = f(-\underline{x})$, F and F^{-1} represent the forward and inverse FFTs respectively, and T_θ represents a rotational transformation by angle θ. For the sake of brevity, we refer to [22] and [16] for the full derivation of the aforementioned cost evaluation using FFT. By solving the problem in this manner, the computational complexity of having to evaluate this cost function is reduced from $O(N^4)$ (if solved in a direct manner) to $O(N^2 \log N)$ for the case where the images are $N \times N$ in size. For example,

based on timing results reported in [22], the speedup factor for globally exhaustive evaluation of translations for images of size 512×512 is approximately $60\times$. The optimal cost can then be evaluated for a number of different rotation angles. Finally, the rotation angle can be fine-tuned using an extension of the FFT method [23].

3.3 Fine-Grained Sub-pixel Registration

After an initial estimate is found using a global optimization scheme based on a rigid transformation model, a local optimization scheme can then be used in conjunction with robust estimation methods to perform fine-grained transformation estimation at the sub-pixel level. In the current implementation of the proposed method, the local optimization scheme used is based on SQP. An overview of SQP can be found in [24]. In the proposed method, the parameters of the transformation model between the images being registered is estimated in an iterative manner using SQP. The images are transformed based on the estimated parameters and the LPCR cost function is evaluated. The parameters of the transformation model are then re-estimated iteratively until convergence is reached to

Figure 7 Data-set 1 for Test 3: **a** T1-weighted MR image, **b** T2-weighted MR image, **c** T1–T2 MR overlay showing initial misregistration, **d** registration using the LPCR method.
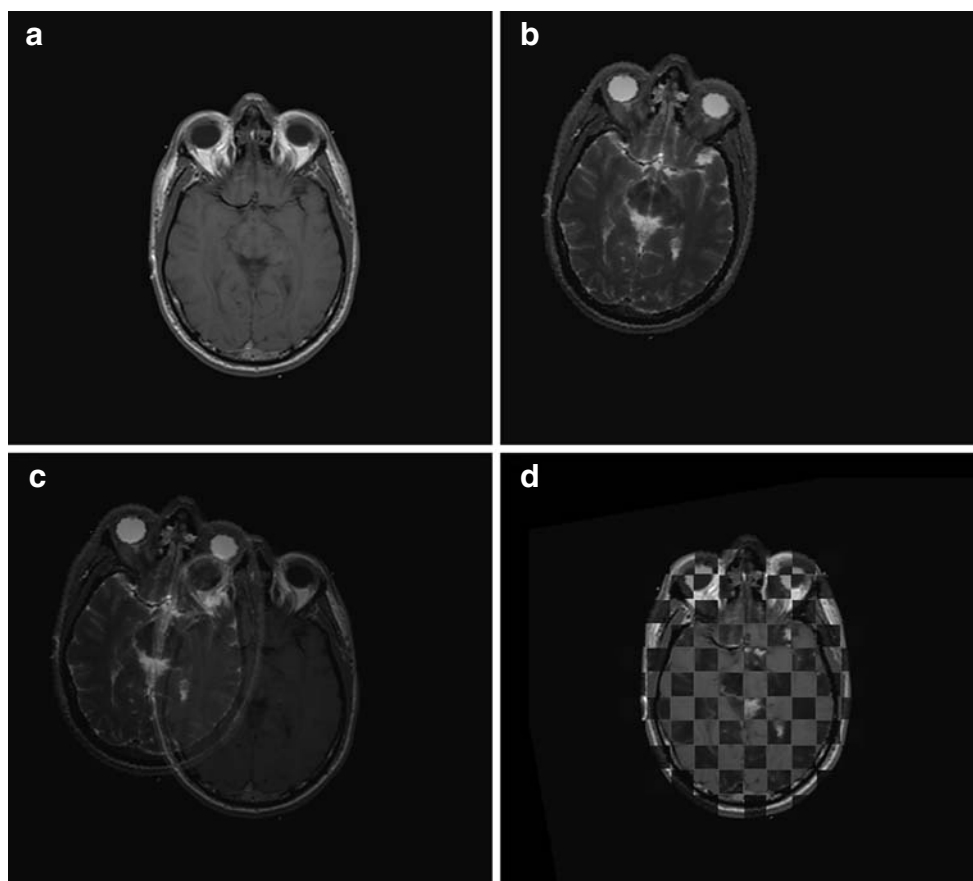
determine the optimal registration. The parameters of the transformation model are estimated using floating-point values, thus allowing the proposed method to achieve sub-pixel accuracy. By using a non-rigid transformation model, the optimal registration between two images can be non-rigid in nature. To improve convergence and reduce the likelihood of being trapped in local optima, a multi-resolution approach involving three different scales ($s = \frac{1}{4}, \frac{1}{2}, 1$) was used. Also, an affine transformation model was used. However, more complex transformation models may be integrated into the proposed method.

4 Testing Methods

The proposed method was tested under three different test scenarios.

1) **Test 1**: The first test scenario evaluates the effect of image non-homogeneity. Three cost functions were tested: intensity gradient methods (using the correlation coefficient on gradient magnitudes), NMI methods, and the LPCR method. This is accomplished by evaluating the cost functions for the registration scenario involving two images with high image non-homogeneity over various rotations within the range $[-\pi/2, \pi/2]$ (radians). The

test set for this scenario is a T1-T2 MR pair of a brain scan in the axial view obtained from the National Library of Medicine (NLM) Visible Human Project (VHP). Contrast variations were synthetically applied to the MR pair to simulate image non-homogeneity. This test scenario is designed to justify the claim that the proposed method does not suffer from the contrast non-homogeneity issues that hinder gradient-based methods.

2) **Test 2**: The second test scenario investigates the monotonicity of the proposed cost function and the NMI cost function. The monotonicity of a cost function is very important as it has a big effect on whether a local optimization scheme will converge to the global optimum. If the cost function exhibits highly non-monotonic behaviour, a local optimization scheme is more likely to be trapped within local optima. The test is performed by evaluating the NMI and LPCR cost functions between two images over various horizontal translations within the ranges of $[-100, 100]$, $[-2, 2]$, and $[-102, -98]$ (pixels). The test set used is the same as that of the first test scenario. This test scenario is designed to justify the claim that the LPCR method is less likely to be trapped in local optima than MI methods.

3) **Test 3**: The third test scenario evaluates the registration performance of the proposed method and

Figure 8 Data-set 4 for Test 3: **a** T1-weighted MR image, **b** PET image, **c** T1-PET overlay showing initial misregistration, **d** registration using the LPCR method.
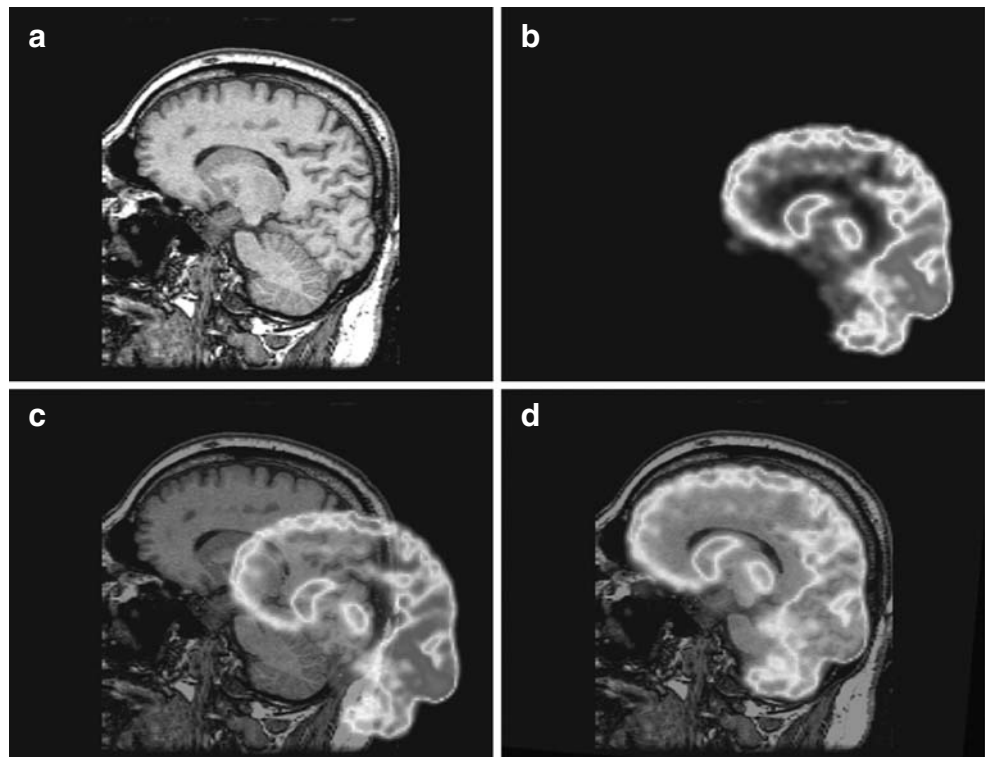
NMI using real multimodal medical image data-sets obtained from the NLM VHP and Whole Brain Atlas [25] (WBA). A summary of each data-set is given below.

a) **Data-set 1**: Brain scan, axial view, 1 mm resolution, T1-T2, VHP.
b) **Data-set 2**: Pelvis scan, coronal view, 1.875 mm resolution, T1-T2, VHP.
c) **Data-set 3**: Torso scan, coronal view, 1.875 mm resolution, T1-T2, VHP.
d) **Data-set 4**: Brain scan, sagittal view, 1 mm resolution, T1-PET, WBA.
e) **Data-set 5**: Brain scan, axial view, 1 mm resolution, PD-CT, VHP.

Each test data-set was synthetically distorted using randomized affine transformations on a sub-pixel level that resulted in small initial overlap. Since the data-sets were originally aligned, the ground-truth transformation parameters are known. The LPCR method and NMI were then used to estimate the transformation that re-aligns the data-sets. To eval-uate the registration accuracy of the methods, the root mean square error (RMSE) of the displacement of 20 ground-truth control point pairs was computed. The same optimization method used to implement the proposed method was used to implement the NMI method, where the local optimization of the cost functions were achieved using SQP within the a multi-resolution framework ($s = \frac{1}{4}, \frac{1}{2}, 1$). NMI was implemented using smoothed histograms computed with 256 intensity bins, as described in [14].

5 Experimental Results

The cost function cross-sections for Test 1 are illustrated in Fig. 5. There is a significant offset between the global minimum of the gradient-based cost function and the ground-truth alignment of the T1–T2 MR images. This is not the case for LPCR and NMI, where the global minimum coincides with the ground-truth alignment of the images. These results demonstrate the

Figure 9 Data-set 5 for Test 3: **a** CT image, **b** PD-weighted MR image, **c** CT-PD overlay showing initial misregistration, **d** registration using the LPCR method.
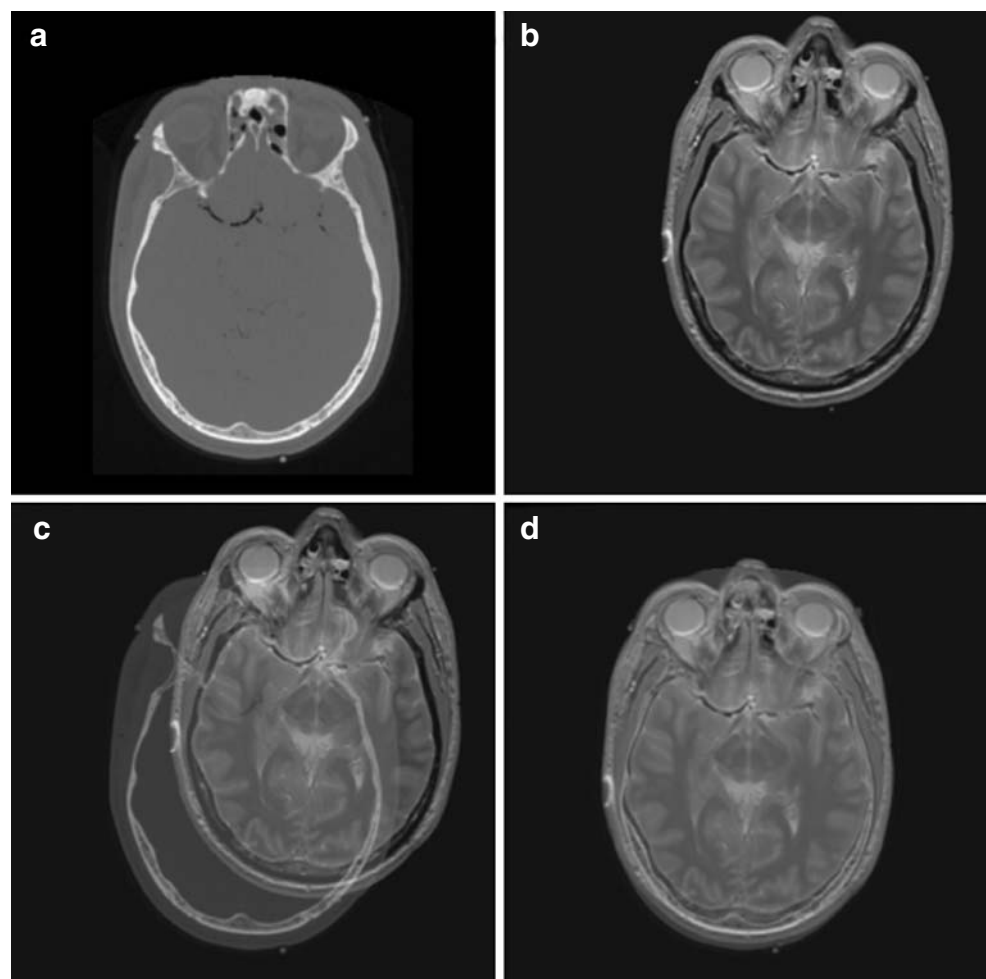
effectiveness of LPCR in situations characterized by high image non-homogeneity.

The cost cross-sections for the NMI and the proposed LPCR cost functions for Test 2 are illustrated in Fig. 6. While both cost functions point to the correct global solution, there is substantial undulation in the NMI cost function. This is not the case for the LPCR method, where the cost function remains smooth. These results demonstrate the monotonicity of the proposed method compared to NMI.

The registration accuracy results for 10 random trials of each test data-set for Test 3 are shown in Table 1. The proposed method yielded lower RMSE than NMI for all of the test sets. Representative examples of T1-T2 registration, T1-PET registration, and CT-PD registration achieved using the proposed method are shown in Figs. 7, 8, and 9 respectively. These results demonstrate the ease by which a local optimization algorithm can converge to the correct global solution using the LPCR-based cost function.

6 Conclusions and Future Work

In this paper, we have introduced a robust approach to the problem of registering multimodal images with high image contrast non-homogeneity and small initial overlap. Experimental results show that improved registration accuracy can be achieved when compared to existing methods. The proposed cost function exhibits improved monotonicity compared to NMI, suggesting why the entropy-based cost functions tend to converge to the wrong solution.

Future work includes an extensive investigation of different robust estimators as well as Monte-Carlo approaches to improve robustness of the proposed framework. Furthermore, we would like to test the proposed method on an even wider range of imaging modalities such as ultrasound images as well as additional data sets with low signal-to-noise ratios. We also plan to study how our method behaves in the presence of severe noise.

References

1. Simmons, A., Tofts, P., Barker, G., & Arridge, S. (1994). Sources of intensity nonuniformity in spin echo images at 1.5T. *Magnetic Resonance in Medicine, 32*(1), 121–128.

2. Oghabian, M., Mehdipour, S., & Alam, N. (2003). The impact of RF inhomogeneity on MR image non-uniformity, *Proc. Image and Vision Computing New Zealand*.

3. Nelder, J., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7*, 308–313.

4. Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., & Marchal, G. (1995). Automated multi-modality image registration based on information theory. *Proceedings of information processing in medical imaging* (pp. 263–274).

5. Viola, P., & Wells, W. (1997). Alignment by maximization of mutual information. *International Journal of Computer Vision, 24*(2), 137–154.

6. Studholme, C., Hill, D., & Hawkes, D. (1999). An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition, 32*(1), 71–86.

7. Crum, W., Hill, D., & Hawkes, D. (2003). Information theoretic similarity measures in non-rigid registration. *Proceedings of information processing in medical imaging* (pp. 378–387).

8. Pluim, J., Maintz, J., & Viergever, M. (2003). Mutual-information-based registration of medical images: A survey. *IEEE Transactions on Medical Imaging, 22*(8), 986–1004.

9. Mellor, M., & Brady, M. (2005). Phase mutual information as a similarity measure for registration. *Medical Image Analysis, 9*, 330–343.

10. Felsberg, M., & Sommer, G. (2001). The monogenic signal. *IEEE Transactions on Signal Processing, 49*(12), 3136–3144.

11. Wong, A., Bishop, W., & Orchard, J. (2006). Efficient multi-modal least-squares alignment of medical images using quasi-orientation maps. *Proceedings of international conference on image processing, computer vision, and pattern recognition*.

12. Haber, E., & Modersitzki, J. (2006). Intensity gradient based registration and fusion of multi-modal images. *Proceedings of the international conference on medical image computing and computer assisted intervention* (pp. 726–733).

13. Orchard, J. (2007). Globally optimal multimodal rigid registration: an analytic solution using edge information. *Proceedings of the IEEE international conference on image processing*.

14. Liu, J., Vemuri, B., & Marroquin, J. (2002). Local frequency representations for robust multimodal image registration. *IEEE Transactions on Medical Imaging, 21*(5), 462–469.

15. Hemmendorff, M., Andersson, M., Kronander, T., & Knutsson, H. (2002). Phase-based multidimensional volume registration. *IEEE Transactions on Medical Imaging, 21*(12), 1536–1543.

16. Wong, A., & Bishop, W. (2008). Efficient least squares fusion of MRI and CT images using a phase congruency model. *Pattern Recognition Letters, 29*(3), 173–180.

17. Morrone, M., & Burr, D. (1998). Feature detection in human vision: a phase-dependent energy model. *Proceedings of Royal Society of London B, 235*, 221–245.

18. Kovesi, P. (2003). Phase congruency detects corners and edges. *Proceedings of Australian pattern recognition society conference* (pp. 309–318).

19. Wang, Z., & Simoncelli, E. (2004). Local phase coherence and the perception of blur. *Advances in neural information processing systems, 16*, May.

20. Selesnick, I., Ivan , W., Baraniuk, R., & Kingsbury, N. (2005). The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine, 22*(6), 123–151.

21. Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12*(7), 629–639.

22. Orchard, J. (2005). Efficient global weighted least-squares translation registration in the frequency domain. *International conference on image analysis and recognition* (Vol. 3656, pp. 116–124). LNCS.
23. Omanovic, M., & Orchard, J. (2006). Efficient multi-modal registration using least-squares. *Proceedings of the international conference on image processing and computer vision*.
24. Boggs, P., & Tolle, J. (1995). Sequential quadratic programming. *Acta Numerica*, 151
25. Johnson, K., & Becker, J. The whole brain atlas. http://www.med.harvard.edu/AANLIB/home.html.
26. Collins, D., Zijdenbos, A., Kollokian, V., Sled, J., Kabani, N., Holmes, C., & Evans, A. (1998). Design and construction of a realistic digital Brain phantom. *IEEE Transactions on Medical Imaging, 17*(3), 463–468.
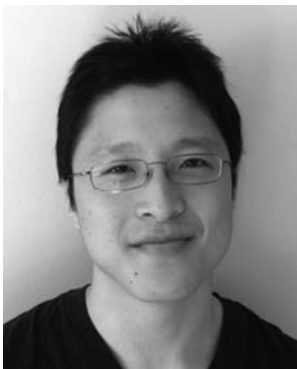
Alexander Wong   received the B.Sc. degree in computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2005 and the M.Sc. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2007. Since 2007, he has been a Ph.D. candidate in the Department of Systems Design Engineering, University of Waterloo. He has published refereed journal and conference papers in various fields such as computer vision, graphics, image processing, and multimedia systems. His research interests revolve around image processing, computer vision, and pattern recognition. He has worked on projects in image registration, image denoising, image super-resolution, image segmentation, biomedical tracking, and image and video coding. At the University of Waterloo, he is affiliated with the Vision and Image Processing Research Group.

Jeff Orchard   received the B.Math. degree in applied mathematics from the University of Waterloo, Waterloo, ON, Canada, in 1994, the M.Sc. degree in applied mathematics from the University of British Columbia, BC, Canada, in 1996, and the Ph.D. degree in computing science from Simon Fraser University, Canada, in 2003. Since 2003, he has been an Assistant Professor in the David R. Cheriton School of Computer Science, University of Waterloo. His research interests revolve around applying mathematics and computation to visual data. He has worked on projects in image registration, motion compensation for medical imaging, functional MRI, medical image reconstruction, and image mosaicking. At the University of Waterloo, he is affiliated with the Scientific Computing Research Group, the Waterloo Institute for Health Informatics Research, and the Centre for Computational Mathematics in Industry and Commerce. In 2005, he organized a workshop called the "Grand Mathematical Challenges in Medical Image Processing."