Differential affine motion estimation for medical image registration

Senthil Periaswamy[a], John B. Weaver[b,a], Dennis M. Healy Jr.[c], Daniel N. Rockmore[d,a], Peter J. Kostelec[d] and Hany Farid[a]

[a] Department of Computer Science, Dartmouth College, Hanover, NH 03755
[b] Department of Radiology, Dartmouth Hitchcock Medical Center, Lebanon, NH 03756
[c] Department of Mathematics, University of Maryland, College Park, MD 20742
[d] Department of Mathematics, Dartmouth College, Hanover, NH 03755

ABSTRACT

We have applied techniques from differential motion estimation in the context of automatic registration of medical images. This method uses optical-flow and Fourier techniques for local/global registration. A six parameter affine model is used to estimate shear, rotation, scale and translation. We show the efficacy of this method with images of similar and different contrasts.

Keywords: differential affine motion estimation, optical-flow, medical image registration, multi-modality registration

1. INTRODUCTION

Medical image registration is a well studied problem and the literature is rich in techniques. Various survey papers have been published on the topic.[1-5] However, we believe that the potential of approaches based on *differential* motion estimation have not been fully realized (see[6-8] for earlier related work). This approach is attractive because its formulation is elegant, and the resulting computations are fast, simple, and non-feature based. We demonstrate the effectiveness of this technique for registration of medical images with images of similar and different contrasts. We recognize that *global* affine registration by itself can not fully characterize complex distortions. However, it is a key component for a more generic elastic registration technique.

2. DIFFERENTIAL AFFINE MOTION ESTIMATION

2.1. Methods

The estimation of parametric motion based on spatio-temporal variations has been extensively studied in the computer vision literature.[9-11] Most of these methods begin by making an assumption of brightness constancy.[12] That is, it is assumed that the difference between a pair of images is due solely to motion and not to any brightness variations. Most approaches are then characterized as differential[13,14]-, energy[15,16]- or Fourier[17]- based. While these methods are equivalent,[10] they can differ in their parameterization of the motion. In the simplest case, it is assumed that the motion is due to translations. A more generic six parameter affine model allows for shear, rotation, scale and translation.[18-20] We describe a slightly modified version of this model that includes symmetric translations. The affine parameters are estimated using a differential framework.

Let $f(x, y)$ and $g(x, y)$ represent *test* and *reference* images. If the brightness constancy assumption holds, and the images are related by an affine transform, then:

$$f(x, y) = g(m_1 x + m_2 y + t_x, m_3 x + m_4 y + t_y),$$ (2.1)

where m_1, m_2, m_3, m_4 form the affine matrix A and t_x, t_y form the displacement vector $\vec{t}$:

$$A = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \quad \text{and} \quad \vec{t} = \begin{pmatrix} t_x \\ t_y \end{pmatrix}.$$ (2.2)

For reasons that will become clear shortly, the translations are made "symmetric" as follows:

$$f\left(x - \frac{t_x}{2}, y - \frac{t_y}{2}\right) = g\left(m_1 x + m_2 y + \frac{t_x}{2}, m_3 x + m_4 y + \frac{t_y}{2}\right). \tag{2.3}$$

Substituting $m_5 = \frac{t_x}{2}$ and $m_6 = \frac{t_y}{2}$, Equation (2.3) is rewritten as:

$$f\left(x - m_5, y - m_6\right) = g\left(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6\right). \tag{2.4}$$

The goal is to estimate the six parameters $\vec{m} = \begin{pmatrix} m_1 & m_2 & m_3 & m_4 & m_5 & m_6 \end{pmatrix}^T$. In order to do this, we first define the following quadratic error function to be minimized:

$$E(\vec{m}) = \sum_\Omega \left[f\left(x - m_5, y - m_6\right) - g\left(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6\right)\right]^2, \tag{2.5}$$

where Ω defines the spatial region of interest in the image. This error function cannot be minimized analytically since it is non-linear in its unknowns. To simplify the minimization, we approximate this error function using a first-order truncated Taylor series expansion:

$$
\begin{aligned}
E(\vec{m}) \quad &\approx \quad \sum_\Omega [(f(x,y) - m_5 f_x(x,y) - m_6 f_y(x,y)) \\
&\quad - (g(x,y) + (m_1 x + m_2 y + m_5 - x)\, g_x(x,y) + (m_3 x + m_4 y + m_6 - y)\, g_y(x,y))]^2
\end{aligned} \tag{2.6}
$$

where $f_x(\cdot), f_y(\cdot), g_x(\cdot)$ and $g_y(\cdot)$ are the spatial derivatives of $f(\cdot)$ and $g(\cdot)$ with respect to x and y. Note that this error function is now linear in its unknowns. Note also that this Taylor series expansion is most accurate for small motions. As a result, by symmetrically distributing the translation terms, the errors in the expansion are reduced. The error function may be expressed more compactly in vector form as follows:

$$E(\vec{m}) = \sum_\Omega \left(k - \vec{c}^T \vec{m}\right)^2, \tag{2.7}$$

where the vector $\vec{c}$ and scalar k are defined as:

$$
\begin{aligned}
\vec{c}^T \quad &= \quad \begin{pmatrix} xg_x(x,y) & yg_x(x,y) & xg_y(x,y) & yg_y(x,y) & (f_x(x,y) + g_x(x,y)) & (f_y(x,y) + g_y(x,y)) \end{pmatrix} \tag{2.8} \\
k \quad &= \quad f(x,y) - g(x,y) + xg_x(x,y) + yg_y(x,y) \tag{2.9}
\end{aligned}
$$

This error function can now be minimized by differentiating with respect to $\vec{m}$:

$$\frac{dE}{d\vec{m}} = \sum_\Omega -2\vec{c}\left(k - \vec{c}^T \vec{m}\right), \tag{2.10}$$

setting equal to zero, and solving for $\vec{m}$:

$$\vec{m} = \left(\sum_\Omega \vec{c}\vec{c}^T\right)^{-1} \left(\sum_\Omega \vec{c}k\right). \tag{2.11}$$

This solution assumes that the 6×6 matrix $\left(\sum_\Omega \vec{c}\vec{c}^T\right)$ is invertible. This can usually be guaranteed by integrating over a large enough spatial region Ω with sufficient image content.

Recall that the error function being minimized in Equation (2.7) is only an approximation to the actual error function in Equation (2.5). The estimation from the approximate error function is improved using a Newton-Raphson style iterative scheme.[21] At each iteration, we estimate the affine parameters between the test and reference image, warp the test image according to this estimate, and repeat.

The required spatio-temporal derivatives have finite support thus fundamentally limiting the amount of motion that can be estimated. A coarse-to-fine scheme is adopted in order to contend with larger motions.[14,22] A Gaussian

pyramid is built for both images, and the affine parameters are first estimated at the coarsest level. These parameters are estimated according to Equation (2.11) using the iterative scheme described above. The estimated parameters are then used in the next level to warp the test image, and the affine parameters re-estimated. This procedure is repeated through each level of the pyramid.

The calculation of the spatio-temporal derivatives is an important step in this procedure. Spatio-temporal derivatives of discretely sampled images are often computed as differences between neighboring sample values. Such differences are typically poor approximations to derivatives and lead to substantial errors in the motion estimation. In computing spatio-temporal derivatives we employ derivative filters specifically designed for multi-dimensional differentiation.[23] These filters significantly improve both the stability and accuracy of the estimate.

2.2. Results

We tested the efficacy of the differential approach on a broad range of medical images with simulated and real distortions. All the images were 8-bit and of size 256×256. In all of our examples, a three-level pyramid was constructed, and at each scale, the affine parameters were iteratively estimated ten times. In the first set of results, a set of *test* images were subjected to a variety of global affine warps to generate the *reference* images. The coarse-to-fine differential technique described above was used to estimate the six affine parameters, Equation (2.2). Shown in Figure 1 are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration (columns 4,5). The error images are displayed on the same log intensity scale and then rescaled into the full intensity range. Shown in Table 1 are the actual and estimated affine parameters and the mean squared error (MSE) before and after registration. In these examples the estimated parameters are nearly perfect.

In the next set of results, the test and reference images were obtained from different subjects, or from subjects at different times. The same approach was used to estimate the global affine parameters. However, in order to minimize violations of the brightness-constancy assumption, the two images were first histogram matched*. Shown in Figure 2 are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration (columns 4,5). The error images are displayed on the same linear intensity scale and then rescaled into the full intensity range. In this case, the actual warp parameters are unknown. Qualitatively we see that the estimates are quite good, and that a global affine warp is reasonably effective in characterizing the overall distortion.

2.3. Weaknesses

While the differential framework yields good results there are several inherent limitations. First, because of the limited support of the derivative filters and the Taylor series approximation, this approach is fundamentally limited by the magnitude of the distortion. Large distortions are typically caused by large translations and cannot always be overcome even by using a coarse-to-fine strategy. Second, because of the brightness constancy assumption, this approach is unlikely to perform well with images of different contrasts. This will be particularly problematic when estimating distortions locally. These issues are addressed in the following section.

3. EXTENSIONS FOR IMAGES OF DIFFERENT CONTRASTS

In this section, we provide a slight variation on the same differential framework described above that contends with large distortions and images of different contrasts. Both of these limitations can be simultaneously addressed by simply estimating the distortions in the Fourier domain. Specifically, the affine parameters are estimated from the normalized power spectra of the test and reference images. By using the power spectrum, the potentially large translations are decoupled from the linear affine parameters. Violations of the brightness-constancy assumption are contended with by using a *normalized* power spectrum.
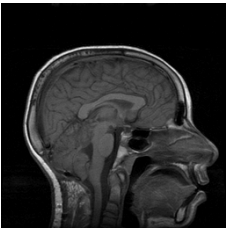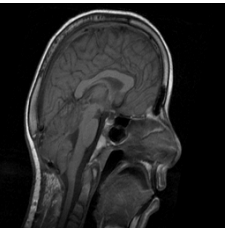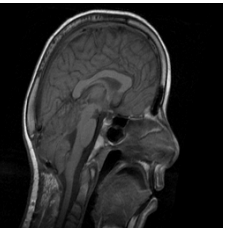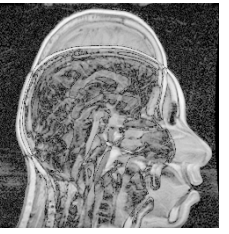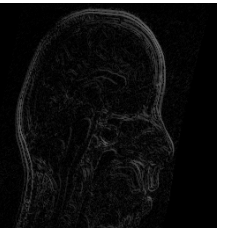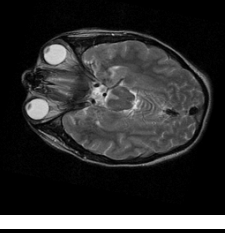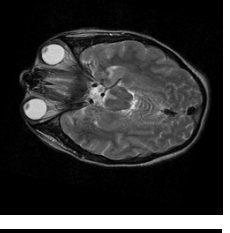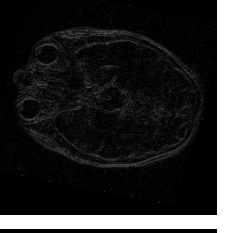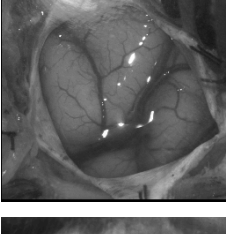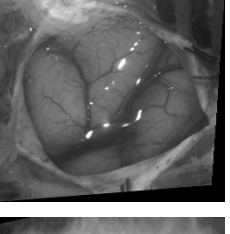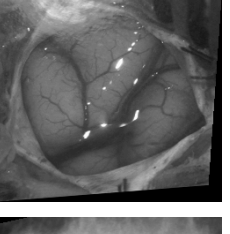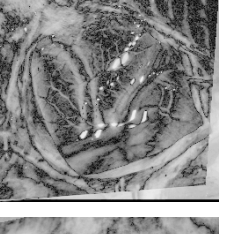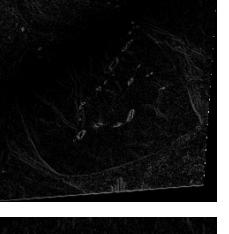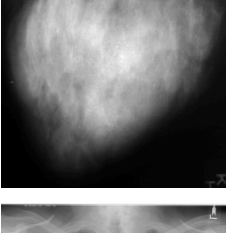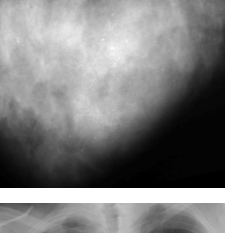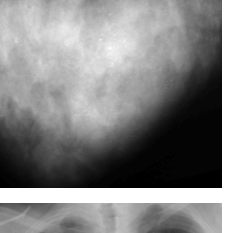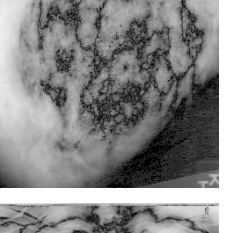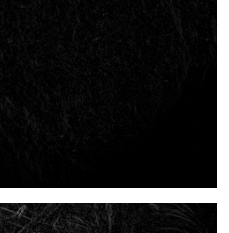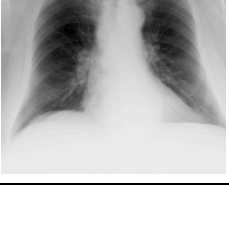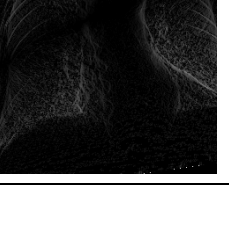
3.1. Methods

It is well known that the power spectrum decouples the translation term and preserves the linear affine transform.[24] More specifically, let $f(\vec{x})$ be the test image with Fourier transform:
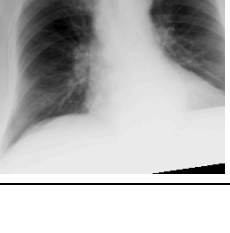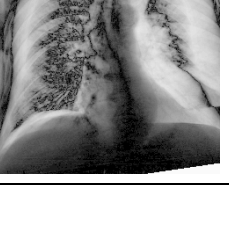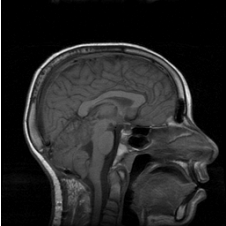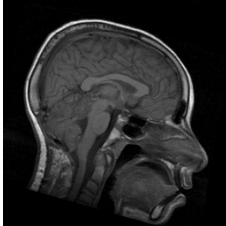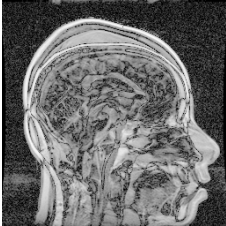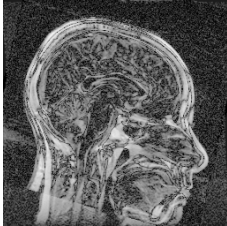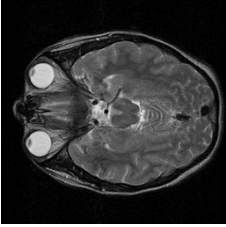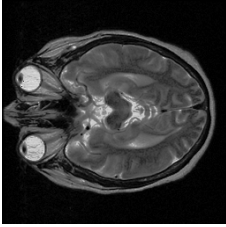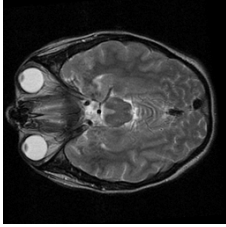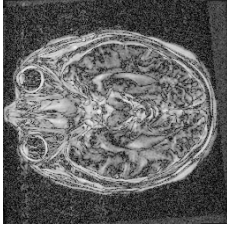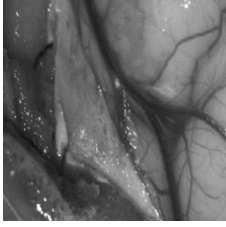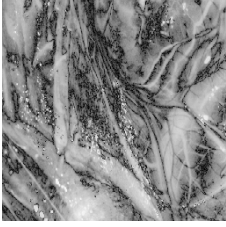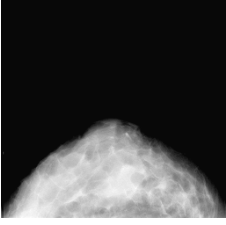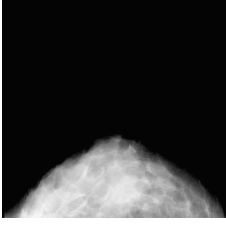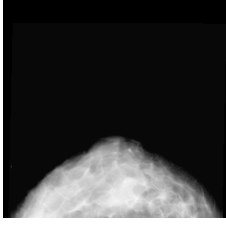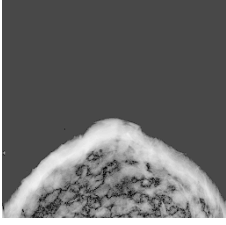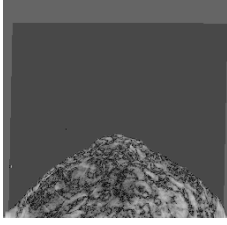
$$F(\vec{u}) = \int_\Omega f(\vec{x}) e^{-i2\pi \vec{u}^T \vec{x}} d\vec{x}. \tag{3.1}$$

---

*Denote $c_1(\cdot), c_2(\cdot)$ as the sampled cumulative distribution functions of images $f_1(\cdot)$, $f_2(\cdot)$. The histogram of $f_2(\cdot)$ is made to match that of $f_1(\cdot)$ by mapping each pixel $f_2(x, y)$ to $c_1^{-1}(c_2(f_1(x, y)))$.

Figure 1. Registration with simulated distortions. Shown are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration (columns 4,5). The error images are displayed on the same log intensity scale. See also Table 1.

Figure 2. Registration with real distortions. Shown are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration (columns 4,5). The error images are displayed on the same linear intensity scale. See also Table 1.

Table 1. Affine parameter estimates with simulated (rows 1-5) and real (rows 6-9) distortions. See also Figures 1 and 2.

	Actual A	$\vec{t}$	Estimated A	$\vec{t}$	MSE before	MSE after
1	$\begin{bmatrix} 1.182 & -0.090 \\ 0.254 & 0.795 \end{bmatrix}$	$\begin{bmatrix} 16 \\ 16 \end{bmatrix}$	$\begin{bmatrix} 1.182 & -0.090 \\ 0.254 & 0.795 \end{bmatrix}$	$\begin{bmatrix} 16.04 \\ 15.95 \end{bmatrix}$	1952.4	0.8
2	$\begin{bmatrix} 1.066 & -0.228 \\ 0.290 & 1.178 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 8 \end{bmatrix}$	$\begin{bmatrix} 1.066 & -0.228 \\ 0.290 & 1.178 \end{bmatrix}$	$\begin{bmatrix} 3.02 \\ 8.04 \end{bmatrix}$	1934.1	0.4
3	$\begin{bmatrix} 0.960 & 0.086 \\ 0.067 & 1.097 \end{bmatrix}$	$\begin{bmatrix} 12 \\ -4 \end{bmatrix}$	$\begin{bmatrix} 0.960 & 0.086 \\ 0.067 & 1.096 \end{bmatrix}$	$\begin{bmatrix} 12.04 \\ -4.03 \end{bmatrix}$	2284.3	2.9
4	$\begin{bmatrix} 0.700 & 0.200 \\ -0.150 & 0.900 \end{bmatrix}$	$\begin{bmatrix} 5 \\ 6 \end{bmatrix}$	$\begin{bmatrix} 0.700 & 0.200 \\ -0.150 & 0.900 \end{bmatrix}$	$\begin{bmatrix} 4.95 \\ 6.01 \end{bmatrix}$	1975.7	0.1
5	$\begin{bmatrix} 0.795 & 0.130 \\ -0.088 & 0.891 \end{bmatrix}$	$\begin{bmatrix} -7 \\ 8 \end{bmatrix}$	$\begin{bmatrix} 0.794 & 0.129 \\ -0.091 & 0.892 \end{bmatrix}$	$\begin{bmatrix} -7.19 \\ 8.02 \end{bmatrix}$	2467.2	7.0
6	-	-	$\begin{bmatrix} 0.973 & -0.271 \\ 0.233 & 0.954 \end{bmatrix}$	$\begin{bmatrix} 5.39 \\ 18.15 \end{bmatrix}$	1683.5	710.0
7	-	-	$\begin{bmatrix} 1.003 & 0.068 \\ -0.096 & 1.006 \end{bmatrix}$	$\begin{bmatrix} 11.14 \\ 1.31 \end{bmatrix}$	1559.5	822.3
8	-	-	$\begin{bmatrix} 0.764 & -0.035 \\ 0.011 & 0.768 \end{bmatrix}$	$\begin{bmatrix} 0.22 \\ -6.56 \end{bmatrix}$	1562.8	252.5
9	-	-	$\begin{bmatrix} 1.047 & -0.002 \\ 0.022 & 1.110 \end{bmatrix}$	$\begin{bmatrix} -2.74 \\ -23.44 \end{bmatrix}$	1328.9	86.7

	f	g	$\widetilde{f}$	$|f - g|$	$|\widetilde{f} - g|$
1					
2					
3					

Figure 3. Registration in the Fourier domain with simulated distortions. Shown are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration (columns 4,5). The error images are displayed on the same linear intensity scale. See also Table 2.
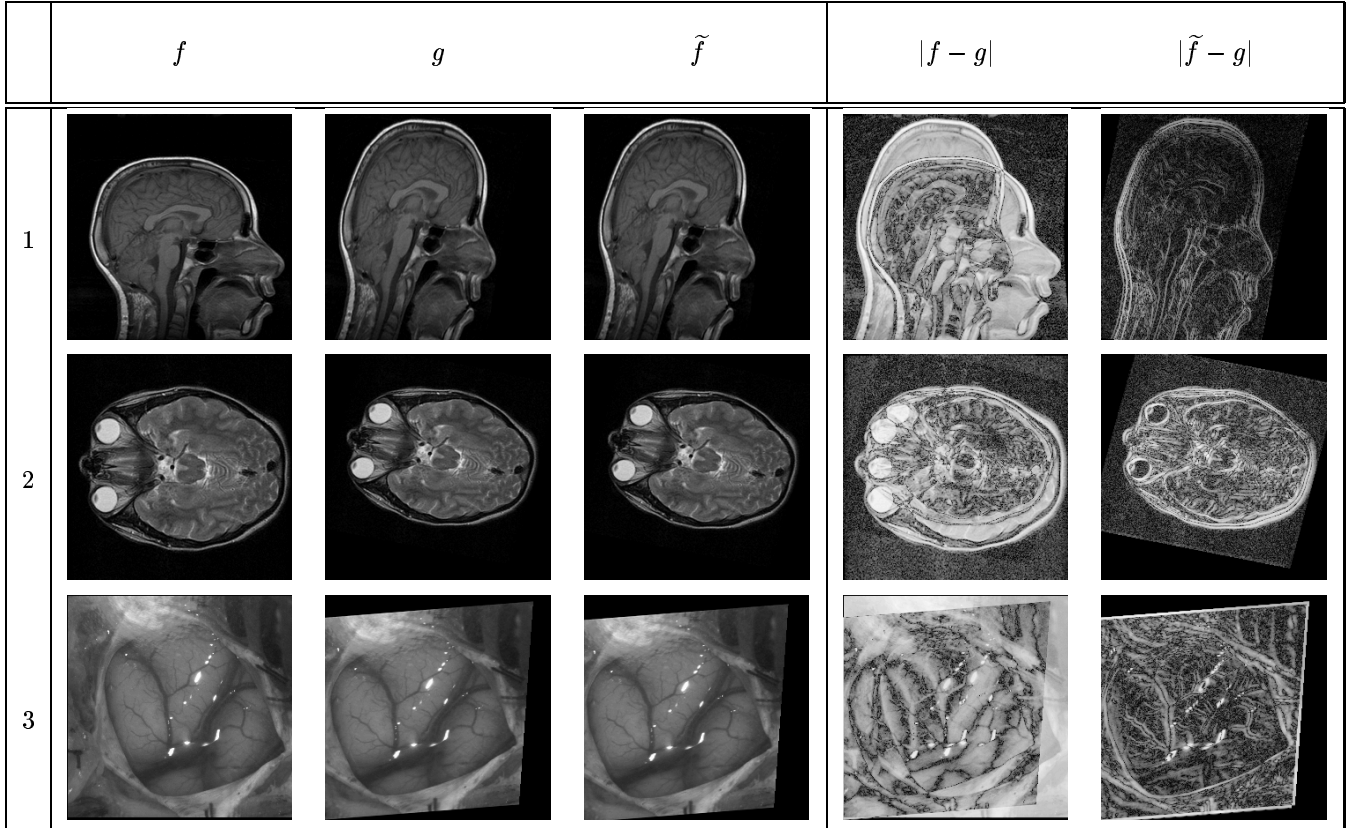
Let $g\left(A\vec{x} + \vec{t}\right)$ be the distorted reference image with Fourier transform $G\left(\vec{u}\right)$, where A and $\vec{t}$ are the affine parameters defined in Equation (2.2). The power spectrum is then given by:

$$|G\left(\vec{u}\right)| \;=\; \frac{1}{|det\,A|}\left|F\left(\left(A^{-1}\right)^{T}\vec{u}\right)\right|.$$ (3.2)

Note that in the Fourier domain the reference image is related to the test image by only the linear portion of the original affine distortion (i.e., the translation component $\vec{t}$ has been decoupled). Also note that the power spectra are related within a scale factor. In order to contend with this scale factor and other potential violations of the brightness-constancy assumption, we normalize the power spectra before estimating the linear affine parameters. Each radial frequency band is divided by the mean energy in that band. This is most easily performed by first converting the power spectra from rectangular to polar coordinates, normalizing and inverting. Given the normalized power spectra of the test and reference images, the estimation of the linear affine transform terms is identical to that described in the previous section. The translation term is then estimated using standard phase-only correlation.[25]

3.2. Results

We tested the efficacy of the Fourier-based approach on a range of medical images with simulated and real distortions. In the first set of results we verify that the Fourier-based estimation yields similar results as previously shown in the spatial domain. In Figure 3 are a subset of the test images from Figure 1. The reference images were distorted using the same linear affine terms but with larger translations. The estimation of the linear affine terms was identical to that of Figure 1, and the translation was estimated separately as described above. Shown in Figure 3 are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration
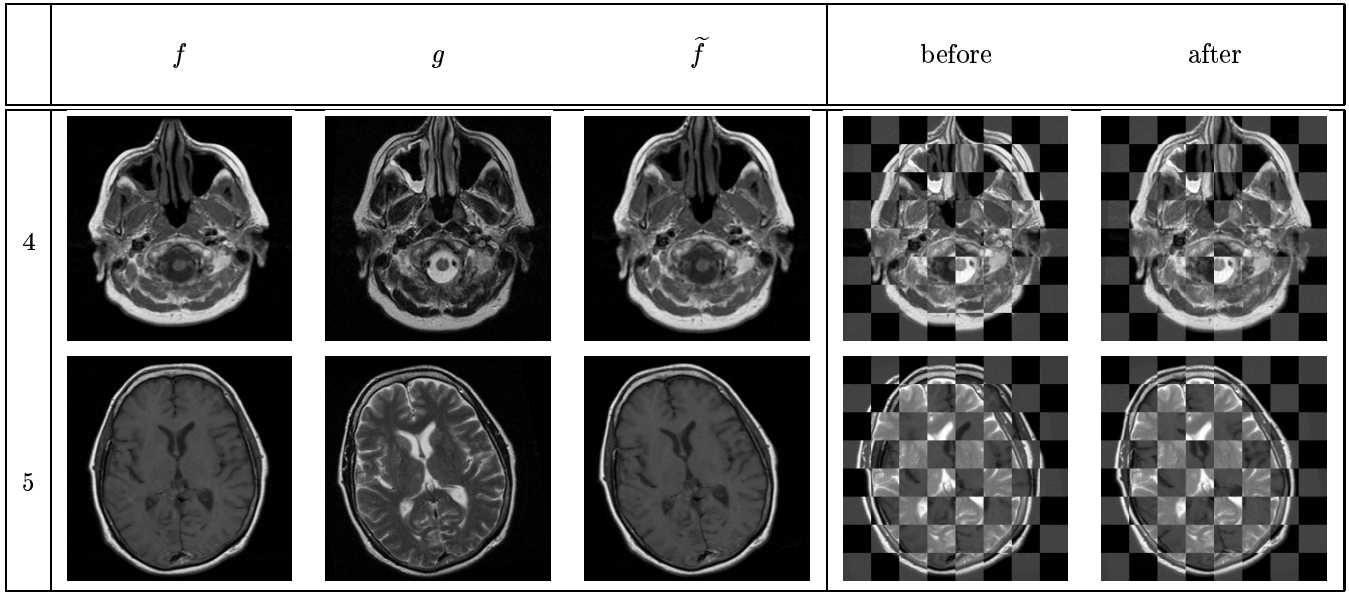
Figure 4. Registration in the Fourier domain using images of different contrasts with simulated distortions. Shown are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration (columns 4,5). See also Table 2.

Table 2. Affine parameter estimates for similar images (rows 1-3) and those with differing contrasts (rows 4-5). See also Figures 3 and 4.

	Actual A	$\vec{t}$	Estimated A	$\vec{t}$	MSE before	MSE after
1	$\begin{bmatrix} 1.182 & -0.090 \\ 0.254 & 0.800 \end{bmatrix}$	$\begin{bmatrix} 32 \\ 16 \end{bmatrix}$	$\begin{bmatrix} 1.186 & -0.091 \\ 0.250 & 0.800 \end{bmatrix}$	$\begin{bmatrix} 31.59 \\ 16.20 \end{bmatrix}$	2483.3	72.2
2	$\begin{bmatrix} 1.066 & -0.228 \\ 0.290 & 1.178 \end{bmatrix}$	$\begin{bmatrix} 6 \\ 16 \end{bmatrix}$	$\begin{bmatrix} 1.069 & -0.230 \\ 0.282 & 1.167 \end{bmatrix}$	$\begin{bmatrix} 2.70 \\ 13.11 \end{bmatrix}$	2071.3	632.2
3	$\begin{bmatrix} 0.960 & 0.086 \\ 0.067 & 1.097 \end{bmatrix}$	$\begin{bmatrix} 32 \\ -4 \end{bmatrix}$	$\begin{bmatrix} 0.961 & 0.085 \\ 0.069 & 1.095 \end{bmatrix}$	$\begin{bmatrix} 35.01 \\ -7.08 \end{bmatrix}$	3002.4	550.4
4	$\begin{bmatrix} 0.999 & -0.051 \\ 0.045 & 0.899 \end{bmatrix}$	$\begin{bmatrix} -7 \\ 5 \end{bmatrix}$	$\begin{bmatrix} 1.000 & -0.047 \\ 0.042 & 0.903 \end{bmatrix}$	$\begin{bmatrix} -8.34 \\ 4.28 \end{bmatrix}$	-	-
5	$\begin{bmatrix} 0.985 & 0.154 \\ -0.174 & 0.988 \end{bmatrix}$	$\begin{bmatrix} 12 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.994 & 0.130 \\ -0.192 & 0.991 \end{bmatrix}$	$\begin{bmatrix} 8.83 \\ 1.12 \end{bmatrix}$	-	-

(columns 4,5). The error images are displayed on the same linear intensity scale and then rescaled into the full intensity range. Shown in Table 2 are the actual and estimated affine parameters. Note that while the affine parameters are accurately estimated, the errors in translation are larger leading to an overall larger MSE.

In the next set of results, the test and reference images are MRI, T1/T2 images with different contrasts. These images were obtained from a single scan suggesting that the distortions between them are minimal. A known affine distortion was applied to the reference image. These affine parameters were then estimated as described above. Shown in Figure 4 are the test and reference images (columns 1,2), the registered test image (column 3) and the errors before and after registration (columns 4,5). Shown in Table 2 are the actual and estimated affine parameters. Since these images are of different contrasts, the MSE is not reported. Note that in the spatial domain, these images violate the brightness-constancy assumption. Nevertheless, by using the normalized power spectrum, the distortion can still be accurately estimated. The slightly larger errors are most likely due to small misalignments in the original images. Note once again that the estimation errors are dominated by errors in the translation.

4. DISCUSSION

We have shown how an elegant and simple framework for differential affine motion estimation can be used for medical image registration. The resulting computations are simple, fast, and non-feature based. This approach has shown itself to be robust over a broad range of medical images with similar and different contrasts.

There are several natural extensions to this work that we are currently pursuing. First, estimating the linear affine parameters in the Fourier domain affords us the ability to contend with images of different contrasts and large distortions. However, the resulting phase-only estimation of translation is sensitive to slight errors in the affine estimation. We are currently investigating how to improve the translation estimates. Second, while we have only tested the proposed technique for global registration, there is no inherent reason why this approach cannot be applied locally. Local parameter estimation coupled with a smoothness constraint will allow for more flexible modeling of distortions as in our earlier work.[26–29] While we believe that this approach is promising, it is unlikely that it will supplant the multitude of existing techniques such as the commonly used mutual information.[30,31] Rather, we expect that this approach will prove to be useful in conjunction with other techniques.

REFERENCES

1. L. Brown, "A survey of image registration techniques," *ACM Computing Surveys* **24**(4), pp. 325–376, 1992.
2. J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis* **2**(1), pp. 1–36, 1998.
3. J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, R. Kessler, and R. Maciunas, "Comparison and evaluation of retrospective intermodality image registration techniques," in *Image Processing, SPIE, Medical Imaging* **2710**, pp. 332–347, (Newport Beach), Apr. 1996.
4. H. Lester and S. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognition, Special issue on Image Registration* **32**(1), pp. 129–149, 1999.
5. P. V. den Elsen, E. Pol, and M. Viergever, "Medical image matching - a review with classification," *IEEE Engineering in Medicine and Biology* **12**(1), pp. 26–39, 1993.
6. D. Barber, W. Tindale, E. Hunt, A. Mayes, and H. Sagar, "Automatic registration of spect images as an alternative to immobilization in neuroactivation studies," *Physics in medicine and biology* **40**, pp. 449–463, 1995.
7. J. Meunier, A. Guimond, C. Janicki, B. Imbert, and J. Soucy, *Computer Assisted Radiology, Vol. 1124 of Excerpta Medica - International Congress Series*, ch. Automatic 3D registration of brain SPECT images, pp. 187–192. Elsevier, Amsterdam, 1996.
8. K.J. Friston, J. Ashburner, J.B. Poline, C.D. Frith, J.D. Heather, and R.S.J. Frackowiak, "Spatial registration and normalization of images," *Human Brain Mapping* **2**, pp. 165–189, 1995.

9. J. K. Aggarwal and N. Nandhakumar, "On the computation of motion from sequences of images - a review," *Proceedings of the IEEE* **76**(8), pp. 917–935, 1988.

10. E. Simoncelli, *Distributed Analysis and Representation of Visual Motion.* PhD thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, Jan. 1993. Also available as MIT Media Laboratory Vision and Modeling Technical Report #209.

11. J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision* **12**(1), pp. 43–77, 1994.

12. B. Horn, *Robot Vision*, MIT Press, Cambridge, MA, 1986.

13. B. K. P. Horn and B. G. Schunck, "Determining optical flow: a retrospective," *Artificial Intelligence* **59**(1-2), pp. 81–87, 1993.

14. B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proceedings of the 7th International Joint Conference on Artificial Intelligence* , pp. 674–679, (Vancouver), 1981.

15. E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *Journal of the Optical Society of America* **2**(2), pp. 284–299, 1985.

16. D. J. Heeger, "Optical flow using spatio-temporal filters," *International Journal of Computer Vision* **1**(4), pp. 279–302, 1988.

17. A. B. Watson and A. Ahumada, *Motion: Perception and representation*, ch. A look at motion in the frequency domain, pp. 1–10. 1983.

18. D.J.Fleet and A. D. Jepson, "Computation of component image velocity from local phase information," *International Journal of Computer Vision* **5**, pp. 77–104, 1990.

19. J. Bergen, P. Anandan, K. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," *Second European Conference on Computer Vision* **588**, pp. 237–252, 1992.

20. M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Computer Vision and Pattern Recognition* **63**, pp. 75–104, Jan. 1996.

21. J. Shi and C. Tomasi, "Good features to track," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* , June 1994.

22. P. Anandan, "A computational framework and an algorithm for the measurement of visual motion," *International Journal of Computer Vision* **2**(3), pp. 283–310, 1989.

23. H. Farid and E. Simoncelli, "Optimally rotation-equivariant directional derivative kernels," *International Conference on Computer Analysis of Images and Patterns* , pp. 207–214, (Kiel, Germany), Sept. 1997.

24. R. Bracewell, K. Chang, A. Jha, and Y. Wang, "Affine theorem for two-dimensional fourier transform," *Electronics Letters* **29**(3), p. 304, 1993.

25. E. D. Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," *IEEE Transactions on Pattern Analaysis and Machine Intelligence* **9**(5), pp. 700–703, 1987.

26. P.J. Kostelec, J.B. Weaver, and D.M. Healy, Jr., "Multiresolution elastic image registration," *Medical Physics* **25**(9), pp. 1593–604, 1998.

27. J. Weaver, D.M. Healy, Jr., S. Periaswamy, and P. Kostelec, "Elastic image registration using correlations," *Journal of Digital Imaging* , pp. 59–65, 1998.

28. J. Weaver, S. Periaswamy, D.M. Healy, Jr., and P. Kostelec, "The windowed correlation; minimizing the adverse effects of the windowing function," *Proceesings of the Society of Magnetic Resonance* , p. 126, (Philadelphia, PA), 1999.

29. S. Periaswamy, J.B. Weaver, D.M. Healy, Jr., and P.J. Kostelec, "Automated multiscale elastic image registration using correlation," *Proceedings of the SPIE - The International Society for Optical Engineering* **3661**, pp. 828–38, 1999.

30. A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, *Information Processing in Medical Imaging*, ch. Automated multimodality image registration using information theory, pp. 263–274. Kluwer, Dordrecht, 1995.

31. P. Viola and W. Wells III, "Alignment by maximization of mutual information," in *International Conference on Computer Vision*, pp. 16–23, (Boston, MA), 1995.