Efficient Multimodal Registration Using Least-Squares

Maja Omanovic
David R. Cheriton School of
Computer Science
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
Email: momanovi@cs.uwaterloo.ca

Jeff Orchard
David R. Cheriton School of
Computer Science
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
Email: jorchard@cs.uwaterloo.ca

Abstract— *Multimodal image registration is a difficult problem in both medical imaging and remote sensing. The least-squares cost function has generally been overlooked for multimodal registration problems due to an underlying assumption that the two images being registered must have corresponding intensities. More recently, methods that employ the least-squares cost function have been developed to efficiently evaluate the globally optimal shift and intensity remapping simultaneously. However, these methods estimate the translation and not the rotation. In this paper we propose a method for using the least-squares cost function efficiently for multimodal registration. By modeling rotation using a linear approximation, we find the globally optimal translation and intensity remapping, and locally optimal rotation angle. In a series of experiments based on registering PD-, T1-, and T2-weighted magnetic resonance images, our method performs better than mutual information.*

Keywords: image registration, multimodal registration, least squares, FFT

1 Introduction

Multimodal image registration has been widely utilized in the medical imaging domain. Its applications include medical image analysis, image guidance, atlas construction, pathology diagnosis, etc. Some imaging modalities distinguish different anatomical features (i.e. bone vs. soft tissue) better than other modalities. Hence, aligning these images provides medical professionals with more information than they get from a single modality, making diagnosis more reliable and accurate. Multimodal image registration is also widely used in remote sensing.

Until recently, least-squares optimization was not used for multimodal image registration because of its underlying assumption that the two images being registered must have corresponding intensities. That is, corresponding features in the two images must be represented with the same pixel intensities. Because of this limitation, the usefulness of least-squares registration has been largely restricted to alignment of serial volumes in applications such as functional MRI where it has enjoyed terrific success [1]–[5].

Entropy-based cost functions such as mutual information [6]–[8], and various forms of normalized mutual information [9], [10] represent the current state-of-the-art in multimodal registration. The advantage of these methods is that they make no assumption about the form of the intensity mapping from one image to the other; merely that pixels with matching intensities in one image are likely to have similar intensities in the other image.

However, mutual information has its share of issues. There is no known closed-form solution to find the optimum, so global optimization must be attempted with only local information of the cost function. Furthermore, the process of estimating the joint probability density function requires binning of pixels into discrete categories, giving the cost function discontinuities. There are interpolation methods to combat this problem, including parzen-windowing and trilinear partial volume distribution [6].

More recently, methods have been developed to efficiently evaluate the least-squares cost function for all integer shifts [11]–[14]. These methods take advantage of the Fast Fourier Transform (FFT) to evaluate the most time-consuming terms in the cost function. Speedups of up to $500\times$ have been reported in the literature [13]. These methods, however, have focussed on estimating translations, and not rotations. Since rigid-body motion is a common form of misregistration in medical imaging, rotations need to be accounted for.

In this paper, we propose an efficient framework for multimodal rigid-body registration using a weighted least-squares cost function that is invariant under linear intensity remappings. The proposed method includes a weighting function (sometimes called an *alpha channel*), allowing one to include or exclude pixels (even on a fractional basis) from the cost function. This is a capability that other Fourier methods, such as phase correlation, lack.

2 Theory

We begin with a brief outline of how the FFT factors into optimizing the least-squares cost function. To register an image (or volume[1]) $\mathbf{f}$ with an image (or volume) $\mathbf{g}$, we seek a rigid-body transformation T, that depends on motion parameters $\mathbf{p}$, that minimizes

$$L(\mathbf{p}) = \sum_{i,j} \left(\tilde{\mathbf{f}}_{i,j} - \mathbf{g}_{i,j} \right)^2 \mathbf{w}_{i,j} \tag{1}$$

where $\tilde{\mathbf{f}} = T(\mathbf{f}, \mathbf{p})$ is a transformed version of $\mathbf{f}$, and $\mathbf{w}_{i,j}$ is the weighting function (alpha map) containing values between 0 and 1. By expanding the brackets, (1) can be written

$$L(\mathbf{p}) = \sum_{i,j} \tilde{\mathbf{f}}^2_{i,j} \mathbf{w}_{i,j} - 2 \sum_{i,j} \tilde{\mathbf{f}}_{i,j} \mathbf{g}_{i,j} \mathbf{w}_{i,j} + \sum_{i,j} \mathbf{g}^2_{i,j} \mathbf{w}_{i,j}. \tag{2}$$

If our rigid-body transformation is limited to only translations, then $\tilde{\mathbf{f}}_{i,j}$ can be written $\mathbf{f}_{i-m,j-n}$, and the cost function is

$$\begin{aligned} L(\mathbf{p}) \;\; = \;\; & \sum_{i,j} \mathbf{f}^2_{i-m,j-n} \mathbf{w}_{i,j} - 2 \sum_{i,j} \mathbf{f}_{i-m,j-n} \mathbf{g}_{i,j} \mathbf{w}_{i,j} \\ & + \sum_{i,j} \mathbf{g}^2_{i,j} \mathbf{w}_{i,j}. \end{aligned} \tag{3}$$

The first and second summations can be written as convolutions, and it is these convolution terms that make this cost function expensive to evaluate. For example, when aligning two $N \times N$ images, a global search over all possible integer shifts would require $\mathcal{O}(N^4)$ operations. However, it is well known that a convolution can be evaluated using the Fast Fourier Transform (FFT) in $\mathcal{O}(N^2 \log N)$ operations. The method stems from the fact that a convolution of two signals is equivalent to performing an element-wise multiplication of their Fourier coefficients, and taking the inverse Fourier transform of the result.

In [13] it was shown that, in conjunction to finding the optimal translation, the globally-optimal linear intensity remapping could also be computed efficiently, thus finding values for m, n, s_0, and s_1 to minimize

$$\min_{m,n,s_0,s_1} \sum_{i,j} \left[s_0 + s_1 \mathbf{f}_{i-m,j-n} - \mathbf{g}_{i,j} \right] \mathbf{w}_{i,j}. \tag{4}$$

In other words, the best brightness and contrast adjustment can be determined simultaneously with the best translation. In [14], this method was generalized to include nonlinear intensity remappings (composed of linear combinations of user-specified basis functions, such as powers of $\mathbf{f}$ or trigonometric functions of $\mathbf{f}$). However, while these methods find the globally-optimal translation, they do not account for any rotation.

Building on theses methods, we now derive a way to include rotation. Consider the linear approximation to the motion transform,

$$T(\mathbf{f}, \mathbf{p}) \approx \mathbf{f} + \mathbf{J}\mathbf{p} \tag{5}$$

[1]From this point on, unless otherwise stated, we will use the word "image" to include higher-dimensional datasets as well.

where $\mathbf{J}$ holds the derivatives of $\mathbf{f}$ with respect to the motion parameters. Since we already know that the optimal translation can be found globally, we model only the rotation using the linear approximation: $R(\mathbf{f}, \theta) \approx \mathbf{f} + \mathbf{J}\theta$. Also including a linear intensity remapping, our cost function is

$$\begin{aligned} & \sum_{i,j} \left[s_0 + s_1 \left(\mathbf{f}_{i-m,j-n} + \mathbf{J}_{i-m,j-n} \theta \right) - \mathbf{g}_{i,j} \right]^2 \mathbf{w}_{i,j} \\ = & \sum_{i,j} \left[s_0 + s_1 \mathbf{f}_{i-m,j-n} + s_2 \mathbf{J}_{i-m,j-n} - \mathbf{g}_{i,j} \right]^2 \mathbf{w}_{i,j} \end{aligned} \tag{6}$$

where $s_2 = s_1 \theta$. Using the method outlined in [14], the optimal shift (m, n) and the optimal coefficients s_0, s_1 and s_2 can be found efficiently. Once found, we compute θ using $\theta = s_2 / s_1$.

3 Methods

We implemented our method in MATLAB (MathWorks Inc., Natick, Massachusetts). Given two images $\mathbf{f}$ and $\mathbf{g}$, and an alpha map $\mathbf{w}$ that specifies a region of interest (ROI), it computes the optimal rotation, translation and intensity remapping required to register the two images over the specified ROI.

The method was tested on a set of images obtained from the Visible Human Project (National Library of Medicine). An axial slice through the head of the Visible Male was chosen, for which we obtained PD-, T1- and T2-weighted magnetic resonance (MR) images. Originally, the three MR images were all registered. Three different alpha maps were designed, each containing a different size of ROI: a small ROI that covered approximately one quarter of the head, a medium ROI that covered half of the head, and a large ROI that covered almost the entire head. The image background was not included in any of the three ROIs. From this set of three MR images and three ROIs, a total of nine different image-ROI pairs could be used as $\mathbf{g}$-$\mathbf{w}$ pairs for registration (see equation (6)).

To objectively test the proposed registration method, a series of trials was set up using the MR images and alpha maps (ROIs). Fifty rigid-body transformations were randomly generated by choosing rotation angles uniformly from the range [-5°, 5°], and x- and y-shifts uniformly from the range [-50, 50]. These rigid-body transformations were applied to each of the three MR images and three alpha maps, creating 9 different $\mathbf{g}$-$\mathbf{w}$ pairs for each of the 50 transformations (for a total of 450 different $\mathbf{g}$-$\mathbf{w}$ pairs). Our experiments involved registering the original MR images with these moved $\mathbf{g}$-$\mathbf{w}$ pairs. The goal was to have the registration method determine the rigid-body transformation (ie. find the motion parameters) used to displace these $\mathbf{g}$-$\mathbf{w}$ pairs.

Each of the three original MR images was registered with each of the 450 displaced $\mathbf{g}$-$\mathbf{w}$ pairs, for a total of 1350 registration scenarios. Of those, 450 correspond to intra-modality (within-modality) scenarios, in which each of the PD-, T1- and T2-weighted MR images were registered with their corresponding displaced versions. The remaining 900 cases correspond to inter-modality (multi-modality) scenarios.

As an error measure, average pixel displacement was calculated for all of the pixels inside the ROI. Since we recorded the original motion parameters, the true shift was known and thus compared to the estimated shift obtained by our method. Smaller average pixel displacement indicates better registration.

For comparison, we also performed the same registrations using FSLs implementation of mutual information, FLIRT [8], [15], using 256 bins and trilinear interpolation.

4 RESULTS

Table I shows the average pixel displacement for intra-modality registration over the small, medium and large ROIs. The average pixel displacement for intra-modality tests was negligible: for small, medium, and large ROIs, the average pixel displacements were 0.0072, 0.0086, and 0.1746, respectively.

	Small ROI	Medium ROI	Large ROI
PD to PD	0.0077	0.0107	0.1604
T1 to T1	0.0058	0.0079	0.2146
T2 to T2	0.0081	0.0074	0.1483
Overall	0.0072	0.0086	0.1746

TABLE I

INTRA-MODALITY AVERAGE PIXEL DISPLACEMENT

Our method's performance in registering inter-modality images is displayed in Figure 1. For each of the ROI sizes, one histogram shows the distribution of the average pixel displacements for the corresponding 300 cases. Figure 2 displays histograms showing the average pixel displacement distribution for the mutual information method for the same dataset. Table II provides the comparison between the two methods.

ROI Size	Our Method	Mutual Information
Small	6.79	79.76
Medium	8.46	118.88
Large	7.53	112.32
Overall	7.59	103.65

TABLE II

INTER-MODALITY AVERAGE PIXEL DISPLACEMENT

5 DISCUSSION AND CONCLUSION

Our method efficiently evaluates the least-squares cost function for all possible integer shifts and rotations. It can be seen from our results that for intra-modality tests the registration was near perfect with the overall average displacement of 0.063 pixels.

The results listed in Table II show that our method considerably outperforms mutual information, the current standard for multimodal image registration. The difference between the average pixel displacements of the two methods is substantial. Using the linear approximation of rotation, our method produced and average error of 7.59 pixels. The average error for the mutual information method was 103.65 pixels. As depicted in Figure 1 and Figure 2, finding the optimal linear intensity remapping when registering two images using our method produced smaller registration errors than mutual information. However, intensity mapping between different modalities typically is not linear. Future work involves incorporating rotation and translation with higher-order intensity remappings. We expect that the estimated rotation and translation parameters obtained with the higher-order intensity remapping will result in better registration and even smaller average pixel displacement.

For mutual information test runs, parameter choices were not necessarily optimized. We did not endeavor to find optimal parameters, but rather made reasonable parameter choices for our test-bed of images used. The success of the mutual information method is hampered by the problems with local minima in the cost function as well as the lack of an efficient way to compute and optimize the cost function. Unlike our method, mutual information does not perform well when the small ROI is used.

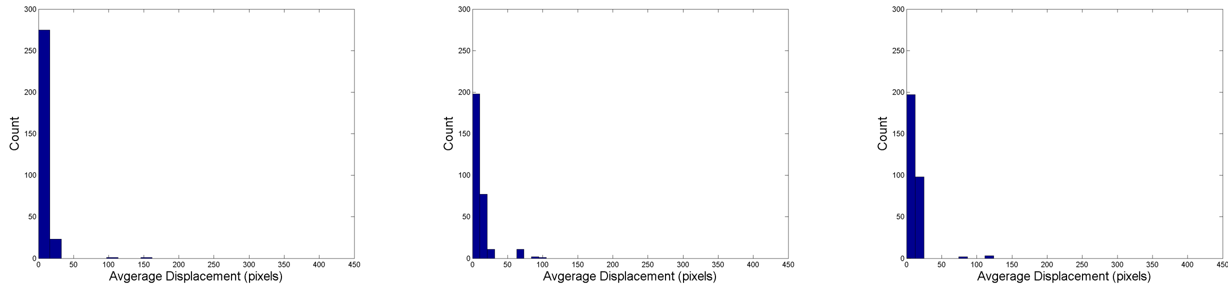
In contrast to other iterative least-squares methods that use local information for gradient descent, our method finds the globally optimal translation. The method is local in terms of determining an optimal angle, but global in terms of determining the optimal translation.

Since the rotation is approximated, our method currently works successfully on small rotation angles in the range of [$-15°$,$15°$]. Future work involves enabling registration with images rotated by a larger angle. Furthermore, spatial scaling is not part of the transformation we are considering here. Incorporating scaling into the least-squares cost function is the next logical step. We are investigating the possibility of integrating scaling into the cost function using a linear approximation.

Currently we are working on using this method to match dental x-rays for forensic human identification purposes. Given a postmortem dental x-ray with an alpha map (ROI), we compare it to set of antemortem dental x-rays in order to retrieve a closest match. Initial experimental results indicate that matching dental records using our method is viable.

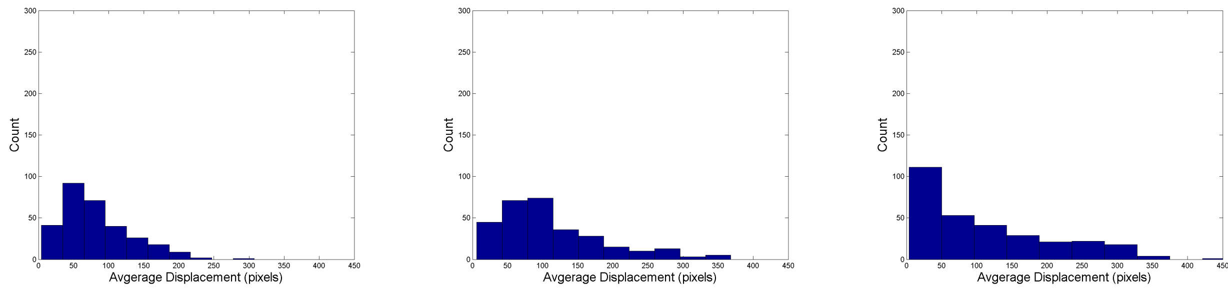
ACKNOWLEDGMENT

The authors would like to thank the National Library of Medicine for the Visible Human Project data. We are also grateful to the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Canadian Foundation for Innovation for funding of this project.



(a) Small ROI (b) Medium ROI (c) Large ROI

Fig. 1. Inter-modality average pixel displacement histograms for a small (a), medium (b), and large (c) ROI.



(a) Small ROI (b) Medium ROI (c) Large ROI

Fig. 2. Mutual information inter-modality average pixel displacement histograms for a small (a), medium (b), and large (c) ROI.

REFERENCES

[1] S. Grootoonk, C. Hutton, J. Ashburner, A. M. Howseman, O. Josephs, G. Rees, K. J. Friston, and R. Turner, "Characterization and correction of interpolation effects in the realignment of fMRI time series," *NeuroImage*, vol. 11, pp. 49–57, 2000.

[2] L. C. Maas, B. D. Frederick, and P. F. Renshaw, "Decoupled automated rotational and traslational registration for fMRI time series data: the DART registration algorithm," *Magn. Reson. Med.*, vol. 37, pp. 131–139, 1997.

[3] S. Thesen, O. Heid, E. Mueller, and L. R. Schad, "Prospective acquisition correction for head motion with image-based tracking for real-time fMRI," *Magn. Reson. Med.*, vol. 44, pp. 457–465, 2000.

[4] Y. Wang, R. Grimm, J. Felmlee, S. Riederer, and R. Ehman, "Algorithms for extracting motion information from navigator echoes," *Magn. Reson. Med.*, vol. 36, pp. 117–123, 1996.

[5] R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta, "Automated image registration: I. General methods and intrasubject, intramodality validation," *Journal of Computer Assisted Tomography*, vol. 22, pp. 141–154, 1998.

[6] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, Eds., 1995, pp. 263–274.

[7] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[8] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[9] C. Studholme, D. L. G. Hill, and D. L. Hawkes, "An overlap invariant entropy measure of 3D medical image alignent," *Pattern Recognition*, vol. 32, pp. 71–86, 1999.

[10] J. P. W. Pluim, A. Maintz, and M. A. Viergever, "Mutual information-based registration of medical images: A survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 986–1004, August 2003.

[11] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Transactions on Image Processing*, vol. 5, no. 8, pp. 1266–1271, 1996.

[12] A. J. Fitch, A. Kadyrov, W. J. Christmas, and J. Kittler, "Fast robust correlation," *IEEE Trans. Image Processing*, vol. 14, no. 8, pp. 1063–1073, August 2005.

[13] J. Orchard, "Efficient global weighted least-squares translation registration in the frequency domain," in *Image Analysis and Recognition (ICIAR)*, M. Kamel and A. Campilho, Eds. LNCS 3656, September 2005, pp. 116–124.

[14] ——, "Efficient least-squares multimodal alignment," *IEEE Transactions on Image Processing*, (submitted Jan. 2006), 2006.

[15] S. Smith, M. Jenkinson, M. Woolrich, C. Beckmann, T. Behrens, H. Johansen-Berg, P. Bannister, M. D. Luca, I. Drobnjak, D. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. Brady, and P. Matthews, "Advances in functional and structural mr image analysis and implementation as fsl," *NeuroImage*, vol. 23, no. S1, pp. 208–219, 2004.