GLOBALLY OPTIMAL MULTIMODAL RIGID REGISTRATION: AN ANALYTIC SOLUTION USING EDGE INFORMATION

Jeff Orchard

University of Waterloo
David R. Cheriton School of Computer Science
Waterloo, Ontario, N2L 3G1 Canada

ABSTRACT

Current multimodal registration methods almost always rely on local gradient-descent type optimization strategies. Such registration methods often converge to an incorrect local optimum, especially when the initial misregistration is large. There are monomodal image registration methods that employ global optimization techniques. This paper introduces the use of these global optimization methods for multimodal image registration. The goal is to robustly bring the images into close enough registration that a local optimization method could fine-tune the solution. The method proposed here is based on edge information extracted from the images. Positive results from a modest set of test cases suggests that this approach is promising.

Index Terms— multimodal, rigid, registration, Fourier, correlation

1. INTRODUCTION

One of the major stumbling blocks facing multimodal registration is the ability to compute the correct registration solution when the two images have a large initial misregistration. In this paper, we address the idea of computing the registration parameters for a multimodal registration scenario using only global operations. This type of globally-exhaustive registration has enjoyed much success in monomodal registration [1, 2, 3, 4], but has yet to be deployed effectively in a multimodal environment.

The vast majority of multimodality registration methods employ local optimization strategies to minimize (or maximize) the desired cost function. One of the most prominent cost functions is mutual information [5, 6] (as well as normalized mutual information [7]). However, these entropy-based cost functions rely on local gradient-decent-type strategies to find the optimal solution. There is no known way to efficiently compute the cost function for a large portion of the motion parameters space.

J. Orchard is financially supported by NSERC and CFI.

When two images are of the same modality, their Fourier Transforms (FT) encode many useful relationships regarding registration. For example, it is commonplace to decouple the rotation and translation parameters via the FT [3]. When an image rotates, its FT also rotates by the same angle. Hence, a rotation can also be assessed from the Fourier coefficients. Typically, one first finds the optimal rotation by resampling the Fourier coefficient magnitudes into polar coordinates, thereby reducing the search for the best angle to a search for the best shift along the θ-axis. Then, after applying the appropriate rotation to the image, one can use the phase of the resulting FT to derive the translation. This can be done in a few different ways. Probably the most popular is the phase correlation technique [1, 2, 4], whereby the corresponding Fourier coefficients of the two images are divided (canceling their magnitudes), leaving only their phase differences; the inverse FT of this phase image yields a single spike impulse whose location indicates the optimal shift of one image to align it with the other. Another approach is to compute either the cross-correlation [8] or sum of squared differences [9] cost function, using the maximum or the minimum (respectively) to indicate the correct shift. Both cost functions involve convolution-like operations that can be computed efficiently using the Fast Fourier Transform (FFT) [10].

In this paper, we propose the use of global optimization methods for multimodal registration. Our goal is to avoid the pitfalls of local optimization methods and transparently handle multimodal registration sceanrios with arbitrary initial misregistration, and bring the images into close enough alignment that one of the well-established local-optimization methods can dependably fine-tune the solution.

2. METHODS

2.1. Estimating the Rotation

We attack the rigid-body registration problem in two steps: rotation first, then translation. Figure 1 shows a schematic of the method for computing the optimal rotation angle. The process starts by extracting edge information from each of the two images using a Canny filter [11]. The success of the

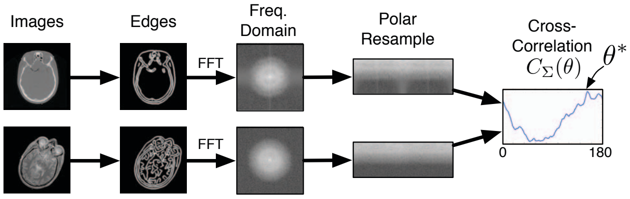


Fig. 1. Determining the optimal rotation

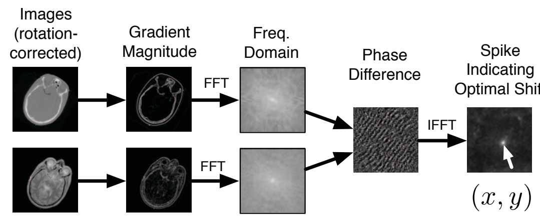


Fig. 2. Determining the optimal shift

method depends on the edges being reasonably similar between the two images. To find the rotation between these two edge images, we take their 2D FFT's and compute the magnitudes of their Fourier coefficients. If the two images have similar edges, then their Fourier Transforms will be approximately rotated versions of one another. A rotation is equivalent to a shift along the θ-axis of a polar-coordinates representation. We resample the two Fourier-magnitude images onto a polar grid. To assess the optimal shift along the θ-axis, we compute the cross-correlation between the two polar plots for each fixed radius. The cross-correlation between two signals f and g is

$$C(\theta) = \sum_n f_{n-\theta} g_n \, . \tag{1}$$

The cross-correlation can be computed efficiently using the FFT [8] by

$$C(\theta) = \mathcal{F}^{-1} \left\{ \overline{\mathcal{F}\{f\}} \mathcal{F}\{g\} \right\}_\theta \, ,$$

where $\mathcal{F}\{\cdot\}$ is the Fourier Transform, and $\overline{\mathcal{F}\{f\}}$ is the complex conjugate of $\mathcal{F}\{f\}$. Hence, we can compute the cross-correlation between signals of length N in $\mathcal{O}(N \log N)$ floating point operations, since that is the computational complexity of the FFT algorithm [10]. This approach can be considerably faster than the $\mathcal{O}(N^2)$ approach of directly computing (1) for every integer shift θ [12].

Once we have computed the cross-correlation $C_r(\theta)$ for each radius r, we combine them over all r using the weighted sum $C_\Sigma(\theta) = \sum_r r C_r(\theta)$, the rationale being that each cross-correlation function should be weighted according to its circumference in the Cartesian coordinate system. The θ-value that corresponds to the maximum value of $C_\Sigma(\theta)$ is our estimate for the rotation required to place our two original images in the same orientation. Because of the conjugate symmetry in the FT of a real-valued image, our cross-correlation function is actually a single $180°$ piece, repeated twice. Hence, the maximum of $C_\Sigma(\theta)$ occurs at two angles separated by $180°$. We carry both candidate rotations on to the next phase of the registration procedure.

2.2. Estimating the Translation

We compute the optimal shift for each of the two candidate angles, and choose the angle that yields a better overall match after the shift is applied. Thus, the following methodology applies to each angle. An overview of this process is shown in Fig. 2.

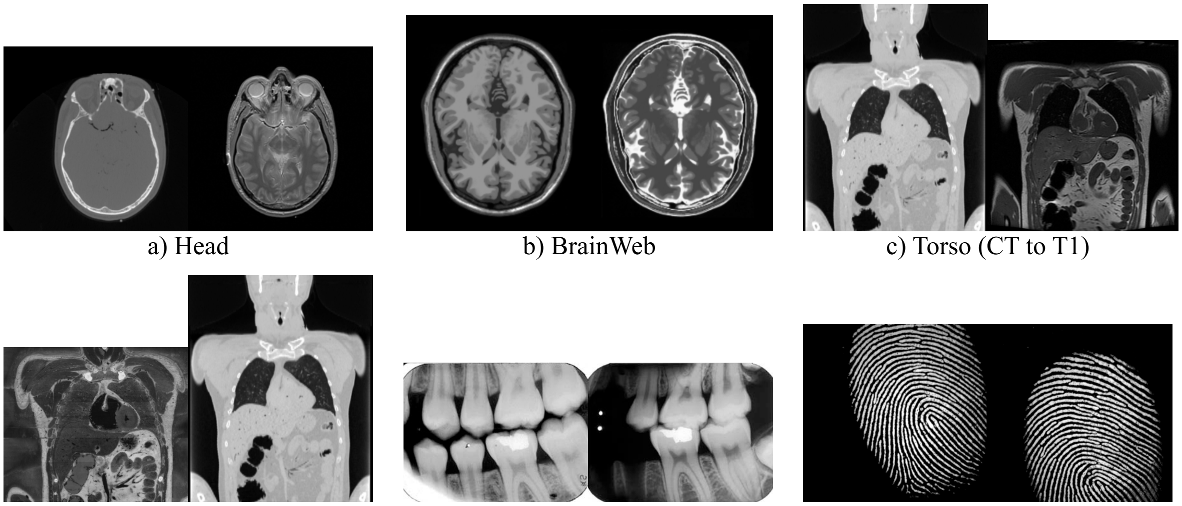
We use our angle to rotate our images into the same orientation. We then compute the magnitude of the gradient vectors for each image. Like the Canny edge images used to compute the optimal rotation, these gradient magnitude images contain edge information while suppressing most of the modality-specific pixel intensity information. We then employ the phase correlation method to find the best shift to align these two edge images. The procedure for computing the phase correlation function starts by taking the 2D FFT of each image. Then, each Fourier coefficient of one image is divided by the corresponding Fourier coefficient from the other image. The magnitudes of the resulting coefficients are all set to unity, and the inverse Fourier Transform of this phase image yields the phase correlation function. If the two images are shifted versions of one another, then the phase correlation function will have a single distinct spike. The position of the spike indicates the optimal shift.

This method is run on each of the two angles arising from the rotation estimation stage (the two angles differ by $180°$). The angle that yields a taller spike is chosen as the optimal rotation angle estimate. The location of that spike determines the optimal shift.

2.3. Experiments

The methodology outlined above was implemented in MATLAB (Mathworks Inc., Natick, Massachusetts). For lack of a better acronym, we will refer to our globally optimal edge-based registration method as GO-EDGE.

We performed registration experiments on the six image pairs depicted in Fig. 3. Some of the images, when zero-padded, contain a clear edge at the original image boundary. That is, some of the images have content that extends right to the image boundaries. These artificial edges could unfairly aid the registration process. To avoid this, images with these boundary edges were first multiplied by a Hann window [13] to taper the pixel intensities gently toward zero at the image boundaries (not shown in Fig. 3). The two images in each image pair were padded with zeros as necessary to make them the same size, followed by zero-padding on all four sides by



a) Head b) BrainWeb c) Torso (CT to T1) d) Torso (photo to CT) e) Teeth f) Fingerprints

Fig. 3. Image pairs used to evaluate the registration method. The "Head" and "Torso" images are from the Visible Human project. The "BrainWeb" images were acquired from the BrainWeb simulated MRI phantom [14]. The "Teeth" images are courtesy of Dr. David Sweet (Bureau of Legal Dentistry, University of British Columbia). The "Fingerprints" images are from the training dataset of the FVC2002 competition [15].

an additional 128 pixels to allow room for rigid-body transformations. Prior to computing the initial 2D FFTs at the beginning of the rotation-estimation stage, the images were zero-padded (in addition to the zero-padding described above) with 128 pixels on all sides to produce super-sampling in the frequency domain.

For each image pair, ten registration trials were created. Each trial consisted of randomly choosing a rotation angle (uniformly from 0° to 360°), as well as randomly choosing horizontal and vertical offsets (using a zero-mean Normal distribution with standard deviation of 20 pixels). Each image pair was initially registered, so the true registration parameters are known for each trial.

We ran our method on all 60 registration cases and compared our estimated motion parameters to the true motion parameters. If all three parameters were within 10 (degrees or pixels) of the true value, then the registration trial was considered a success. If any one of the three parameters was more than 10 (degrees or pixels) away from the true value, then the trial was considered a failure.

For comparison, we also registered all 60 trials using FLIRT [16], a state-of-the-art medical image registration application developed by Oxford's FMRIB group [17]. The cost function was set to the correlation ratio (CR). The developers of FLIRT have gone to great lengths to make it as robust as possible in the face of large initial misregistrations. However, FLIRT's optimization engine is essentially a local optimization technique built into a multiresolution framework, in addition to a rudimentary coarse parameter sampling. We developed a FLIRT schedule file (a script that dictates its optimization strategy) to mimic FLIRT's default 3D rigid-body registration behaviour in a 2D rigid-body context.

3. RESULTS AND DISCUSSION

The proportion of successful trials for both registration methods is reported in Table 1. Both methods were successful on all the Head and BrainWeb trials. FLIRT was also quite effective at registering the Torso CT-T1 image pair. However, FLIRT did not successfully register any of the remaining trials involving the Torso graylevel photo, the Teeth, or the Fingerprints (these image pairs are shown in Fig. 3). It is worth noting that FLIRT's four failed cases in the Torso CT-T1 image pair were close to the 10°/10-pixel criterion. In fact, FLIRT consistently computed the correct rotation for these trials, but consistently computed a slightly erroneous translation that sometimes resulted in success, and sometimes resulted in failure, depending on the orientation of the true solution. In general, FLIRT's failed cases were the result of repeatedly converging to a wrong, though consistent, solution.

Six of the 16 failed GO-EDGE cases involved rotation estimates that were approximately 180° away from the true rotation, suggesting that the method correctly assessed the angle, but the shift estimation stage yielded a higher correlation for the incorrect of the two angles. In total, FLIRT succeeded in 43% of the trials, while the GO-EDGE method succeeded in 73% of the trials.

Table 1. Registration success rate

Dataset	GO-EDGE	FLIRT
Head	100%	100%
BrainWeb	100%	100%
Torso (CT → T1)	100%	60%
Torso (photo → CT)	50%	0%
Teeth	40%	0%
Fingerprints	50%	0%

4. CONCLUSIONS

Based on the small set of test cases, the use of global optimization methods shows promise for multimodal rigid-body registration. Such global methods are almost entirely used for monomodality registration. However, we have shown that image edge information can be sufficient for multimodal registration. We feel that the use of global optimization techniques is key with these edge images; edge images are often quite sparse, which may spell trouble for a local optimization method. The use of global optimization methods avoids these pitfalls.

Future work includes investigating the use of phase information to choose which of the two candidate angles to use.

5. REFERENCES

[1] Hassan Foroosh, Josiane B. Zerubia, and Marc Berthod, "Extension of phase correlation to subpixel registration," *IEEE Trans. Image Processing*, vol. 11, no. 3, pp. 188–200, 2002.

[2] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," in *Proc. Int. Conf. Cybernetics and Society*, 1975, pp. 163–165.

[3] 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.

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

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

[6] William M. Wells III, Paul Viola, Hideki Atsumi, Shin Nakajima, and Ron Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–51, 1996.

[7] 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.

[8] James G. Pipe, "Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging," *Magn. Reson. Med.*, vol. 42, pp. 963–969, 1999.

[9] Jeff Orchard, "Efficient least-squares multimodal registration with a globally exhaustive alignment search," *IEEE Trans. Image Processing*, vol. (to appear), 2007.

[10] James W. Cooley and John W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Mathematics of Computation*, vol. 19, no. 90, pp. 297–301, 1965.

[11] John Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 6, pp. 679–698, 1986.

[12] Jeff Orchard, "Efficient global weighted least-squares translation registration in the frequency domain," in *ICIAR'05, Lecture Notes in Computer Science, LNCS*, Mohamed Kamel and Aurelio Campilho, Eds., September 2005, vol. 3656, pp. 116–124.

[13] Alan V. Oppenheim and Ronald W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, second edition, 1999.

[14] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, June 1998.

[15] "FVC2002, second international fingerprint verification competition," .

[16] M. Jenkinson and S. Smith, "A global optimization method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, 2001.

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