A Pattern Search Method for Image Registration

Hong Zhou¹ and Benjamin Ray Seyfarth²

¹ Saint Joseph College,
West Hartford, CT 06117
hzhou@sjc.edu

² The University of Southern Mississippi,
Hattiesburg, MS 39406, USA

Abstract. This paper presents a pattern search method for transformation function search in automatic image registration. In this search method, affine transformation parameters are represented as a 5-parameter vector and the search proceeds toward the direction resulting in higher similarity values between the reference and sensed images. Experiments show that this method can successfully find an optimal affine transformation function for edge images generated by computing the local standard deviation of the images. In addition, a series of overlapped aerial images of the Mississippi Delta area are successfully registered into a large reference image automatically by employing this search method.

1 Introduction

Image registration is a process of aligning two different images of the same object such that corresponding points in the two images represent the same physical location. The image to be registered is usually called the sensed image, while the image to which the sensed image is registered is commonly called the reference image. In general, image registration involves three steps including feature detection, feature matching and transformation function construction, and image transformation. Feature detection is a step to identify salient and distinctive objects. In the second step, feature matching and transformation function construction, the correspondence between the features detected in the sensed image and those detected in the reference image is established in terms of similarity measurement. Based on the correspondence, transformation functions (also called mapping functions) which are used to align the two images are constructed. In the third step, by means of the transformation function, the sensed image is transformed to register with the reference image.

The key work in image registration is to find the optimal transformation function f such that most of the points of the sensed image can be accurately mapped to the reference image by means of f [5]. However, reaching the optimal transformation function has never been a trivial task because of the large search space and possible image distortions. The search space is characterized by the transformation model underlying each registration process. Affine transformation is the most frequently used global transformation model [3, 4, 13, 15, 16], since it is reliable, efficient and can correct some global distortions. Global transformation methods are typically either a search for the allowable transformation which maximizes the used metric, or a search for the optimal

parameters of the transformation function [3]. Though most of the current research on search strategies are related to concepts of multiresolution, wavelet, relaxation matching, and Hausdorff distance [2, 3, 4, 6, 8, 10, 11, 16], direct search related techniques are still in active use because of their easy implementation and freedom from derivative computation [9, 12].

Direct search methods started from the work of Hooke and Jeeves [7]. The idea can be explained as the following. Starting from a base point p , repeatedly search for a new point p' from various directions until either a p' is found such that p' is a better point than p and therefore p is replaced with p' , or no better p' can be found. If no better p' can be found, the search is finished. One key step in direct search methods is how to select a new point. Based on the tactic of the new point selection, direct search methods are usually categorized into

1. pattern search methods,
2. simplex search methods,
3. and methods with adaptive sets of search directions.

In this work, a pattern search method that works with the affine transformation model is presented for image registration. It is demonstrated that with the classic similarity metric, normalized cross-correlation (NCC), this method could successfully find optimal affine transformation function parameters for edge-images generated by computing the local standard deviation of the images.

2 The Search Method

To precisely address the proposed search method, it is necessary to explain the edge-image used in the method first. The edge-image is generated by computing the local standard deviation of the image.

2.1 Local Standard Deviation for Edge Detection

Normalized cross correlation is one of the most used similarity metrics in image registration. However, edge-based correlation, which is computed on the edges extracted from the images rather than on the original images, is advantageous over the general correlation method in that it is less sensitive to the intensity differences between the reference and sensed images [1, 14, 16]. In the proposed search method, the NCC value is computed between two edge-images.

While there are a large number of edge detection methods, the local standard deviation (LSD) method in which edges are detected by computing the standard deviation inside a local small window shows some important properties and advantages:

1. With proper window sizes, LSD can be used for successful line edge detection. Sharp lines can be detected as line edges by LSD with small window sizes. Wider lines require larger window sizes.
2. LSD widens the detected edges. The width of the detected edge is proportional to the size of the local window used. This property plays an important role in the proposed search method.

2.2 The Search Method

An affine transformation is precisely a function of 6 parameters

$$\begin{aligned}x' &= a_0 + a_1x + a_2y, \\y' &= b_0 + b_1x + b_2y.\end{aligned}\tag{1}$$

An affine transformation without shear can also be written as:

$$\begin{vmatrix} x' \\ y' \end{vmatrix} = \begin{vmatrix} s_x & 0 \\ 0 & s_y \end{vmatrix} \begin{vmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{vmatrix} \begin{vmatrix} x \\ y \end{vmatrix} + \begin{vmatrix} t_x \\ t_y \end{vmatrix}.\tag{2}$$

Clearly, an affine transformation without shear can be fully determined by 5 parameters: s_x and s_y which are the x and y scale factors, t_x and t_y which are the x and y translations, and θ which is the rotation angle. These 5 parameters constitute a vector fully describing an affine transformation composed of rotation, translation and scaling. Each vector element determines the amount of transformation of a specific type. Thus, the search for an optimal affine transformation becomes the task of the search for an optimal vector function, and the key is the search direction. The search method can be detailed as the following:

1. Start with the affine transformation defined by the vector $A = \langle A_0, A_1, A_2, A_3, A_4 \rangle$. The values of A must be somewhat close to optimal and might be determined by tie points selected from each image or by other methods.
2. Start with a small *step* size.
3. For each element of A , determine whether increasing or decreasing that element by the *step* size improves the transformation. The reference image is transformed to match the sensed image and the NCC between the warped reference image and the sensed image is computed. Improvement is defined to be an increase in the NCC value. Record the amount of NCC improvement for this element and its direction.
4. Based on the amount of improvement and the direction for each vector element of A , construct a 5-dimensional vector $a = \langle a_0, a_1, a_2, a_3, a_4 \rangle$ that determines the search direction. The value of each element in a is proportional to the amount of NCC value improvement incurred by its corresponding element in A . The sum of the absolute values of the elements in a equals 1. However, if there is no element in A whose variation results in any improvement, *step* is divided by two and the process repeats from step 3 until either at least one element incurs some improvement or the *step* is smaller than a predefined limit. In case that *step* is smaller than this predefined limit, the search is finished.
5. The search starts toward the direction determined by the vector a . The vector A is recomputed, i.e. $A = A + \text{step} \times a$. If the newly computed A generates a better NCC value, it is kept and the old vector is discarded.
6. Repeat step 5 until it fails to generate a higher NCC value. Then *step* is divided by two. If *step* is smaller than a predefined limit, the search toward this vector direction is over and the search restarts from step 2. Otherwise, the search repeats from step 5 with the decreased *step*.

In brief, for the current optimal direction, the search continues until there is no gain in this direction. At this point, a new direction is computed and expressed in the vector

a. The search stops when there is no direction resulting in a larger NCC value. The computation of the vector direction can be done in other ways. The way presented here is a very simple method, in which each vector element is weighted equally.

The proposed search method is basically the pattern search method described by Hooke and Jeeves with some variations [7]. A critical factor in pattern search methods is the *step* value. Large *step* sizes may result in quickly finding the optimal solution, but may seriously degenerate the search result [9]. It turns out that for different image registration cases, the optimal *step* size should be different. As a general rule, the *step* size is small. However, a small *step* size is more likely to run into a local maximum when the starting point is relatively far from the optimal solution.

To successfully minimize the local maximum problem, the LSD edge detection technique is employed to generate line edges of different widths in the method. In the beginning when the search is relatively far from the optimal parameters, the LSD technique is applied with larger window sizes to generate wider edges. Images with wide line edges have less content details, thus the search is focused on major features of the image, i.e. the search is on a coarse level. As the solution approaches the optimum, the LSD processing is performed with smaller window sizes to generate sharp edges, i.e. more image details are taken into consideration. Using a large initial window size for LSD processing has another valuer, which can be explained in Fig. 1. In Fig. 1, (a) and (b)

Fig. 1. The significance of the width of line edges in image registration

are the reference and sensed images each with only one line. If the width of the detected line edges are small at the initial search step as shown by (c), the NCC metric between the warped reference image and the sensed image generates no significant result (0, in this case) to guide any proper search direction. However, if the width of the two line edges are large enough as shown in (d), then there is a meaningful NCC value that could be used to guide the next step search. As the search is closer to the optimal parameters as shown by (e) and (f), the width of the two line edges are cut down for finer alignment. It is important to observe the limit on the window sizes for LSD processing. Large window sizes can significantly diminish prominent features. They can make close edges indistinguishable. Thus, a proper starting LSD window size should be large enough to account for the distance between the initial warped reference image and the sensed image, but still leave separate edges distinguishable.

2.3 Refinement

Shear is an important component of affine transformation and can be used to correct some global distortions. An affine transformation without shear is not complete. Thus, shear is applied after the search is finished to refine the search transformation function parameters. As there exist both x and y shears, we take the shear as a vector of 2 elements and repeat the above process again specifically for the shear components.

The iterative refinement procedure is based on the fact that different search starting points usually lead to slightly different optimal solutions. In the iterative refinement procedure, after the first iteration of the search and shear refinement, the searched transformation parameters are slightly modified and another iteration of the search and shear refinement is carried out. However, in this iteration, the LSD is always processed with a small window size.

3 Experimental Results

In order to test the method and demonstrate its feasibility, the proposed search method was implemented and aerial images of the Mississippi Delta area were used as test images. The reference and sensed images were taken seven years apart and have significant differences. First, we artificially generated a number of rotated, translated, and scaled images from the original aerial images to perform some experiments to determine the impact of the initial LSD window size on the registration quality. The results showed that when the *step* sizes are small, proper LSD window sizes could significantly improve the registration quality. Therefore, in all the experiments presented below, the initial and final window sizes for LSD processing are 21×21 and 5×5 , the initial *step* values applied to rotation, x -translation, y -translation, x -scaling, and y -scaling are 0.2, 1.0, 1.0, 0.01, and 0.01 respectively. The similarity metric NCC was computed between the warped reference image and the sensed image with both images being LSD processed and smoothed. In addition, shear refinement is applied after the search was finished to improve the transformation.

Figures 2 and 3 show the registration results of two sets of aerial images of the Mississippi Delta area. Both figures are rotated 90° counterclockwise. In Fig. 2 from left to right, the first image is the reference image, the middle image is a combination of the reference image and the unregistered sensed image in such a way that the unregistered image lies in the middle surrounded by the reference image. The right image is also a combination of the reference and the registered sensed images arranged as the middle image. The purpose of such two combined images is to demonstrate the registration quality visually. Clearly, the proposed search method can render good registration quality for aerial images of the Mississippi Delta area.

We performed another set of experiments in which a series of the individual aerial sensed images were registered to a large map automatically based on the proposed search method. The Mississippi Delta area is primarily an agricultural area with many fields and ponds. To successfully monitor the variation of the ground cover in the Delta area, aerial images are taken periodically and these aerial images are registered to construct an integrated overview image of the area. These aerial images were taken in order. They have similar sizes, and are overlapped sequentially. We designed an algorithm

Fig. 2. Registration of a sample aerial image

Fig. 3. Automatic registration of 12 aerial images of the Mississippi Delta area

in which the first sensed image was manually aligned with the large reference image roughly well, then the proposed search method was applied to register it. After the first image was registered, based on the overlapping properties between the sensed images, the search starting point of the second sensed image could be obtained automatically, and the proposed registration procedure was applied to register the second image. This process was repeated for 12 partially overlapped aerial images. In addition, 3 levels of iterative refinement were applied to the registration process of the 12 aerial images, and the result is shown in Fig. 3. This result shows that the proposed search method can find an optimal transformation function for the registration of each individual aerial image in the automatic registration procedure.

4 Conclusion

The proposed search method is a pattern search method in which the affine transformation parameters are represented as a 5-parameter vector. Applied to edge images generated by LSD, the proposed search method is demonstrated to be able to find optimal parameters for affine transformation. We plan to extend the search method to higher-order polynomial transformation models in the future.

References

1. Anuta, P.: Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transform. *IEEE Transactions on Geoscience Electronics*. **8** (1970) 353–368
2. Baker, S., Matthews, I.: Equivalence and efficiency of image alignment algorithms. In *Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition*. (2001) 1090–1097
3. Brown, L.: A survey of image registration techniques. *ACM Computing Surveys*. **24** (1992) 326–376
4. Chalermwat, P.: High performance automatic image registration for remote sensing. PhD thesis, George Mason University. (1999)
5. Chmielewski, L., Kozinska, D.: Image registration. In *Proceedings of the 3rd Polish Conference on Computer Pattern Recognition Systems*, Poland. (2001) 163–168
6. Fonseca, L., Costa, M.: Automatic registration of satellite images. <http://www.inf.uni-konstanz.de/cgip/lehre/ss03-proj/papers/papers/FoCo97.pdf>
7. Hooke, R., Jeeves, T.: Direct search solution of numerical and statistical problems. *Journal of the Association for Computing Machinery*. **8** (1961) 212–229
8. Huttenlocher, D., Klanderman, G., Rucklidge, W.: Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **15** (1993) 850–863
9. Isaacs, A.: Direct-search methods and dace. Seminar Report 02401002, Indian Institute of Technology Bombay. (2003)
10. Kruger, S., Calway, A.: Image registration using multiresolution frequency domain correlation. In *British Machine Vision Conference*. (1998) 316–325
11. Lai, S., Fang, M.: Robust and efficient image alignment with spatially varying illumination models. http://www.cs.nthu.edu.tw/~lai/CVPR99_Lai_Fang.pdf
12. Lewis, R.: Direct search methods: then and now. ICASE Report 2000-06, Institute for Computer Applications in Science and Engineering, National Aeronautics and Space Administration Langley Research Center. (2000)
13. Image Fusion Systems Research: Transformation functions for image registration. http://www.imgfsr.com/ifsr_tf.pdf. (2003)
14. Wie, P., Stein, M.: A landsat digital image rectification system. *IEEE Transaction on Geoscience Electronics*. **15** (1977) 130–136
15. Wang, X., Feng, D., Jin, J.: Elastic medical image registration based on image intensity. In *Pan-Sydney Area Workshop on Visual Information Processing (VIP2001)*, Sydney, Australia. (2001) 139–142
16. Zitová, B., Flusser, J.: Image registration methods: a survey. *Image and Vision Computing*. **21** (2003) 977–1000