Automatic multi-resolution image registration based on genetic algorithm and Hausdorff distance

Famao Ye (叶发茂), Lin Su (苏 林), and Shukai Li (李树楷)

Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101

Received February 27, 2006

Image registration is a crucial step in all image analysis tasks in which the final information is gained from the combination of various data sources, and it is difficult to automatically register due to the complexity of image. An approach based on genetic algorithm and Hausdorff distance to automatic image registration is presented. We use a multi-resolution edge tracker to find out the fine-quality edges and utilize the Hausdorff distance between the input image and the reference image as similarity measure. We use wavelet decomposition and genetic algorithm, which combine local search methods with global ones balancing exploration and exploitation, to speed up the search of the best transformation parameters. Experimental results show that the proposed approach is a promising method for registration of image.

OCIS codes: 100.0100, 100.2000.

Registration is a fundamental task in image processing used to match two or more images taken at different times, from different viewpoints, and/or by different sensors. It geometrically aligns two images—the reference and input images. The present differences between images are introduced due to different imaging conditions. Typically, registration is required in remote sensing (multispectral classification, environmental monitoring, change detection, image mosaicing, creating super-resolution images, integrating information into geographic information systems (GIS)), cartography (map updating), computer vision (target localization, automatic quality control)[1], medicine[2] (combining CT and NMR data to obtain more complete information about the patient, monitoring tumor growth, treatment verification), and so on.

During the last decades, image acquisition devices have undergone rapid development and growing amount and diversity of obtained images invoked the research on automatic image registration. The current automatic registration techniques generally fall into two categories: area-based methods and feature-based methods. Area-based methods use the whole image content to estimate the transformation parameters. In contrast to area-based methods, feature-based methods utilize extracted features to estimate the registration parameters.

In this paper, we use wavelet-based method to extract large-scale edge and ignore the sharp textures. Hausdorff distance is taken as similarity measure because it has three advantages: relative insensitivity to small perturbations; simplicity and speed of computations; natural allowance for portions of one shape to be compared with another. We use wavelet decomposition and genetic algorithm, which combine local search methods with global ones balancing exploration and exploitation, to speed up the searching of the best transformation parameters.

Given two two-dimensional (2D) images, reference image I_0 and input image I_1 subject to registration, a transformation has to be found that correctly maps points of I_1 into corresponding points of I_0. There are numerous types of spatial transformations including both global (e.g., rigid, affine, and perspective), where all pixels are displaced according to a single rule, and local where displacement of a pixel depends on its location in the image[1]. When accurate registration is required, the process is often split into two steps. First, a global transformation is determined that takes into account most of the warping between the images being aligned. Then, the result is refined by computing local displacements. These two stages require very different types of registration algorithms[3], so we only focus on the first stage and consider only the "rotation-scale-translation" (RST) transformation:

$$\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = \begin{pmatrix} k\cos\theta & k\sin\theta & t_x \\ -k\sin\theta & k\cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix}, \quad (1)$$

where $\{t_x, t_y\}$ are the translations in x and y directions, θ is the rotation angle, and k is the isometric scaling.

Hausdorff distance has been widely applied in 2D image matching. The classical Hausdorff distance measure[4] between two point sets $A = \{a_1, \cdots, a_{N_a}\}$ and $B = \{b_1, \cdots, b_{N_b}\}$ of sizes N_a and N_b, respectively, is defined as

$$H(A,B) = \max\left(h(A,B), h(B,A)\right), \quad (2)$$

where $h(A,B) = \max\limits_{a\in A}\min\limits_{b\in B}\|a-b\|$, and $\|a-b\|$ is the Euclidean distance between the points of a and b.

The two functions, $h(A,B)$ and $h(B,A)$, represent the directed distances between two sets A and B. The directed function $h(A,B)$ identifies the points $a \in A$ that is farthest from any point of B, and measures the distance from a to its nearest neighbor in B. The Hausdorff distance, $H(A,B)$, is defined as the maximum of $h(A,B)$ and $h(B,A)$. Thus it measures the degree of mismatch between the two sets by measuring the distance of the point of A that is farthest from any point of B and vice versa.

However, the Hausdorff distance is very sensitive to even a single outlier point of A or B. To obtain more efficient image matching results, it is common to use the directed partial Hausdorff distance that is defined as

$$H^f(A,B) = f\text{th}\min\|a-b\|. \quad (3)$$

In the above definition, $f\text{th}g(x)$ $(x \in X)$ denotes the fth quantile value of $g(x)$ over the set X, for some values of f between 0 and 1[5]. For example, the 1st quantile value is the maximum and the 1/2th quantile value is the median. The physical explanation of $H^f(A,B)$ is as follows: there are at least a fraction f of the points in set B so that the distances from these points to their nearest neighbors in set A will not exceed $H^f(A,B)$.

The feature sets as input for the Hausdorff distance-based registration method are the fine-quality edges, and we combine gradient-based edge detection and wavelet-based multi-resolution edge tracking to extract edges[6]. The approach can preserve the large- resolution edges and ignore the sharp textures. The wavelet transform decomposes an image into different-scale and different-frequency subbands that can suppress the noise, and we produce multi-resolution shift-invariant gradient images from the high-frequency subbands. The proposed contextual-filter edge detector detects the reliable edges from the finest scale gradient images. Then, the edge tracker refines the detected edges, connects reasonable broken edge segments, and ignores the small edges by exploring multi-scale gradient images. The proposed approach can effectively reduce the noise influence and produces proper edges for the following applications.

Using Hausdorff distance as similarity measure, the problem of finding the correct transformation of Eq. (1) can be restated as the problem of finding the optimal vector of parameters $\mathbf{V}^* = \{t_x, t_y, \theta, k\}$ minimizing the measure of difference F that is the Hausdorff distance between the edge maps of images I_0 and I_1 here, so that

$$F(\mathbf{V}) > F^*(\mathbf{V}^*), \quad \text{for all} \quad \mathbf{V} \neq \mathbf{V}^*, \tag{4}$$

the measure of difference F can be defined as

$$F = H^f(t(E_1), E_0), \tag{5}$$

where E_0 and E_1 are the edge images from the feature extraction algorithm; $t(E_1)$ denotes the transformed version of E_1 by $\mathbf{V}$.

The difference F in Eq. (5) is a nonlinear function that has to be computed numerically for every trial parameter vector. In the case when the reference image I_0 contains many objects, and the object corresponding to the image I_1 is significantly distorted during the transformation, the problems of Eqs. (4) and (5) turn into large space, multimodal optimization problems. Traditional optimization methods fail to be trapped in one of the local minima of the nonlinear multimodal function F. Genetic algorithm[7], a powerful heuristic method from the toolkit of evolutionary computations, can be used to solve the problem. It is based on a natural concept that diversity helps to ensure a population's survival under changing environmental conditions. Genetic algorithm is a simple and robust method for optimization and search and has intrinsic parallelism.

According to terminology used in evolutionary computations, vector $\mathbf{V}$ is called a chromosome, and objective function F corresponding to $\mathbf{V}$ is called its fitness. The algorithm concurrently works with a set of chromosomes $\{V_i, i = 1, \cdots, P_V\}$, called a population, and attempts to find a chromosome $\mathbf{V}^*$ that has the minimal value of its fitness $F^*(\mathbf{V}^*)$. The initial population can be selected heuristically or randomly. For each generation, each candidate is evaluated and assigns the fitness value. The "selection" selects these candidates for the reproduction in the next generation based on their fitness values. The selected candidates are combined using the genetic recombination operation "crossover". The crossover operator exchanges portions of bit strings hopefully to produce better candidates with higher fitness for the next generation. Several crossover schemes have been suggested in the literature, such as single point, multipoint, or uniform crossover. The "mutation" is then applied to perturb the bits of the chromosome as to guarantee that the probability of searching a particular subspace of the problem space is never zero. It also prevents the algorithm from becoming trapped on local optima. The whole population is evaluated again in the next generation and the process continues until it reaches the termination criteria. The termination criteria may be triggered by finding an acceptable approximate solution, reaching a specific number of generations, or until the solution converges.

Figure 1 shows a flow diagram of our multi-resolution image registration based on genetic algorithm. It takes two images (reference and input) as input. The two images are decomposed using wavelet decomposition to produce multi-resolution edge images. The algorithm initially searches from the level that contains the smallest wavelet-compressed edge images toward the highest resolution wavelet-compressed edge images. For each level

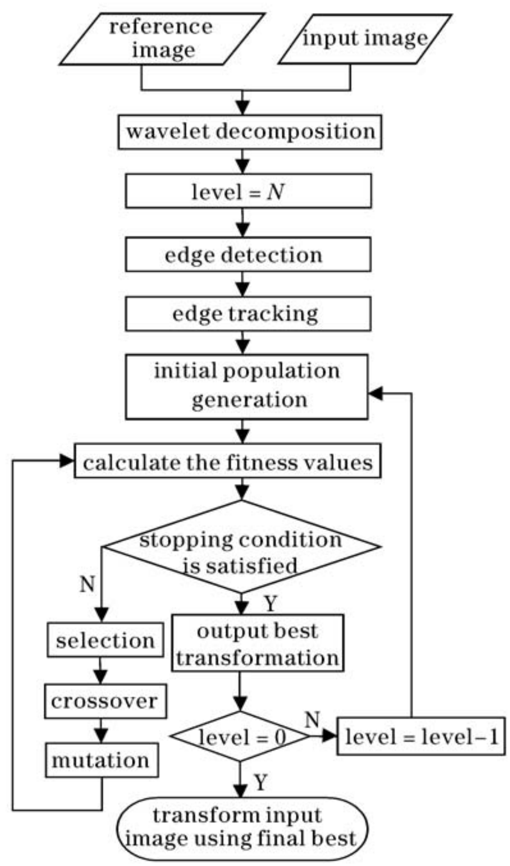

Fig. 1. Flow diagram of multi-resolution image registration based on genetic algorithm.

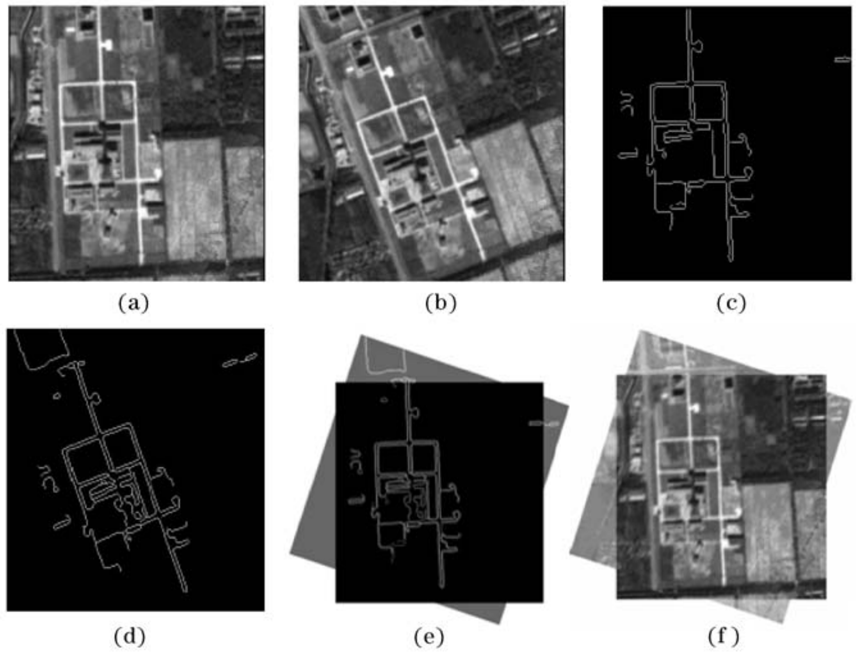

Fig. 2. (a) Reference image; (b) test image; (c) edge image of (a); (d) edge image of (b); (e) mosaic of (c) and (d); (f) mosaic of (a) and (b).

of multi-resolution, the best result found from the previous level is used as a center of the search.

In order to test the proposed approach based on Hausdorff distance and the multi-resolution evolutionary algorithm, experiments were conducted on 2D grayscale images. In the experiment, the input image and the reference image have been decomposed into three scale gradient images in edge extraction. We use a bit encoding scheme for chromosome string, and the length of chromosome string is 30. An 8-bit field is used to represent possible relative rotation of the input image to the reference image. 8 bits also are used to express translation in x-axis, 8 bits for the y-axis and 6 more for the isometric scaling. The population size is 30 and the iteration number is up to 200. Figure 2 shows two air images with much misalignment. The registration result is shown in Fig. 2(f) and final parameters are: $t_x = 40.9304$, $t_y = -72.2818$, $\theta = 18.2986$, and $k = 1.0024$. It shows that the proposed method can correctly register the two images with much misalignment.

In conclusion, an image registration based on Hausdorff distance and multi-resolution genetic algorithm is presented. We use a multi-resolution edge tracker to find out the fine-quality edges and take Hausdorff distance as similarity measure. We use wavelet decomposition and genetic algorithm to speed up the search of the best transformation parameters. The experimental results illustrate that the method provides excellent, robust performance in finding correct registrations with 2D images.

F. Ye's e-mail address is yfm_email@126.com.

References

1. B. Zitovt and J. Flusser, Image and Vision Computing **21,** 977 (2003).
2. X. Yang and J. Pei, Chin. Opt. Lett. **3,** 510 (2005).
3. L. Brown, ACM Comput Surv. **24,** 325 (1992).
4. D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, IEEE Trans. Pottern Anal. Machine Intell. **9,** 850 (1993).
5. W. J. Rucklidge, International Journal of Computer Vision **24,** 251 (1997).
6. M. Y. Shih and D. C. Tseng, Image and Vision Computing **23,** 441 (2005).
7. M. Zhou and S. D. Sun, *Genetic Algorithms: Theory and Applications* (in Chinese) (National Defence Industry Press, Beijing, 1999).