Progressive refinement for robust image registration

Li Song (宋 利), Yuanhua Zhou (周源华), and Jun Zhou (周 军)

Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai 200030

Received September 30, 2004

A new image registration algorithm with robust cost function and progressive refinement estimation is developed on the basis of direct method (DM). The robustness lies in M-estimation to avert larger local noise and outliers. Moreover, the progressive refinement model estimation under the multi-resolution framework, where the initial parameter values of coarse level are estimated by Mellin transformation, is adopted so as to get global optimization and reduce search space. Experiments show that the proposed algorithm greatly extends the scope of the conventional DM of image registration.

OCIS codes: 100.0100, 100.2000.

Image registration is the process of spatially matching two or more partially overlapping images so that corresponding pixels converge at the same physical region of the scene. It is very important in many vision applications, e.g., multi-source data fusion, time-varying image analysis for changes, image mosaic, and object recognition, etc.

In the past years, numerous image registration methods were contributed to improving the accuracy, generality, robustness, or computation speed[1,2]. Direct method (DM)[3], which recovers the unknown parameters directly from intensity of every pixel in the image, attains great attention in practice for high accuracy. However, challenges lie in the following issues: large displacement, noisy images, and outlier's rejection. At present, DM is capable of handling image displacement typically up to 10—15 percent of image size with coarse-to-fine refinement, so that an initial estimate is required for larger misalignments. The inevitable outliers from false correspondences or local noise points would greatly destroy the supposed parameters model.

In this paper, we extend DM in the two aspects. Firstly, we define a robust similarity cost function, which uses M-estimator to reject the outliers. Secondly, the progressive refinement model estimation under the multi-resolution framework is adopted to get global optimization and reduce complexity. Mellin transformation is utilized to get initial parameters on the basis of geometry analysis.

Image shots, which consist of planar objects or a camera rotation round its optical axis, are assumed to meet the following brightness constancy model with the projective transformation

$$I_1(x', y') = I_1(g(m); x, y) = I_2(x, y), \tag{1}$$

where $I_i(x, y)(i = 1, 2)$ is image intensity of position (x, y), $g(m)$ is some two dimensional (2D) mapping function depending on a parameter vector $\vec{m} = \{m_0, m_1, m_2, m_3, m_4, m_5, m_6, m_7\}$, which is to be determined by

$$\begin{bmatrix} sx' \\ sy' \\ s \end{bmatrix} = g(m) \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{2}$$

where s is a scale coefficient, (x, y) and (x', y') are position variables of the same object in their own image coordinates.

The sum of squared differences (SSD) is a most popular similarity measurement in vision with the tradeoff between accuracy and efficiency; however, its disadvantage is the frangibility to outliers. A generalized M-estimator is defined as a cost function, which is defined as

$$E = \sum_i \rho(r_i; \sigma), \quad r_i = I_1(x', y') - I_2(x, y), \tag{3}$$

where $\rho(r; \sigma)$ denotes robust estimation functions over the residuals r, with a given scale factor σ; and i is the index of corresponding pixel. In the paper, German-Mclure function is used so that weights decease rapidly with large residuals

$$\rho(r; \sigma) = r^2/(\sigma^2 + r^2). \tag{4}$$

Because an influence function for each data constraint to the cost function's solution is the derivative of the ρ function, the weight is defined as

$$w(r) = \rho'(r)/r = 2\sigma^2/(\sigma^2 + r^2)^2. \tag{5}$$

The plots of the ρ function and the weight for $\sigma = 1.0$ are shown in Fig. 1.

The optimal parameter $\vec{m}$ is attained by minimizing E, which can be solved according to an iterative weighting least-squares problem[4]

$$\min \sum_i w(r_i) r_i^2. \tag{6}$$

It is reasonable to model the residuals using a contaminated Gaussian distribution, where the components of

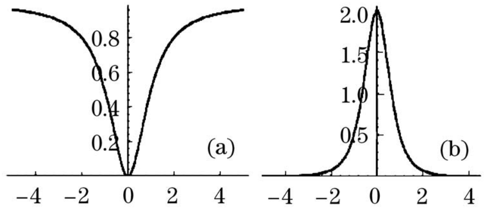

Fig. 1. German-Mclure function $\rho(r)$ (a) and the weight function $w(r)$ (b).

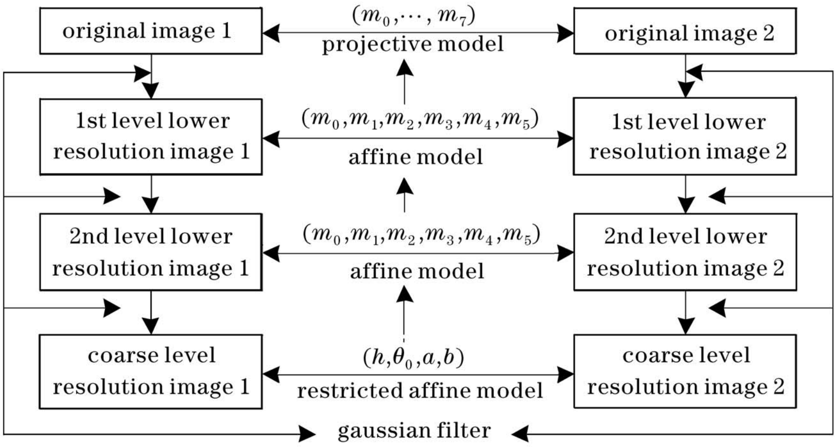

Fig. 2. Progressive refinement model estimation under four-level multi-resolution for image registration.

the residuals are the normal noise and outliers. Therefore, σ can be decided by[5]

$$\sigma = 1.4826 \cdot \mathrm{median}_i |r_i|. \tag{7}$$

We use Levenberg-Marquardt (L-M) algorithm to solve Eq. (6). Unfortunately, L-M is locally optimized and very sensitive to initial values, especially when images are taken from wide viewpoints. To get the global minimization and good stability, we employ progressive refinement model estimation under the multi-resolution framework. Specifically, we construct a pyramid of Gaussian-filtered and sub-sampled images, which consist of four-level of pyramid representation of image. Different parameter models are adopted at different levels. At the coarse level, a restricted affine model is used with initial parameters estimated by Mellin transform; at intermediate levels, rigor affine model is used; and at original resolution of image, a projective model is used. These upper levels utilize parameters which are estimated from the previous level and further refined via L-M algorithm described above. This scheme is depicted in Fig. 2 and derived from the following geometry analysis.

Because the projective model can be progressively simplified into two restricted models, the first is an affine model g_1, and the second is a model g_2 that consists of a rotation about the optical axis (θ_0), a translation (a, b), and a scale factor (h)

$$g_1 = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ 0 & 0 & 1 \end{bmatrix},$$

$$g_2 = \begin{bmatrix} h\cos\theta_0 & -h\sin\theta_0 & a \\ h\sin\theta_0 & h\cos\theta_0 & b \\ 0 & 0 & 1 \end{bmatrix}. \tag{8}$$

Then the subsequent assignment is initialized to estimate four parameters of g_2. It is noticeable that Fourier-Mellin transforms (FMT)[6], which is based on phase correlation and the properties of Fourier analysis so as not to be biased by outliers, becomes a well-known method for registering coplanar images. Two images (f_1, f_2) that differ with the translation displacement (a, b) presents the relationship

$$\frac{F_1(u,v)\cdot F_2^*(u,v)}{|F_1(u,v)\cdot F_2(u,v)|} = \mathrm{e}^{[-j2\pi(ua+vb)]}, \tag{9}$$

where F_i is a Fourier transform of $f_i (i = 1, 2)$, F_i^* is the complex conjugate of F_i. By taking inverse Fourier transform of the representation, we will attain a function that produces an impulse in (a, b).

Let f_1 and f_2 are related by g_2, and their Fourier magnitudes (M) in polar representation are related by

$$\begin{cases} M_1(\log\rho,\theta) = M_2(\log\rho - \log h, \theta - \theta_0) \\ \rho = (x^2+y^2)^{1/2}, \quad \theta = \tan^{-1}(x/y) \end{cases}. \tag{10}$$

By imposing phase correlation technique on M_1 and M_2, the scale h and the angle θ_0 can be obtained, so that the image with the higher resolution is scaled and rotated by amounts h and θ_0, moreover, the amount of translation displacement is still solved using phase correlation to transformed image pairs.

Among the various applications of image registration, we will mainly focus on image mosaic that is an automatic alignment of multiple images into a larger one.

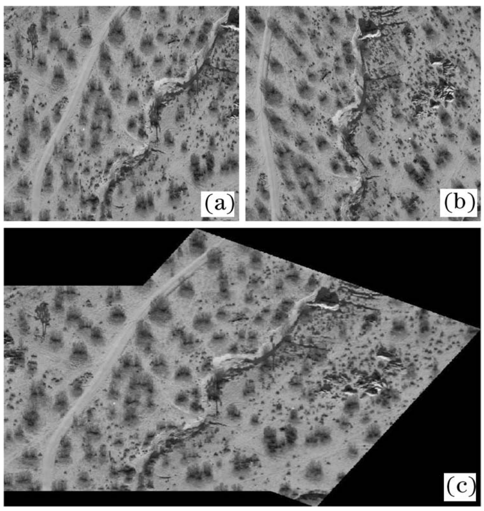

Fig. 3. Two SPOT satellite image (a) and (b) from Ref. [7] and the alignment image (c).

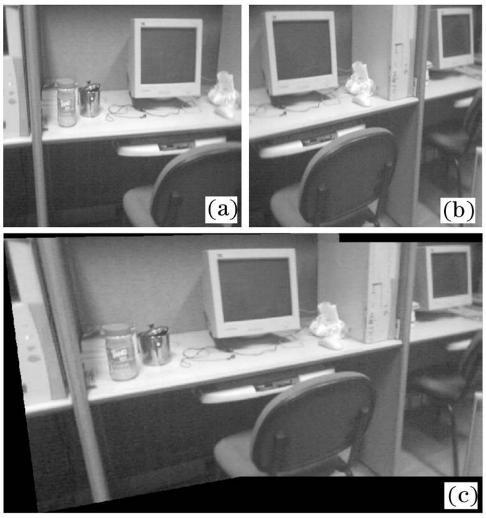

Fig. 4. Two indoors image (a) and (b) from our lab and the alignment image (c).

Although only overlapped areas meet the Eq. (1), our algorithm can get good results because FMT is very robust to shift, rotation and scale change and L-M is locally converging optimization. Figure 3 shows an example of alignment of two satellite images[7] with dimension 512 × 512, whose geometric variation is quite large. Conventional DM algorithms cannot deal with them because initial parameter values of model are far away from zero. The registration result of the proposed algorithm is shown in Fig. 3(c) and final parameters are: $m_0 = 0.85081$, $m_1 = -0.49326$, $m_2 = 300.61$, $m_3 = 0.48933$, $m_4 = 0.86303$, $m_5 = -143.82$, $m_6 = -0.000024$, $m_7 = 0.0000063$.

Figure 4 is another example of indoor images taken by approximately rotating video camera; the overlap of two images with dimension 640 × 480 is only 50% and rotation is about 20°, scale is about 1.3. Figure 4(c) shows the alignment of second with first one and the final parameters are: $m_0 = 1.2668$, $m_1 = 0.14183$, $m_2 = -400.9$, $m_3 = -0.024675$, $m_4 = 1.027$, $m_5 = -8.5873$, $m_6 = -0.000387$, $m_7 = 0.000054$.

In this paper, we have described a new image registration algorithm with robust cost function and progressive refinement estimation. Experimental results show that the proposed algorithm can greatly extend the scope of classical direct image registration method.

This work was supported by the National Natural Science Foundation of China under Grant No. 69905003. L. Song's e-mail address is songli@qantsoft.com.

References

1. L. G. Brown, ACM Computing Surveys **24,** 325 (1992).
2. B. Zitova and J. Flusser, Image and Vision Computing **21,** 977 (2003).
3. R. Szeliski, IEEE Computer Graphics and Applications **16,** 22 (1996).
4. Z. Zhang, Image and Vision Computing Journal **15,** 59 (1997).
5. P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection* (John Wiley & Sons, New York, 1987) p.329.
6. B. S. Reddy and B. N. Chatterji, IEEE Trans. Image Processing **5,** 1266 (1996).
7. http://vision.ece.ucsb.edu/registration/satellite/.