An Improved Optical Flow Method for Image Registration with Large-scale Movements

XIONG Jing-Yi[1] LUO Yu-Pin[1] TANG Guang-Rong[1]

Abstract In this paper, an improved optical flow method for image registration is proposed. It is novel in the way that it improves the optical flow method with an initial motion estimator: extended phase correlation technique (EPCT), using merits of the latter to compensate deficiencies of the former. In a more detailed manner, it can be said that the optical flow method can reach the sub-pixel accuracy and calculate complex distortion patterns like chirping and tilting but is weak with large-scale movements. Because EPCT covers measurements of large translations and rotations with pixel level accuracy and is efficient in the calculating load, it can be treated as a good initial motion estimator for optical flow method. Tests have proved that this improved method will significantly enhance the registration performance, especially, for images with large-scale movements and robust against random noises.

Key words Image registration, improved optical flow method, motion estimator, extended phase correlation technique (EPCT)

Image registration, a fundamental image processing problem, is the process of overlaying a sequence of images of the same scene taken at different times, from different viewpoints, and/or by different sensors. During this process, coordinate transformations are calculated so that images taken from the same static scene are related. Image registration has been widely used as an intermediate step in many research fields, such as super-resolution, panorama mosaics, medical photos analysis, etc.

Generally speaking, two categories of registration algorithms exist: feature-based and nonfeature-based. In the former category, there are algorithms using low-level features like edges and corners, and high-level features, such as identified objects, or relations between features[1]. While in the nonfeature-based category, there are algorithms using frequency domain information[2], and differential optical flow equation method[3−5].

For the optical flow, people find it is on the basis of the Taylor expansions and differential theory, and thus, it is weak to estimate large-scale movements between images. It was proposed by Gibso in 1950, and that the optical flow method is sensitive to noises because it is on the basis of differential technology. Some filters (both high-pass and low-pass) are used to reduce this bad effect. Moreover, Gaussian pyramids are used to construct different levels of resolution and incrementally accumulate the track of each level using the law of composition[5]. This is useful for enhancing the ability to estimate large-scale movements, but it is still not enough. Some ideas have been reported to further improve the registration by using an initial motion estimator providing a rough input for following calculation[6−7]. Some works have been done based on this, like log-polar mapping relied on nonlinear least square iterative optimization algorithm in [8], however, this method has a heavy calculating load. And phase correlation method relied on nonrigid optical flow estimation in [9], which is only specified in indocyanine green angiography (ICGA) fundus without rotations.

On the basis of the previous work, a new separate motion pre-estimator, EPCT (Extend phase correlation technique), is regarded as a more suitable choice in this paper. The proposed approach takes great advantage of the properties of EPCT-measurements of large translations and rotations, and calculating efficiency, so that the increase of workload is a linear growth of the image size. Although EPCT only reaches pixel level accuracy and can not deal with more complex distortions like keystoning or chirping, these deficiencies will be refined by the optical flow method later.

Another work done is defining a way to gauge the registration accuracy and to revise the loop time in optical flow from fixed to self-adaptive according to accuracy. Thus, it will save time and effort in calculation.

This paper is organized as follows. Section 1 outlines the principles of the projective model, optical flow method, as well as EPCT, then, describes how to improve optical flow method by using EPCT as a processor to estimate an initial input. Section 2 presents some registration results using the improved optical flow method and compares the results of EPCT with that of the original optical flow method. In Section 3, conclusions and some research perspectives are given.

1 Principles

1.1 Projective model

The projective model is chosen to describe movements between two images. It is an eight-parameter-equation pair dealing with rotation, translation, scaling, chirping, and tilting.

$$x' = \frac{m_0x + m_1y + m_2}{m_6x + m_7y + 1}, \quad y' = \frac{m_3x + m_4y + m_5}{m_6x + m_7y + 1} \tag{1}$$

It can be rewritten as a matrix

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{2}$$

where $x' = u/w, y' = v/w$.

In this model, m_2 and m_5 stand for translations in horizontal and vertical directions separately, m_0, m_1, m_3, and m_4 for scaling and rotation, and m_6 and m_7 are in charge of chirping and keystoning effects. An example is listed in Fig. 1 (see next page) for details of every parameter.

1.2 Optical flow method

In the optical flow assumption, for each point (x, y) in frame t, there will be a corresponding point in frame $t+\Delta t$, which means

$$E(x, y, t) = E(x + \Delta x, y + \Delta y, t + \Delta t) \tag{3}$$

Apply Taylor expansions to the right side of (3), we get the two-dimensional optical flow constraint:

$$\frac{\partial E_t}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial E_t}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\mathrm{d}E_t}{\mathrm{d}t} = 0 \tag{4}$$

Received May 14, 2007; in revised form August 9, 2007
1. Department of Automation, Tsinghua University, Beijing 100084, P. R. China
DOI: 10.3724/SP.J.1004.2008.00760


(a) (b)


(c) (d)


(e) (f)

(g)

Fig. 1 Relationship between parameters and image motions
((a) The original frame, $m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; (b) Horizontal translation, $m = \begin{bmatrix} 1 & 0 & 20 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; (c) Vertical translation, $m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 20 \\ 0 & 0 & 1 \end{bmatrix}$; (d) Chirping effect, $m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.0015 & 0 & 1 \end{bmatrix}$; (e) Keystoning effect, $m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0.0015 & 1 \end{bmatrix}$; (f) Scaling, $m = \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; (g) Rotation, $m = \begin{bmatrix} \cos 30° & \sin 30° & 0 \\ -\sin 30° & \cos 30° & 0 \\ 0 & 0 & 1 \end{bmatrix}$)

Define cost function as

$$\mathcal{E}_{\text{flow}} = \sum_x (u_m e_x + v_m e_y + e_t)^2 \tag{5}$$

where $u_m = x' - x$, $v_m = y' - y$, $e_x = \frac{\partial E_t}{\partial x}$, $e_y = \frac{\partial E_t}{\partial y}$, and $e_t = \frac{\mathrm{d}E_t}{\mathrm{d}t}$.

Minimize the cost function so that it satisfies the optical flow constraint as much as possible. The smaller the value of $\mathcal{E}_{\text{flow}}$, the better the registration. More details can be found in [5].

Because optical flow method is on the basis of Taylor expansions, and $\frac{\mathrm{d}x}{\mathrm{d}t}$ and $\frac{\mathrm{d}y}{\mathrm{d}t}$ are substituted by $x' - x$ and $y' - y$ in calculation, it is weak in estimating large-scale movements. In some circumstance, it might even lead to failures. If there is a motion estimation for the optical flow method as an initial input to turn the larger-scale movements into smaller-scale movements, it will be beneficial for registration effect. This initial estimation does not need to be precise because optical flow is good at accuracy and will refine its results, and the less cost the preprocessor, the better the algorithm. A good choice for this preprocessor is EPCT according to our previous analysis in the introduction section.

1.3 EPCT

Assume that $f_2(x, y)$ is a translated and rotated replica of $f_1(x, y)$, such that

$$\begin{aligned} f_2(x, y) = f_1(&x \cos\theta_0 + y \sin\theta_0 - x_0, \\ &-x \sin\theta_0 + y \cos\theta_0 - y_0) \end{aligned} \tag{6}$$

According to the properties of Fourier transform, (6) can be transferred from space domain to the frequency domain as

$$\begin{aligned} F_2(\xi, \eta) =& \mathrm{e}^{-\mathrm{j}2\pi(\xi x_0 + \eta y_0)} \times \\ & F_1(\xi \cos\theta_0 + \eta \sin\theta_0,\ -\xi \sin\theta_0 + \eta \sin\theta_0) \end{aligned} \tag{7}$$

Let M_1 and M_2 be the magnitudes of F_1 and F_2. From (7), because the magnitude of $\mathrm{e}^{-\mathrm{j}2\pi(\xi x_0 + \eta y_0)}$ is 1, we have

$$M_2(\xi, \eta) = M_1(\xi \cos\theta_0 + \eta \sin\theta_0,\ -\xi \sin\theta_0 + \eta \cos\theta_0) \tag{8}$$

Mapping the magnitude from Cartesian coordinates to polar coordinates, we have

$$M_1(\rho, \theta) = M_2(\rho, \theta_0 - \theta) \tag{9}$$

where $\rho = \sqrt{\xi^2 + \eta^2}$, $\theta = \tan^{-1}\left(\frac{\eta}{\xi}\right)$.

After the operations mentioned above, this problem has become a typical one for PCT in polar coordinates by which θ_0 is calculated. By rotating one of the frames according to θ_0, there will be only translations between images. PCT should be applied again in Cartesian coordinates so that translations in vertical and horizontal directions can be calculated.

1.4 The improved optical flow algorithm

Since EPCT and optical flow method have been discussed in previous sub-sections, now let us come to the improved version of optical flow method. First, EPCT is used to get the rotation and translation between images, from which θ_0 and x_0, y_0 are estimated. Because the matching images are of the same size, no scaling effect is supposed. Chirping and keystoning effects are also unknown at the moment so m_6 and m_7 are both set to be 0. Then, an initial pixel-accurate value of the projective model matrix is

$$m = \begin{bmatrix} \cos\theta_0 & \sin\theta_0 & x_0 \\ -\sin\theta_0 & \cos\theta_0 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

It is coarse but good enough for the optical flow method as an initial input, as compared with $m = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right]$, which means the matching images are identical. From experiments the optical flow method is sensitive to the initial m, and a better estimation will greatly improve the ultimate result. So using the result of EPCT instead of an identity matrix will significantly enhance the registration. An overview of this improved algorithm is shown in Fig. 2.

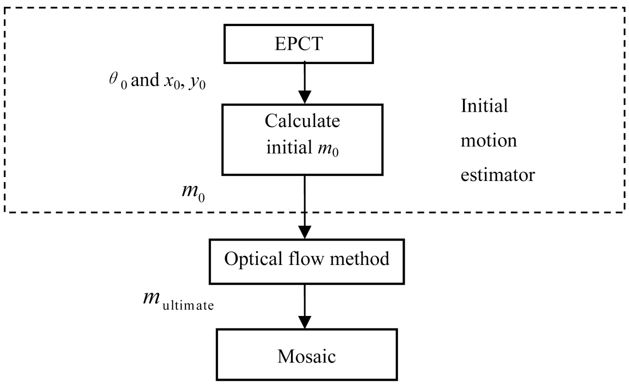

Fig. 2 Overview of the registration algorithm

2 Implementation and experimental results

2.1 Implementation

First, in the preprocessor of EPCT where θ_0 and x_0, y_0 are estimated, the initial input for optical flow method is calculated as in (10).

Second, to make the optical flow method more effective, a two-layer loop structure is applied, as shown in Fig. 3. Suppose the size of images being registered is $M \times N$. The first layer is Gaussian pyramids of reference and current frames from coarse to fine (usually three-level, with the sizes of images $\frac{M}{4} \times \frac{N}{4}$, $\frac{M}{2} \times \frac{N}{2}$, $M \times N$, respectively), and the second layer is repeating the optical flow calculating process in each Gaussian pyramid level, using the projective model parameters obtained last time as the initial input for next loop.

To make the loop times of the second layer self-adaptive, to save time and workload, two criteria are applied as follows.

Criterion 1.

$$loopTime_{\max} > \alpha$$

Criterion 2.

$$\frac{OMSE(this\_time) - OMSE(last\_time)}{OMSE(last\_time)} \leq \beta$$
$$\text{for } \gamma \text{ consecutive times}$$

In Criteria 1 and 2, α, β, and γ are parameters defined by user according to the accuracy they require. And overlapped mean square error (OMSE) is calculated as

$$OMSE = \sum_{(x,y)\in\Gamma} \frac{\left[\dfrac{r(x,y) - c(x,y)}{255}\right]^2}{N} \tag{11}$$

In (11), Γ is the overlapped part of images, $r(x, y)$ and $c(x, y)$ are the pixel-values of point (x, y) in mosaic of the reference and current frame, respectively, and there are totally N points existing in the overlapped part.

Either Criterion 1 or 2 will make the loop in the second layer stop and turn to next Gaussian pyramid level in layer one. Compared with fixed loop, there are two advantages for using these two criterias to make loop stopping decisions:

1) Different images might have diverse motion estimation convergence speeds. Thus, self-adaptability will enhance the calculation efficiency.

2) Through making adjustments to α, β, and γ, it is easy to set different registration accuracy. To some extent, by enlarging α and γ, or decreasing β, registration accuracy will be boosted.

2.2 Experimental results

The performance of this improved optical flow method is tested on different sets of images. Random noises are also added to test the algorithm's robustness. Two experimental results are shown. Fig. 4 is on the simulated data, whereas Fig. 5 is on the real data.

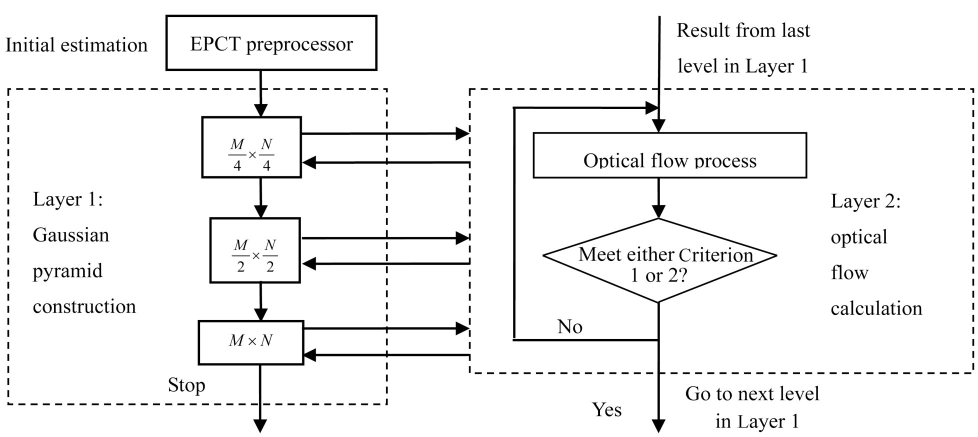

Fig. 3 Two-layer loop structure of optical flow calculation

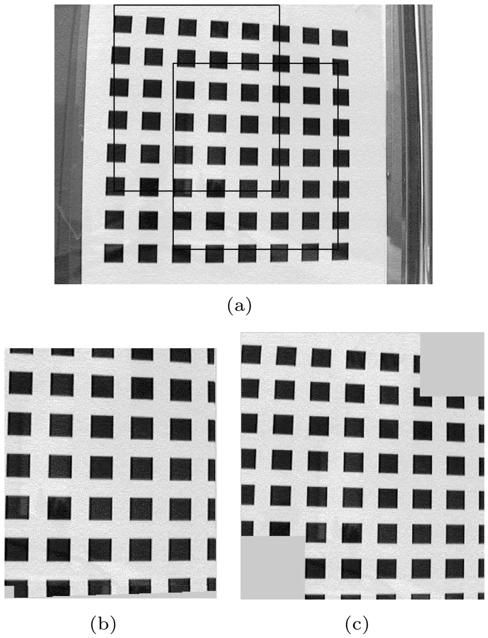

Fig. 4 Grids ((a) The two rectangles marked are the reference and current images, respectively; (b) The mosaic result of optical flow method; (c) The mosaic result of improved optical flow method)

The registration difference between optical flow method and the improved version is shown in Fig. 4. In Fig. 4 (a), an artificial movement of 100 pixels is made both vertically and horizontally between the reference and current images marked by the rectangles. It is easily seen that optical flow registration of Fig. 4 (b) is an error, and the improved optical flow of Fig. 4 (c) is much better. Actually, the motion estimation from Fig. 4 (c) is 100.0224 pixels in both directions, which is an excellent estimation.

In Fig. 5, (c) from traditional optical flow method is clearly a failure. And (d) from EPCT is properly right, yet some slight mismatches exist, marked by a black rectangle, and more details can be checked in an expanded version (e). The best mosaic is (f) without any perceptible mismatch, for details see (g). Make a careful comparison between (e) and (g), it is easy to find that (g), the registration result of improved optical flow, is much better than (e) from optical flow. In (h), some random noise is added to the current and reference frames, and the improved method works pretty well as we have previously analyzed.

Some quantitative evaluations are shown in Tables 1 and 2.

Table 1 $OMSE$ of registration in Fig. 5

Method	Optical flow	Improved optical flow
Frame $OMSE$	0.0323	0.0017

$OMSE$ is calculated as (11), and the value of OMSE measures the overall difference of overlapped parts of reference and current frames in their mosaic. The smaller the values, the better the registration results, and vice versa.

Table 2 Time consuming of registration in Fig. 5

Method	Fixed loop	Self-adaptive loop
Time consumed (s)	45.916	28.721

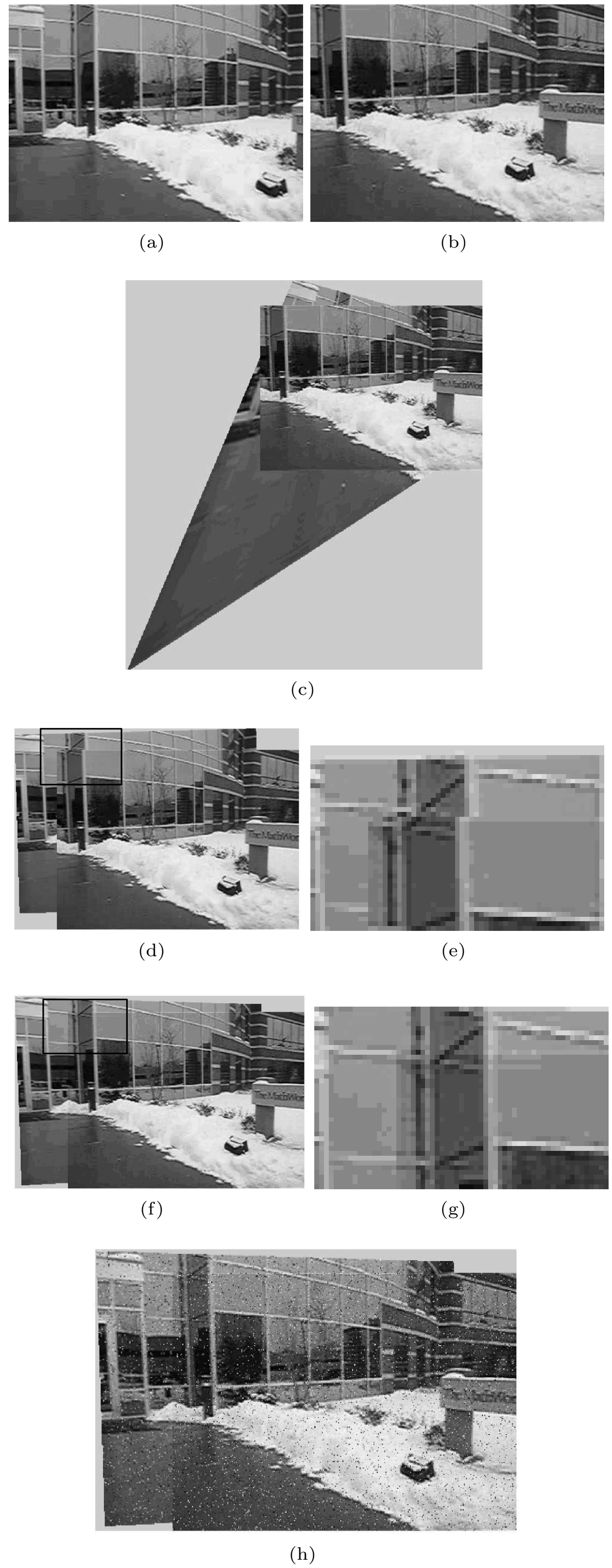

Fig. 5 Snow ((a) The current frame; (b) The reference frame; (c) Result of optical flow method; (d) Result of EPCT; (e) Expanded version of the part marked by a black rectangle in (d); (f) Result of improved optical flow; (g) Expanded version of the part marked by a black rectangle in (f); (h) Result of improved optical flow added with salt and pepper noise)

In fixed loop times category, the time is fixed to 10, whereas in self-adaptive loop times, $\alpha = 10$, $\beta = 0.1$, and $\gamma = 2$. Registration results are identical in these two situations, but time consuming of the latter is nearly half of the former.

Table 3 Time consuming distribution in different part in Fig. 5 (Note that self-adaptive optical flow is abbreviated as SAOF.)

Part	EPCT	SAOF	Total
Time consumed (s)	8.692	20.029	28.721
Percentage (%)	30.3	69.7	100

Comprehensively speaking, from experiment results listed above and some other tests which have not been listed in this paper, it is implied that the improved optical flow method using EPCT as a preprocessor does refine the mosaics in some cases, compared with EPCT or traditional optical flow method, especially, for those with large-scale movements. Moreover, the two-layer self-adaptive loop structure will significantly reduce calculation time and thus, enhance efficiency.

Some random noises are added to images to check if the improved algorithm keeps working well. And it does. The reason for this merit has been explained in the introduction.

3 Conclusion

In this work, an improved optical flow method, using EPCT as a preprocessor and a two-layer structure in the optical flow calculation, is presented to deal with image registration, especially, for those with large-scale movements.

This improved algorithm has the merits of both methods. There are two steps in the calculating process. First, EPCT is used to handle large-scale movements to get a coarse estimation of projective model parameters. After that, optical flow is used to refine the result from the preprocessor, from pixel to subpixel accuracy.

Experiments show that it is effective to those images with large-scale movements, which cannot be handled by traditional optical flow, and to a robust algorithm against random noises. However, it is also found that for some images containing too large-scale movements or few clear edges in the contents, this improved method reaches a poor result (although the same for EPCT or optical flow). Some investigation should be done to improve the algorithm to estimate a broader range of movements in the future.

References

1 Brown L G. A survey of image registration techniques. *ACM Computing Surveys*, 1992, **24**(4): 325−376

2 Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 1996, **5**(8): 1266−1271

3 Tsai C J, Galatsanos N P, Katsaggelos A K. Optical flow estimation from noisy data using differential techniques. In: Proceedings of International Conference on Acoustics, Speech, and Signal Processing. Phoenix, USA: IEEE, 1999. 3393−3396

4 Bereziat D, Herlin I, Younes L. A generalized optical flow constraint and its physical interpretation. In: Proceedings of Conference on Computer Vision and Pattern Recognition. Hilton Head Island, USA: IEEE, 2000. 487−492

5 Mann S, Picard R W. Video orbits of the projective group: a simple approach to featureless estimation of parameters. *IEEE Transactions on Image Processing*, 1997, **6**(9): 1281−1295

6 Wu Jian-Ning, Feng Zong-Zhe, Guo Bao-Long. An image mosaic method based on PCT combined with optimization. *Microelectronics and Computer*, 2006, **23**(1): 117−120 (in Chinese)

7 Ren X F, Tang G R. Improving video resolution by image mosaics. In: Proceedings of World Academy of Science, Engineering and Technology. Istanbul, Turkey: WASET, 2005. 138−140

8 Li Zhong-Xin, Mao Yao-Bin, Wang Zhi-Quan. A method of image mosaicing using log-polar coordinate mapping. *Journal of Image and Graphics*, 2005, **10**(1): 59−63 (in Chinese)

9 Zhou Yong-Jin, Tang Li-Rong, Li Jun-Bo, Wan Ming-Xi, Zhang Peng. ICGA fundus image mosaicing based on phase-correlation method and nonrigid optical flow estimation. *Chinese Journal of Scientific Instrument*, 2002, **23**(6): 600−603 (in Chinese)

XIONG Jing-Yi Received her bachelor and master degrees both from the Department of Automation at Tsinghua University in 2005 and 2007, respectively. Her research interest covers image registration and video mosaics. Corresponding author of this paper.
E-mail: xjy@mails.tsinghua.edu.cn

LUO Yu-Pin Professor at Tsinghua University. He received his bachelor degree from Hunan University in 1982, and master and Ph. D. degrees both from Nagoya Institute of Technology, Japan in 1987 and 1990, respectively. His research interest covers image processing, graph theory, and artificial intelligence.
E-mail: luo@mail.tsinghua.edu.cn

TANG Guang-Rong Professor at Tsinghua University. He received his bachelor degree from Tsinghua University in 1970. His research interest covers computer applications and image processing.
E-mail: tanggr@mail.tsinghua.edu.cn