

基于块匹配的鲁棒视频图像镶嵌算法

李忠新¹⁾ 刘光杰²⁾ 茅耀斌²⁾ 王执铨²⁾

¹⁾(南京理工大学机械工程学院, 南京 210094) ²⁾(南京理工大学自动化系, 南京 210094)

摘要 为了获取宽视野的场景表示,提出了一种基于块匹配的视频图像镶嵌算法,该算法首先采用基于相位相关的块匹配方法估计出视频图像间的运动矢量场,并剔除其中由于图像噪声或运动物体的遮挡而导致外点运动矢量,然后根据图像的运动矢量场确定出图像子块之间的对应点对,进而利用得到的对应点对迭代求解图像间的变换模型参数以实现视频图像的自动镶嵌。针对真实场景的视频图像序列进行实验,获得了较好的镶嵌结果,表明了该算法的有效性。

关键词 视频图像镶嵌 块匹配 相位相关

中图分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2006)10-1468-06

Robust Video Mosaicing with Block Matching Method

LI Zhong-xin¹⁾, LIU Guang-jie²⁾, MAO Yao-bin²⁾, WANG Zhi-quan²⁾

¹⁾(School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094)

²⁾(Department of automation, Nanjing University of Science and Technology, Nanjing 210094)

Abstract A robust image mosaic method is proposed to construct a panoramic view of the scene from video sequence. Firstly, block matching method based on phase correlation is employed to estimate the motion vector field of video frames. After removing the outlier motion vector caused by the image noise or moving object occlusion, the corresponding point pairs between image blocks can be determined by the dominant motion vectors, and then the parameters of projective transformation model can be iteratively estimated with the obtained correspondences to implement the video mosaicing. All the work of video mosaicing is automatically completed without manual manipulation. Experimental results using video sequence with real scene have demonstrated the availability of the method.

Keywords video mosaicing, block matching, phase correlation

1 引言

视频图像镶嵌是一种将从真实世界获取的视频图像序列通过图像配准和融合获取一幅宽视场、高分辨率的全维场景图像(Panorama)的技术,它提供了一种有效的且完整的场景表示方法,使人们能更加客观形象地认识和理解真实世界。近年来,视频图像镶嵌技术在摄影测量学、计算机图形学、图像处理 and 计算机视觉等学科领域中得到了广泛的应用,如大型航空照片和卫星遥感图像的镶嵌、环境监测、

虚拟场景的构建、视频压缩、视频检索与浏览等等^[1]。

图像镶嵌过程主要包括图像配准和图像融合两个步骤,其中如何实现图像的配准是进行图像镶嵌的关键问题,其实质就是寻找图像之间的变换模型参数。Szeliski等人^[2]利用Levenberg-Marquardt非线性优化算法最小化图像间灰度差的平方和(sum of square difference)来估计图像间的变换模型参数,虽然具有较高的准确度,但计算量大且易陷入局部极值。Mann等人^[3]利用图像光流场求最小均方差的方法求解变换模型参数,从而实现无特征点对

收稿日期:2005-06-23; 改回日期:2005-10-11

第一作者简介:李忠新(1976~),男,讲师。2004年于南京理工大学获控制科学与工程专业博士学位。主要从事图像处理、虚拟现实等方面的研究。E-mail:lxznj@sohu.com

应的图像镶嵌,能够快速有效地获得图像配准参数,但当图像间存在较大的位移时,将得不到正确的结果。Traka 等人^[4]提出了一种图像镶嵌算法,通过块匹配或角点匹配的方法获得图像间的对应点,最后利用 M-估计的方法求解图像间的变换模型参数进行图像镶嵌,得到了的较好镶嵌结果。本文提出了一种新的图像镶嵌算法,首先利用基于相位相关的块匹配方法估计出图像的运动矢量场,与原有基于灰度的块匹配方法相比大大减少了计算量,同时根据图像运动矢量的直方图剔除大部分与图像主运动不一致的运动矢量(文中称之为外点运动矢量),然后利用图像的运动矢量场确定图像间的对应关系,最后通过迭代求解图像间的变换模型参数进行图像的镶嵌。具体流程如图 1 所示。

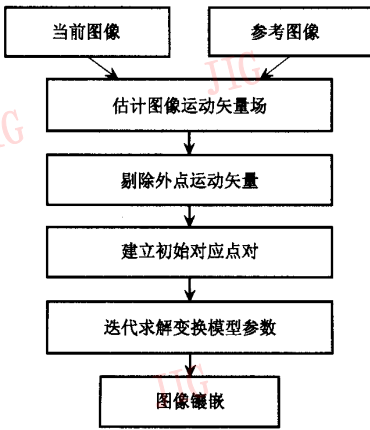


图 1 图像镶嵌流程
Fig. 1 Steps of image mosaicing

2 图像间的变换模型

对于同一视点拍摄的场景图像,文献[5]指出相邻图像间的变换关系可以采用如下的 8 参数投影变换模型来描述:

$$x'_i = Mx_i \text{ 或}$$

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \\ m_7 & m_8 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (1)$$

式中, M 为投影变换矩阵, $m_i (i = 1, 2, \dots, 8)$ 为投影变换参数; $(x_i, y_i, 1)$ 、 $(x'_i, y'_i, 1)$ 分别为相邻两幅图像间对应点 x_i, x'_i 的齐次坐标。

若已知图像间的对应点对 (x_k, x'_k) , $(k = 1, \dots,$

$n, n \geq 4)$, 则求解投影变换模型参数的优化问题可描述为

$$\hat{M} = \arg \min_M \sum_{k=1}^n \|x'_k - Mx_k\|^2 \quad (2)$$

3 图像运动矢量场的计算

为了确定图像间的对应关系,将块匹配算法与相位相关法^[6]结合起来,首先将图像划分为 $n \times n$ 的子块(n 一般取为 16),然后两幅图像的子块之间运用相位相关法,计算出各子块的运动矢量,从而获得整幅图像的运动矢量场,最后利用图像的运动矢量场来确定图像间的对应点。由于相位相关法与频域中图像的振幅无关,仅与所有频率的相位差有关。与原有的直接基于图像灰度的块匹配方法相比,在图像存在灰度变化,以及图像具有大量重复性内容(如草地、森林)的情形下,相位相关法依然可以得到正确的运动矢量,而块匹配等方法往往得不到正确的结果。此外,利用现有的快速傅里叶变换算法进行块匹配也大大提高了计算效率。

3.1 基于相位相关的块匹配算法

相位相关法最早是由 Kuglin 和 Hines 在 1975 年提出来的,是一种基于傅氏功率谱的频域相关技术。该方法只利用了互功率谱中的相位信息进行运动估计,对图像间的亮度变化不敏感,具有一定的抗干扰能力,而且所获得的相关峰尖锐突出,位移检测范围大,具有较高的匹配精度。设 $f_k(x, y)$ 为当前图像帧, $f_{k+1}(x, y)$ 为下一帧图像,两者间的位移量为 $(\Delta x, \Delta y)$, $F_k(\xi, \eta)$ 、 $F_{k+1}(\xi, \eta)$ 分别为 $f_k(x, y)$ 、 $f_{k+1}(x, y)$ 的傅里叶变换,即有

$$f_k(x, y) = f_{k+1}(x - \Delta x, y - \Delta y) \quad (3)$$

根据傅里叶变换的平移性质,对式(3)两端进行傅里叶变换后可得

$$F_k(\xi, \eta) = e^{-j2\pi(\xi\Delta x + \eta\Delta y)} |F_{k+1}(\xi, \eta)| \quad (4)$$

两幅图像的标准互功率谱为

$$cps(\xi, \eta) = \frac{F_{k+1}(\xi, \eta) F_k^*(\xi, \eta)}{|F_{k+1}(\xi, \eta) F_k(\xi, \eta)|} \quad (5)$$

式中, $F_k^*(\xi, \eta)$ 表示 $F_k(\xi, \eta)$ 的复共轭。

将式(4)代入式(5)得

$$cps(\xi, \eta) = \frac{|F_{k+1}(\xi, \eta)| |F_{k+1}(\xi, \eta)| e^{j2\pi(\xi\Delta x + \eta\Delta y)}}{|F_{k+1}(\xi, \eta)| |F_{k+1}(\xi, \eta)|} = e^{j2\pi(\xi\Delta x + \eta\Delta y)} \quad (6)$$

对式(6)进行傅里叶逆变换可以得到相位相关函数

表达式

$$cps(x, y) = \delta(x - \Delta x, y - \Delta y) \quad (7)$$

由上式可见相位相关函数是一个脉冲函数,它只在位置 $(\Delta x, \Delta y)$ 处不为零,而在其他各处为零,这个位置正是要计算的位移量。

利用相位相关法计算图像运动矢量场的具体算法如下:

(1) 将第 k 和 $k+1$ 帧图像按上述方法划分为 32×32 的子块 ($k=1, 2, 3, \dots$);

(2) 计算第 k 和 $k+1$ 帧图像中对应子块的 2 维离散傅里叶变换;

(3) 按式(5)计算图像子块的互功率谱相位;

(4) 计算互功率谱相位的 2 维离散傅里叶逆变换;

(5) 检测得到的相位相关函数的最大峰值所处的位置,该位置即为对应块平移量,从而可以得到图像的运动矢量场。

在具体实现过程中,为了避免块与块之间运动的丢失,准确地估计出子块与子块之间的运动,以原有的 16×16 子块为中心,向周围进行扩展得到 32×32 的子块进行相位相关计算,如图 2 所示。从图中可以看出,若采用尺寸大小为 16×16 的子块进行相位相关计算,则图像子块之间的重叠区域比较小,图像子块之间的相关性也比较小,此时往往得不到正确的结果。而采用尺寸大小为 32×32 的子块进行计算时,图像子块之间的重叠区域扩大,增加了图像之间的相关性,进一步提高了相位相关法估计结果的准确性。采用上述方法进行相位相关计算时,图像子块之间是相互重叠的,其所能检测的运动矢量范围是 $[-16, 16]$ 。

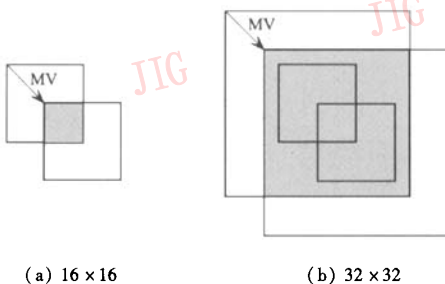


图 2 图像子块的划分

Fig. 2 The partition of image block

3.2 外点运动矢量的剔除

实际上,由于图像中运动物体的遮挡和图像噪

声的影响,计算出来的运动矢量场中可能会存在与图像主运动不一致的运动矢量,直接利用最小二乘法计算变换模型参数会导致较大的误差。所以在求解投影变换模型参数时,选择图像间的对应点对必须事先剔除那些与图像主运动不一致的运动矢量,使用与主运动一致的运动矢量来确定图像间的对应点对。

采用文献[7]提出的根据水平和垂直运动矢量直方图的 1/4 值点和 3/4 值点位置的算法,剔除大部分与主运动不一致的运动矢量,具体算法如下:

(1) 建立图像运动矢量场中水平运动矢量的直方图 $H(v_x)$, 记其总数为 N ;

(2) 计算 $H(v_x)$ 的 1/4 值点 s , 满足

$$\sum_{i=-M_{vx}}^{s-1} h(i) < \frac{N}{4}; \quad \sum_{i=-M_{vx}}^s h(i) \geq \frac{N}{4}$$

其中 M_{vx} 为水平运动矢量的最大可能值;

(3) 计算 $H(v_x)$ 的 3/4 值点 t , 满足

$$\sum_{i=-M_{vx}}^{t-1} h(i) < \frac{3}{4}N; \quad \sum_{i=-M_{vx}}^t h(i) \geq \frac{3}{4}N$$

(4) 计算水平运动矢量分布区间 $[r_1, r_2]$,

$$r_1 = \frac{3s}{2} - \frac{t}{2}, \quad r_2 = \frac{3t}{2} - \frac{s}{2};$$

(5) 将水平运动矢量不在区间 $[r_1, r_2]$ 的运动矢量看作外点矢量加以滤除;

(6) 对集中垂直运动矢量重复上述操作;

(7) 返回水平和垂直方向的运动矢量均在给定区间内的运动矢量。

对图像运动矢量场剔除外点运动矢量后,用保留下来的运动矢量来确定图像帧之间的对应点对,根据这些对应点进行投影变换参数的初始估计。

4 变换模型参数的迭代算法

在剔除外点运动矢量之后,图像运动矢量场中基本不含有与主运动截然不同的运动矢量,但仍然含有少量与主运动不完全一致的运动矢量,这是因为图像运动矢量的直方图只是反映了运动矢量的大小分布,但不能反映运动矢量的几何分布。例如图中,最上行中间的运动矢量明显与其相邻的运动矢量不同,但因其与图像底部的运动向量相近,采用上述基于运动矢量直方图的方法并不能有效的剔除,若不考虑这种情况,直接利用剔除外点运动矢量后的图像运动矢量场来确定图像间对应点的对应关

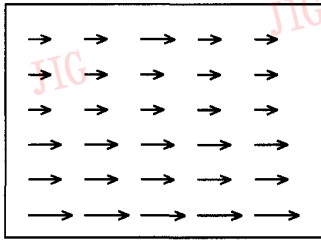


图 3 运动矢量的不一致性

Fig. 3 The inconsistency of motion vector

系,则同样会导致计算出来的变换模型参数不准确。为此,采用迭代估计的方法求解图像间的变换模型参数。

假设当前图像和参考图像分别为 I_{cur} 、 I_{ref} , 迭代求解变换模型参数具体步骤如下:

(1) 令投影变换模型参数的初始值 $M_0 = \mathbf{I}$, 迭代次数 $i = 1$, $I_{cur}^{(i)} = I_{cur}$, 其中 \mathbf{I} 为 3×3 单位矩阵。

(2) 利用相位相关法计算当前图像 $I_{cur}^{(i)}$ 和参考图像帧 I_{ref} 间的运动矢量场, 再采用 3.2 节介绍的方法剔除大部分的外点运动矢量, 然后按下式确定当前图像 $I_{cur}^{(i)}$ 和参考图像帧 I_{ref} 间对应点的坐标

$$\mathbf{x}'_k = \mathbf{x}_k + \Delta\mathbf{x} \quad (8)$$

其中, $\Delta\mathbf{x}$ 由剔除外点运动矢量后的运动矢量场来确定。

(3) 根据所确定的对应点对 $(\mathbf{x}_k, \mathbf{x}'_k)$, ($k = 1, \dots, n$), 采用最小二乘法求解如下优化问题

$$M_i = \arg \min_M \sum_{k=1}^n \|\mathbf{x}'_k - M\mathbf{x}_k\|^2 \quad (9)$$

(4) 计算由图像运动矢量场确定的对应点与通过上一步得到投影变换参数计算出的对应点之间的距离, 此处称之为残差。对于第 k 个对应点, 其残差

$$\mathbf{r}_k = \mathbf{x}'_k - M_i \mathbf{x}_k \quad \text{即} \quad \begin{bmatrix} r_{k1} \\ r_{k2} \end{bmatrix} = \begin{bmatrix} x'_k - \frac{(m_1^{(i)} x_k + m_2^{(i)} y_k + m_3^{(i)})}{(m_7^{(i)} x_k + m_8^{(i)} y_k + 1)} \\ y'_k - \frac{(m_4^{(i)} x_k + m_5^{(i)} y_k + m_6^{(i)})}{(m_7^{(i)} x_k + m_8^{(i)} y_k + 1)} \end{bmatrix} \quad (10)$$

计算残差的均值

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \frac{1}{n} \sum_{k=1}^n \mathbf{r}_k = \begin{bmatrix} \frac{1}{n} \sum_{k=1}^n r_{k1} \\ \frac{1}{n} \sum_{k=1}^n r_{k2} \end{bmatrix} \quad (11)$$

计算残差的方差

$$\sigma_1^2 = \frac{1}{n-1} \sum_{k=1}^n (r_{k1} - \mu_1)^2$$

$$\sigma_2^2 = \frac{1}{n-1} \sum_{k=1}^n (r_{k2} - \mu_2)^2 \quad (12)$$

若 $|r_{k1}| > \sigma_1$ 或 $|r_{k2}| > \sigma_2$, 则从图像运动矢量场中剔除对应的运动矢量, 同时剔除相应的对应点对。

(5) 根据剩余下来的对应点对重新估计投影变换参数, 即求解下述优化问题

$$M_i = \arg \min_M \sum_{k=1}^{n-m} \|\mathbf{x}'_k - M\mathbf{x}_k\|^2 \quad (13)$$

式中, m 表示在上一步骤中去掉的对应点对的个数。

(6) 利用重新估计得到的投影变换参数 M_i 将图像 $I_{cur}^{(i)}$ 进行图像 Warping 得到新的图像 $I_{cur}^{(i+1)}$ 。

(7) 计算在第 i 次迭代和第 $i-1$ 次迭代之间投影变换模型参数中的最大变化量, 若最大变化量小于给定的阈值, 或迭代次数大于给定的次数, 则停止迭代, 投影变换模型参数最终的估计值为

$$\hat{M} = M_i M_{i-1} \cdots M_0 \quad (14)$$

若最大变化量大于给定的阈值, 则 $i = i + 1$, 转第 2 步继续迭代。

5 实验结果与分析

首先, 从一段定点拍摄视频图像序列中提取出相邻的两幅图像 (第 20、21 帧), 如图 4 所示, 然后利用基于相位相关的块匹配算法计算出这两幅图像间的运动矢量场, 同时在实验中也给出了采用三步快速搜索块匹配算法得到的图像间的运动矢量场, 如图 5 所示。

从图中可以看出采用基于相位相关的块匹配算法得到的图像间的运动矢量场比较平滑, 与主运动不一致的运动矢量比较少, 有利于图像间变换模型参数的估计, 相比之下三步快速搜索块匹配算法得到的运动矢量场含有较多的与主运动不一致的运动矢量, 这对于估计图像间的变换模型参数是不利的, 同时在相同的运行环境下 (P III 550、256M SDRAM、Matlab6.5), 基于相位相关的块匹配算法花费的计算时间为 0.972s, 而三步快速搜索块匹配算法花费时间为 1.772s, 可见基于相位相关的块匹配算法大大提高了计算效率。

考虑到图像场中外点运动矢量对图像间变换参数估计的影响, 以及计算效率等因素, 采用基于相位相关的块匹配方法来进行图像运动矢量场的计算。在计算出图像间的运动矢量场后, 再利用 3.2 节介绍的方法剔除图像运动矢量场中的外点运动矢量,



图 4 原始视频图像

Fig. 4 Original video images

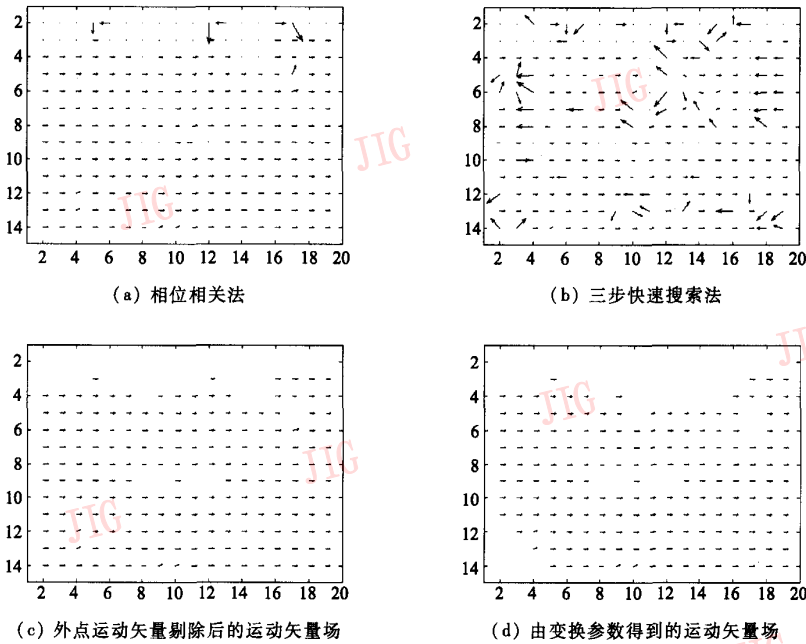


图 5 视频帧间的运动矢量场

Fig. 5 The motion vector field between video frames

结果如图 5(c) 所示, 与图 5(a) 比较, 可以看出外点运动矢量基本上都被剔除。最后利用变换模型参数迭代算法求解出图像间的变换模型参数, 图 5(d) 给出了通过变换模型参数计算出的图像的运动矢量场, 从图中可以看出图像运动矢量场中保留的运动矢量具有很强的一致性。在迭代估计变换模型参数过程中, 若变换模型参数中的最大变化量小于 0.05 时, 则停止迭代, 本次试验经过 3 次迭代即满足了上述条件, 得到的变换模型参数为

$$M = \begin{bmatrix} 0.9935 & 0.0015 & 3.9131 \\ -0.0020 & 0.9949 & 0.4636 \\ 0.0000 & 0.0000 & 1 \end{bmatrix}$$

最后, 利用提出的图像镶嵌算法针对定点拍摄一段视频图像序列进行镶嵌, 该视频图像序列共包含有 90 帧图像, 为提高计算效率, 以每隔 2 帧提取一帧图像参与图像的镶嵌。图 6 给出的最终的镶嵌结果, 其中是以视频图像序列的第 1 帧为参考图像。从图中可以看出尽管图像中存在大量的重复性内容, 采用本文提出的方法仍能获得较好的镶嵌结果。同时, 需要值得注意的是, 由于块匹配算法的局限性, 当图像序列中图像间存在较为明显几何变形时, 或者在拍摄过程中, 若存在镜头的缩放等运动, 采用该算法进行镶嵌时, 将得不到正确的镶嵌结果。

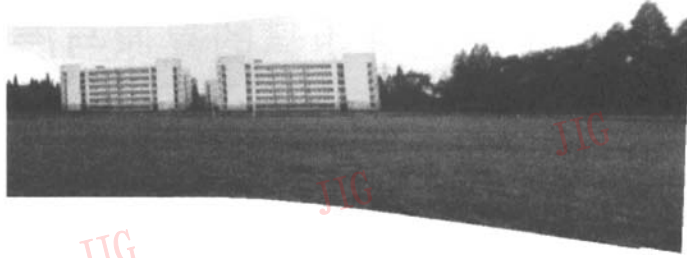


图 6 视频图像序列镶嵌结果

Fig. 6 Result of video mosaicing

6 结 论

基于块匹配的视频图像镶嵌算法,首先利用基于相位相关的块匹配算法计算出图像间的运动矢量场,再根据计算出来的运动矢量场确定图像子块之间的对应关系,然后由这些图像子块间的对应关系估计出图像间的变换模型参数进行图像镶嵌。在估计图像运动矢量场时,由于图像噪声、运动物体的遮挡等原因会出现一些与主运动不一致的运动矢量(即外点运动矢量),若直接利用估计出的运动矢量场确定图像间对应点的关系,则会严重影响图像间的变换参数的准确估计,由此本文算法首先根据图像运动矢量场的直方图剔除大部分外点运动矢量,由于运动矢量的直方图只是反映了运动矢量大小的统计特性,不能够反映运动矢量在图像运动矢量场中的几何分布,因此在经过上述处理后,仍然会存在一些与主运动一致但与其邻域内运动矢量不一致的运动矢量,对此,通过迭代求解图像间变换参数,以消除这类运动矢量对图像间变换参数估计的影响。最后,利用提出的算法对实际拍摄的一段视频图像序列进行镶嵌,获得了比较好的镶嵌结果,通过实验

验证了该算法的有效性。

参考文献 (References)

- 1 Irani M, Anandan P, Bergen J, *et al.* Efficient representations of video sequences and their applications[J]. *Signal Processing: Image Communication*, 1996, 8(4):327 ~ 351.
- 2 Szeliski R. Video mosaics for virtual environments [J]. *IEEE Computer Graphics and Applications*, 1996, 16(2): 22 ~ 30.
- 3 Mann S, Picard R W. Video orbits of the projective group: A simple approach to featureless estimation of parameters [J]. *IEEE transactions on image processing*, 1996, 6(9):1280 ~ 1295.
- 4 Traka M, Tziritas G. Panoramic view construction [J]. *Signal Processing: Image Communication*, 2003, 18(6):465 ~ 481.
- 5 Mann S, Picard R W. Virtual bellows; constructing high quality stills from video[A]. In: *Proceedings of 1st International Conference on Image Processing[C]*, Los Alamitos, CA, USA, 1994:363 ~ 367.
- 6 Kuglin C D, Hines D C. The phase correlation image alignment mMethod[A]. In: *Proceedings of IEEE International Conference on Cybernetics and Society [C]*, San Francisco, CA, USA, 1975: 163 ~ 165.
- 7 BAI Xue-sheng, JIN Guo-ying, XU Guang-you, *et al.* Video segmentation based on inter-frame motion [J]. *Journal of Tsinghua University (Sci & Tech)*, 2002, 42(1):76 ~ 79. [白雪生,金国英,徐光祐等. 基于帧间运动的视频分割[J]. *清华大学学报*, 2002, 42(1):76 ~ 79.]