

文章编号:1004-9037(2012)05-0528-07

# 基于二维投影变换的视频全景图拼接算法

邹晓春<sup>1</sup> 何明一<sup>1</sup> 赵歆波<sup>2</sup>

(1. 西北工业大学电子信息学院, 西安, 710072; 2. 西北工业大学计算机学院, 西安, 710072)

**摘要:**视频序列全景图拼接技术是视频图像分析的关键技术之一。针对帧间运动模型为二维投影变换的常见视频序列,提出了一种全景图拼接算法。该算法首先采用改进的基于RANSAC的特征点匹配算法对相邻帧的运动参数进行估计,接着采用直接法对各帧与全景图间的运动参数进行精确调整,最后运用自适应中值滤波绘制出全景图。实验结果表明,该方法能够生成较高质量的全景图,基本能够实现无缝连接。

**关键词:**视频处理;全景图像;RANSAC算法;运动估计

**中图分类号:**TP391

**文献标识码:**A

## Video Panorama Mosaic Algorithm Based on 2D Projection Transform

Zou Xiaochun<sup>1</sup>, He Mingyi<sup>1</sup>, Zhao Xinbo<sup>2</sup>

(1. School of Electronics and Information, Northwestern Polytechnical University, Xi'an, 710072, China;

2. School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, China)

**Abstract:** Panorama mosaic is one of key techniques in video analysis. Aiming at the two-dimensional projective motion between frames used in the video sequence frequently, a mosaic algorithm for video sequence panoramic image is proposed. Firstly, the motion parameters of adjacent frame are estimated by using improved RANSAC-based matching algorithm for feature points. Secondly, the motion parameters between every frame with panorama are adjusted accurately by a direct estimation method. Finally, the video sequence panoramic image is plotted by self-adaptive median filter. Experimental results show that the proposed algorithm can generate the panoramic image with high quality.

**Key words:** video processing; panoramic image; RANSAC algorithm; motion estimation

## 引 言

视频序列全景图绘制算法是计算机视觉、医学图像分析、遥感数据分析等领域中的一项重要技术。其目的在于对取自同一场景、不同视角、不同焦距、围绕同一光学中心拍摄、相互有部分重叠的视频序列图像进行无缝拼接,即先利用图像配准算法求得各帧图像间运动参数,然后合成一张大的静态宽视角图像。要求拼接算法能够使合成后的图像最大程度地与所拍摄的真实场景接近,没有明显的缝合线。

可见,视频序列图像配准是全景图绘制的关键

技术。图像配准技术的精度与鲁棒性对全景图的质量会产生重大的影响<sup>[1-3]</sup>。图像配准的方法通常分为两类:基于图像灰度的配准(又称为直接法)和基于图像特征的配准<sup>[4]</sup>。本文在研究两类配准方法各自特点的基础上,针对帧间运动模型为二维投影变换的常见视频序列,提出了一种使特征匹配算法和直接法相结合的新图像配准算法,该算法具有以下5个特点:

(1)采用二维投影变换作为视频帧间模型,该模型具有旋转、缩放和位移不变性,适用于以下3种情形所拍摄的视频序列:空间为平面场景和任意摄像机运动;任意空间场景,摄像机为旋转或放缩运动;远距离场景拍摄,即景深远小于场景到摄像

机的距离(如航拍或卫星成像等),其中最后一种情形近似满足适用条件。

(2)提出一种改进的基于RANSAC(Random sample consensus)的特征点匹配算法,该算法由一组从匹配特征点中得到的运动矢量来计算视频序列相邻帧间运动模型的参数,并采用内循环的方式提高了采样数据的使用效率。

(3)采用改进的基于RANSAC的特征点匹配算法对相邻帧的运动参数进行估计,利用特征点的空间信息有效地去除了光照和噪声的影响。

(4)采用直接法对各帧与全景图间的运动参数进行精确调整。这种基于图像像素操作的方法,利用了图像像素的全部信息,可以达到很高的亚像素级精度。

(5)利用基于特征的运动参数估计算法得到的各帧图像与背景图像间的运动参数,作为直接法的初始值,有效地提高了直接法的收敛速度,避免陷入局部极小值。

由此可见,该算法充分利用了特征匹配算法鲁棒性好和直接法精度高的特点。实验表明,本文算法可很好地用于视频序列的全景图绘制中。

## 1 算法描述

### 1.1 改进的基于RANSAC的特征点匹配算法

为了剔除视频序列图像中的错误以及与全局运动不符的像素点,从而得到摄像机运动模型的参数,本文提出了一种改进的基于RANSAC的特征点匹配算法。算法共包括特征点探查、特征点间对应关系的确定和运动参数估计3个部分。考虑帧间运动模型为二维投影变换的常见视频序列 $I = \{I_1(\mathbf{x}), \dots, I_{t-1}(\mathbf{x}), I_t(\mathbf{x}), I_{t+1}(\mathbf{x}), \dots, I_n(\mathbf{x})\}$ ,其中 $I_t(\mathbf{x})$ 为第 $t$ 帧图像,大小为 $k \times l$ , $\mathbf{x} = [x, y]^T$ 为第 $t$ 帧图像中像素点的坐标,算法描述如下:

(1)运用特征提取算法(Harris 算法<sup>[5]</sup>)在当前帧 $I_t(\mathbf{x})$ 和后相邻帧 $I_{t+1}(\mathbf{x})$ 中分别提取特征点集 $\{\mathbf{p}_i\}$ 和 $\{\hat{\mathbf{p}}_k\}$ ;

(2)对当前帧 $I_t(\mathbf{x})$ 中的特征点,以前相邻帧 $I_{t-1}(\mathbf{x})$ 和当前帧 $I_t(\mathbf{x})$ 间的运动模型参数为预测值(设定 $I_0(\mathbf{x}) = I_1(\mathbf{x})$ ),来确定相应的搜索区域,并运用某种相似性原则(如最小绝对距离<sup>[6]</sup>)计算匹

配特征点对集合 $C = \{\mathbf{p}_i = (x_i, y_i, 1)^T \leftrightarrow \hat{\mathbf{p}}_k = (\hat{x}_k, \hat{y}_k, 1)^T\}$ ;

(3)运用RANSAC鲁棒估计方法<sup>[7]</sup>得到集合 $C$ 中最大的内点集合和二维投影运动模型参数 $\mathbf{h}_{t,t+1}$ 的估计,即

①从 $C$ 中随机抽取子集 $\zeta$ , $|\zeta| = 4$ (因为估计参数 $\mathbf{h}_{t,t+1}$ 至少需要4对数据点),用于估计 $\mathbf{h}_{t,t+1}$ 参数;

②利用子集 $\zeta$ 中的对应关系,计算二维投影运动模型参数 $\mathbf{h}_{t,t+1}$ ,即

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1\hat{x}_1 & -y_1\hat{x}_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1\hat{y}_1 & -y_1\hat{y}_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2\hat{x}_2 & -y_2\hat{x}_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2\hat{y}_2 & -y_2\hat{y}_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3\hat{x}_3 & -y_3\hat{x}_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3\hat{y}_3 & -y_3\hat{y}_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4\hat{x}_4 & -y_4\hat{x}_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4\hat{y}_4 & -y_4\hat{y}_4 \end{bmatrix} \cdot \begin{bmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \end{bmatrix} = \begin{bmatrix} \hat{x}_1 \\ \hat{y}_1 \\ \hat{x}_2 \\ \hat{y}_2 \\ \hat{x}_3 \\ \hat{y}_3 \\ \hat{x}_4 \\ \hat{y}_4 \end{bmatrix} \quad (1)$$

式中 $\mathbf{h}_{t,t+1} = (h_{00} \ h_{01} \ h_{02} \ h_{10} \ h_{11} \ h_{12} \ h_{20} \ h_{21})_{t,t+1}^T$ 。

③对 $C$ 中的每一个匹配点对 $\{\mathbf{p}_i \leftrightarrow \hat{\mathbf{p}}_k\}$ ,计算拟合误差 $d(\hat{\mathbf{p}}_k, \mathbf{H}_{t,t+1} \mathbf{p}_i)$ ( $d$ 为欧几里德距离函数<sup>[8]</sup>),其中

$$\mathbf{H}_{t,t+1} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & 1 \end{bmatrix}_{t,t+1}$$

④设置门限值 $\epsilon$ ,当拟合误差小于此门限值时,表示此匹配点服从当前运动模型,记为内点,确定内点集合 $\chi = \{\mathbf{p}_i \leftrightarrow \hat{\mathbf{p}}_k \in C \mid |d(\hat{\mathbf{p}}_k, \mathbf{H}_{t,t+1} \mathbf{p}_i)| < \epsilon\}$ ,并统计内点的数目;

⑤利用最小平方估计算法<sup>[9]</sup>估计该内点集合的运动参数 $\mathbf{h}_{t,t+1}$ ,即

$$\left( \sum_{\mathbf{p}_i \in \chi} A_i \right) \mathbf{h}_{t,t+1} = \sum_{\mathbf{p}_i \in \chi} \mathbf{b}_i \quad (2)$$

式中

$$A_i = \begin{bmatrix} x_i^2 & x_i y_i & x_i & 0 & 0 & 0 & -x_i^2 \hat{x}_i & -x_i y_i \hat{x}_i \\ x_i y_i & y_i^2 & y_i & 0 & 0 & 0 & -x_i y_i \hat{x}_i & -y_i^2 \hat{x}_i \\ x_i & y_i & 1 & 0 & 0 & 0 & -x_i \hat{x}_i & -y_i \hat{x}_i \\ 0 & 0 & 0 & x_i^2 & x_i y_i & x_i & -x_i^2 \hat{y}_i & -x_i y_i \hat{y}_i \\ 0 & 0 & 0 & x_i y_i & y_i^2 & y_i & -x_i y_i \hat{y}_i & -y_i^2 \hat{y}_i \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i \hat{y}_i & -y_i \hat{y}_i \\ x_i^2 \hat{x}_i & x_i y_i \hat{x}_i & x_i \hat{x}_i & x_i^2 \hat{y}_i & x_i y_i \hat{y}_i & x_i \hat{y}_i & -x_i^2 (\hat{x}^2 + \hat{y}^2) & -x_i y_i (\hat{x}^2 + \hat{y}^2) \\ x_i y_i \hat{x}_i & y_i^2 \hat{x}_i & y_i \hat{x}_i & x_i y_i \hat{y}_i & y_i^2 \hat{y}_i & y_i \hat{y}_i & x_i y_i (\hat{x}^2 + \hat{y}^2) & -y_i^2 (\hat{x}^2 + \hat{y}^2) \end{bmatrix}$$
$$b_i = (x_i \hat{x}_i \quad y_i \hat{x}_i \quad \hat{x}_i \quad x_i \hat{y}_i \quad y_i \hat{y}_i \quad \hat{y}_i \quad x_i (\hat{x}^2 + \hat{y}^2) \quad y_i (\hat{x}^2 + \hat{y}^2))^T$$

⑥ 重复步骤③~⑤ 三次,得到当前子集 $\zeta$ 所对应的最大内点集合 $\chi'$ ;

⑦ 重复步骤①~⑥,直至达到最大抽样次数  $N$  并将 $|\chi'|$ 最大的集合记为内点集合 $\chi_m$ ,由该集合得到的参数  $\mathbf{h}_{t,t+1}$  即为当前帧  $I_t(\mathbf{x})$  和后相邻帧  $I_{t+1}(\mathbf{x})$  间的二维投影运动模型参数;

(4) 令  $t=t+1$ ,重复步骤(1~3),直至  $t=n-1$ 。

上述算法中的步骤(1~3)分别完成特征点探查、特征点间对应关系的确定和运动参数估计的工作。传统算法的主要计算量集中在步骤(2)中的特征点对匹配搜索上<sup>[1]</sup>。为了减少传统算法的计算量,上述算法对搜索区域的确定进行了改进。由于循环算法中前相邻帧  $I_{t-1}(\mathbf{x})$  和当前帧  $I_t(\mathbf{x})$  间的运动模型参数  $\mathbf{h}_{t-1,t}$  已经求得,因此可以用其来指导确定当前帧  $I_t(\mathbf{x})$  和后相邻帧  $I_{t+1}(\mathbf{x})$  间匹配特征点对的搜索区域,假设  $\mathbf{p}_i$  为当前帧  $I_t(\mathbf{x})$  中的某个特征点,则后相邻帧中匹配特征点的搜索区域可确定为以  $\hat{\mathbf{p}}_i = \mathbf{H}_{t-1,t} \mathbf{p}_i$  为中心的矩形邻域,大小一般选为  $5 \times 5$  或  $3 \times 3$ ,这样就可大大减少搜索时间,以提高计算速度。

利用 RANSAC 算法找出最大的内点集合,是决定上述算法精度的关键。通常,对数据样本进行随机采样的次数越多,算法的精度就越高,但计算量也随之增大。为了在保证算法精度的同时,减少计算量,上述改进算法步骤(3)中的步骤⑥采用内循环的方式计算某一子集 $\zeta$ 所对应的最大内点集合 $\chi'$ ,因而提高了采样数据的使用效率,减少了采样的次数。

1.2 基于直接法的全景图合成算法

利用上节描述的基于 RANSAC 的特征点匹配算法,得到了当前帧  $I_t(\mathbf{x})$  和后相邻帧  $I_{t+1}(\mathbf{x})$  间的运动模型参数矩阵  $\mathbf{H}_{t,t+1}$ 。利用这些帧间的变换关

系,构建各帧图像与背景图像  $I_B(\mathbf{x})$  间的精确运动参数矩阵  $\mathbf{W}_t$ ,是全景图像合成算法的中心任务。

首先,假设原始透明背景图像  $I_B(\mathbf{x})$  各点像素值均为 0,图像大小  $s \times t$  由  $\mathbf{I}$  决定,如图 1 所示。然后,选定当前视频序列图像第一帧  $I_1(\mathbf{x})$  作为参考帧。再定义将  $I_1(\mathbf{x})$  逐点复制于全景图  $I_B(\mathbf{x})$  的变换为单位矩阵  $\mathbf{W}_1 = \overline{\mathbf{W}}_1 = \mathbf{I}$ ,则可利用参数矩阵  $\mathbf{H}_{t,t+1}$  的可逆性和级联性来得到任意帧与全景图间的初始变换关系  $\overline{\mathbf{W}}_t$

$$\overline{\mathbf{W}}_t = \overline{\mathbf{W}}_{t+1} \cdot \mathbf{H}_{t,t+1} \tag{3}$$

以第 5 帧为例

$$\begin{aligned} \overline{\mathbf{W}}_5 &= \overline{\mathbf{W}}_4 \cdot \mathbf{H}_{4,5}^{-1} = \overline{\mathbf{W}}_3 \cdot \mathbf{H}_{3,4}^{-1} \cdot \mathbf{H}_{4,5}^{-1} = \\ &\overline{\mathbf{W}}_2 \cdot \mathbf{H}_{2,3}^{-1} \cdot \mathbf{H}_{3,4}^{-1} \cdot \mathbf{H}_{4,5}^{-1} = \\ &\overline{\mathbf{W}}_1 \cdot \mathbf{H}_{1,2}^{-1} \cdot \mathbf{H}_{2,3}^{-1} \cdot \mathbf{H}_{3,4}^{-1} \cdot \mathbf{H}_{4,5}^{-1} = \\ &\mathbf{H}_{1,2}^{-1} \cdot \mathbf{H}_{2,3}^{-1} \cdot \mathbf{H}_{3,4}^{-1} \cdot \mathbf{H}_{4,5}^{-1} \end{aligned} \tag{4}$$

由式(3)可以看出,当多帧图像拼接时,级联算法通常会导致累计误差的出现,这些误差往往会带来全景图像的模糊,甚至裂痕。因此,有必要通过将每一幅输入图像高精度排列到所构建的背景图像中,来进一步优化  $\overline{\mathbf{W}}_t$ ,得到精确运动参数  $\mathbf{W}_t$ 。

对于相邻帧能够很好地满足光流约束的常见视频序列,本文采用基于直接法的全景图合成算法,对各帧图像与全景图间的运动参数进行精确调整,其算法描述如下:

(1) 利用变换  $\overline{\mathbf{W}}_1 = \mathbf{I}$ ,将参考帧  $I_1(\mathbf{x})$  逐点复制于全景图  $I_B(\mathbf{x})$  中,令  $t=2$ ;

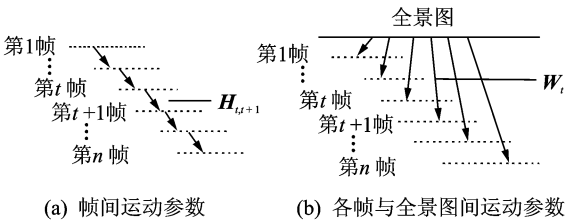


图 1 全景图合成的投影关系

(2) 以式(3)得到的初始变换 $\overline{W}_t$ 为预测值,利用式(5)求解当前帧 $I_t(\mathbf{x})$ 与背景图像 $I_B(\mathbf{x})$ 间的精确运动参数矩阵 $W_t$

$$W_t = \underset{\overline{W}_t}{\operatorname{argmin}} \sum_{\mathbf{x} \in \operatorname{area} I_t(\mathbf{x})} \phi(I_B(W_t \mathbf{x}), I_t(\mathbf{x})) \quad (5)$$

$$\text{式中 } W_t = \begin{bmatrix} w_0 & w_1 & w_2 \\ w_3 & w_4 & w_5 \\ w_6 & w_7 & w_8 \end{bmatrix}, I_B(W_t \mathbf{x}) \text{ 利用双线性插值}$$

算法<sup>[10]</sup>进行求值, $\phi$ 函数定义为

$$\phi(I_B, I_t) =$$

$$\begin{cases} 0 & I_B = 0 \\ (I_B(W_t \mathbf{x}) - I_t(\mathbf{x}))^2 & |I_B - I_t| < \lambda \\ \lambda^2 & \text{其他} \end{cases} \quad (6)$$

式中 $\lambda$ 表示 $I_B(W_t \mathbf{x})$ 与 $I_t(\mathbf{x})$ 间平方误差的上限。

(3) 利用求得的 $W_t$ ,将当前帧 $I_t(\mathbf{x})$ 添加到背景图像 $I_B(\mathbf{x})$ 上,操作时只改动 $I_B(\mathbf{x})$ 上像素值为0的透明点。即

$$I_B(W_t \mathbf{x}) = \begin{cases} I_t(\mathbf{x}) & I_B(W_t \mathbf{x}) = 0 \\ I_B(W_t \mathbf{x}) & I_B(W_t \mathbf{x}) \neq 0 \end{cases} \quad (7)$$

(4) 令 $t=t+1$ ,重复步骤(2)和步骤(3),直至 $t=n$ 。

此外, $I_B(W_t \mathbf{x})$ 的求解是决定上述算法精度的关键。在实际应用中, $W_t \mathbf{x}$ 的结果往往不在整数位置,直接取整势必降低其精度,因此本文利用双线性插值算法求取 $I_B(W_t \mathbf{x})$ ,以达到亚像素精度。

步骤(2)中 $W_t$ 的求解是本算法的主要工作<sup>[11]</sup>。为了降低迭代的次数,本文采用Levenberg-Marquardt算法<sup>[12]</sup>来最小化代价函数

$$E_t = \sum_{\mathbf{x} \in \operatorname{area} I_t(\mathbf{x})} \phi(I_B(\mathbf{x}'), I_t(\mathbf{x})) \quad (8)$$

式中 $\mathbf{x}' = [x', y']^T = W_t \mathbf{x} = \begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{w} \end{bmatrix}^T$ 。考虑满足二维投影运动模型的常见视频序列,得到

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{w} \end{bmatrix} = \begin{bmatrix} w_0 & w_1 & w_2 \\ w_3 & w_4 & w_5 \\ w_6 & w_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (9)$$

以 $\phi(I_B, I_t) = (I_B - I_t)^2$ 为例

$$E_t = \sum_{\mathbf{x} \in \operatorname{area} I_t(\mathbf{x})} (I_B(\mathbf{x}') - I_t(\mathbf{x}))^2 = \sum_{\mathbf{x} \in \operatorname{area} I_t(\mathbf{x})} (I_B(\frac{w_0 x + w_1 y + w_2}{w_6 x + w_7 y + 1}, \frac{w_3 x + w_4 y + w_5}{w_6 x + w_7 y + 1}) - I_t(x, y))^2 \quad (10)$$

分别计算 $E_t$ 的梯度矢量 $\nabla E_t = [\frac{\partial E_t}{\partial w_0}, \frac{\partial E_t}{\partial w_1},$

$\frac{\partial E_t}{\partial w_2}, \frac{\partial E_t}{\partial w_3}, \frac{\partial E_t}{\partial w_4}, \frac{\partial E_t}{\partial w_5}, \frac{\partial E_t}{\partial w_6}, \frac{\partial E_t}{\partial w_7}]$ 和海森矩阵 $\nabla^2 E_t =$

$$\begin{bmatrix} \frac{\partial^2 E_t}{\partial w_0^2} & \frac{\partial^2 E_t}{\partial w_0 \partial w_1} & \cdots & \frac{\partial^2 E_t}{\partial w_0 \partial w_7} \\ \frac{\partial^2 E_t}{\partial w_1 \partial w_0} & \frac{\partial^2 E_t}{\partial w_1^2} & \cdots & \frac{\partial^2 E_t}{\partial w_1 \partial w_7} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E_t}{\partial w_7 \partial w_0} & \frac{\partial^2 E_t}{\partial w_7 \partial w_1} & \cdots & \frac{\partial^2 E_t}{\partial w_7^2} \end{bmatrix} \text{代入得到}$$

$$\Delta W'_t = [\nabla^2 E_t]^{-1} \sum_{\mathbf{x} \in \operatorname{area} I_t(\mathbf{x})} [\nabla E_t]^T (I_B(\mathbf{x}') - I_t(\mathbf{x})) \quad (11)$$

不断修正 $W_t \leftarrow W_t + \Delta W_t$ ,直至 $W_t$ 有稳定解,其中

$$\Delta W'_t = [\Delta w_0, \Delta w_1, \Delta w_2, \Delta w_3, \Delta w_4, \Delta w_5, \Delta w_6, \Delta w_7]^T$$

$$\Delta W_t = \begin{bmatrix} \Delta w_0 & \Delta w_1 & \Delta w_2 \\ \Delta w_3 & \Delta w_4 & \Delta w_5 \\ \Delta w_6 & \Delta w_7 & \Delta w_8 \end{bmatrix}$$

### 1.3 基于自适应中值滤波的全景图融合算法

利用上两节描述的图像配准策略,得到了各帧图像与背景图像 $I_B(\mathbf{x})$ 间的精确运动参数矩阵 $W_t$ ,并绘制出具有亚像素精度的视频全景图。但是,由于光照等拍摄条件对视频序列各帧色度和亮度的影响,会在全景图像的不同区域产生人眼可分辨的差异。因此本文对全景图像的每个像素点运用基于自适应中值滤波的融合算法<sup>[13]</sup>来确定其取值,从而绘制出最终全景图。在 $I_B(\mathbf{x})$ 上确定以 $\mathbf{x}$ 为中心的矩形邻域 $S_{xy}$ ,最大尺寸选为 $S_{\max} \times S_{\max}$ ,自适应中值滤波算法工作在两个层次上,定义为A层和B层,如下所示:

A层: $A_1 = z_{\text{med}} - z_{\min}$

$A_2 = z_{\text{med}} - z_{\max}$

如果 $A_1 > 0$ 且 $A_2 < 0$ ,转到B层

否则增大窗口尺寸

如果窗口尺寸 $\leq S_{\max}$ ,重复A层

否则输出 $I_B(\mathbf{x})$

B层: $B_1 = I_B(\mathbf{x}) - z_{\min}$

$B_2 = I_B(\mathbf{x}) - z_{\max}$

如果 $B_1 > 0$ 且 $B_2 < 0$ ,输出 $I_B(\mathbf{x})$

否则输出 $z_{\text{med}}$

其中: $z_{\min} = S_{xy}$ 中灰度级的最小值, $z_{\max} = S_{xy}$ 中

灰度级的最大值,  $z_{\text{med}}=S_{xy}$  中灰度级的中值。

由该算法的运行机理可以看出,  $A$  层的目的是决定中值滤波器的输出  $z_{\text{med}}$  是否是一个类冲激式的噪声成分,  $B$  层的目的是决定当前矩形窗口  $S_{xy}$  中心点  $I_B(\mathbf{x})$  本身是否是一个类冲激式的噪声成分。

若  $z_{\text{med}}$  是一个类冲激式的噪声成分, 则增大窗口尺寸并重新计算, 直到  $z_{\text{med}}$  满足非噪声条件。若  $I_B(\mathbf{x})$  是一个类冲激式的噪声成分, 算法作为中值滤波器输出  $z_{\text{med}}$ ; 若  $I_B(\mathbf{x})$  不是冲激噪声, 为了减少图像中的失真, 算法输出  $I_B(\mathbf{x})$ 。

执行该算法的优点在于: 它能够有效地去除冲激噪声, 平滑其他非冲激噪声, 并减少诸如物体边界细化或粗化等失真。

2 实验结果

为验证本文算法的正确性及有效性, 对其进行了计算机仿真实验。实验的硬件环境为, Intel 酷睿 2 双核 E7300 2.66 GHz CPU 以及 2 GB RAM 的 PC 机, 软件平台为 Matlab。

**实验 1** 验证改进的基于 RANSAC 的特征点匹配算法的有效性 with 正确性。实验利用 Harris 算法进行特征提取, 并构造出已知特性的匹配点对集合  $C$ , 其中服从已知理想帧间运动模型  $H_{t,t+1}$  的内点集合所占比例为  $p_i$ , 剩余的外点集合占百分比  $p_0$ 。实验采用 Stefan 和 Coastguard 序列前 80 帧作为测试序列, 利用 1.1 节算法对已知数据集  $C$  进行特征点匹配, 得到帧间运动模型  $H'_{t,t+1}$ , 计算出集合  $C$  中服从  $H'_{t,t+1}$  的内点集合所占比例  $p'_i$  与外点集合比

例  $p'_0$ , 并由  $1-(1-p'_0)^{|s|}$  得到  $p'_f$ 。表 1 给出了改进的基于 RANSAC 的特征点匹配算法在不同抽样次数下  $p'_f$  的平均值。为了方便比较, 同时给出了理想情况下基于 RANSAC 的特征点匹配算法的平均  $p_f$  值, 其中算法的成功率设为  $P=99.9\%$ 。

由表 1 可看出, 在相同成功率下, 改进算法能够在较少的抽样次数下, 得到较小的外点集合。由于改进算法采用误差门限来确定内点, 对外点集合中靠近内点的外点可能产生误判, 因此得到小于  $p_f$  的  $p'_f$  值, 但个别外点的误判并不会影响内点集合收敛于正确的运动模型。实验结果说明, 改进算法能更好地估计出全局运动模型参数, 有效性与正确性得到验证。

**实验 2** 采用大小为  $352\times 289$  的标准 Flower 图像序列, 共 40 帧, 和大小为  $360\times 288$  的标准 Library 图像序列, 共 29 帧。像素点的灰度均为 8 bit, 即 256 个灰度级。图 2(a) 与图 3(a) 为两个原始视频序列的样本, 各列举 2 帧。图 2(b), 3(b) 分别为利用本文算法拼接的两个序列的全景图像。由实验结果不难看出, 本文算法的拼接结果消除了拼接部分的亮度差异和“鬼影”, 基本达到了无缝拼接。

**实验 3** 采用大小为  $640\times 480$  的自拍 Garden 彩色图像序列, 共 62 帧。图 4(a) 列举了该序列的第 1 帧, 30 帧和 60 帧图像。图 4(b) 是采用文献[14]的方法对此序列拼接的结果, 而图 4(c) 则是采用本文方法拼接而成的全景图。从比较结果不难看出, 由于文献[14]的方法未对帧图像进行二维投影变换, 因此拼接而成的全景图存在严重的场景变形, 而采

表 1 两种特征点匹配算法得到的成功率比较

设 $P=99.9\%$		Stefan $p_0=11\%$			Coastguard $p_0=18\%$		
基于 RANSAC 的特征点匹配算法的理想结果	最大抽样次数	7			12		
	$p_f$ 平均值/ $\%$	37.3			54.1		
改进的基于 RANSAC 的特征点匹配算法的实验结果	最大抽样次数	5	6	7	10	11	12
	$p'_f$ 平均值/ $\%$	42.1	36.7	30.8	53.6	46.5	38.9



(a) Flower 视频序列的第 1 帧和 40 帧 (b) 采用本文算法生成的静态全景图

图 2 Flower 序列的结果全景图

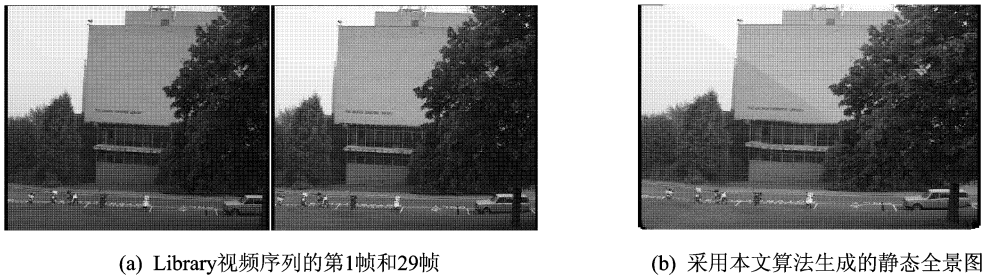


图3 Library序列的结果全景图

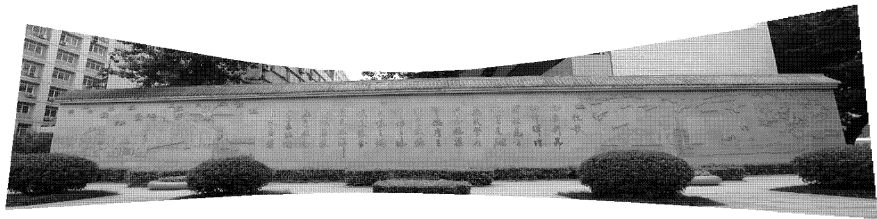
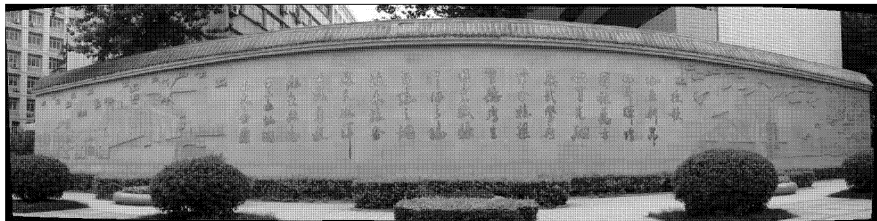
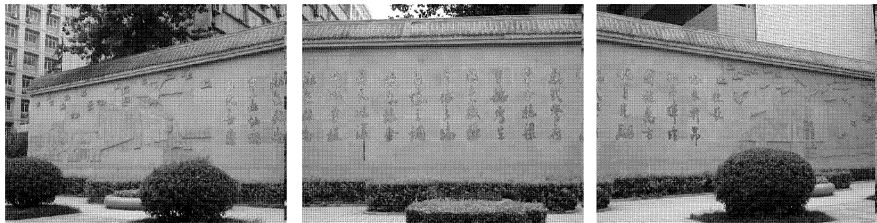


图4 不同算法的实验结果比较

用本文基于二维投影变换的方法得到的全景图,没有大的变形,视觉效果良好。

3 结束语

本文提出并实现了一种基于二维投影变换的视频序列静态全景图拼接算法。该算法针对帧间运动模型为二维投影变换的常见视频序列,提出了一种结合特征匹配算法和直接法的图像配准策略,即首先采用改进的基于RANSAC的特征点匹配算法对相邻帧的运动参数进行估计;接着采用直接法对各帧与全景图间的运动参数进行精确调整。该策略以特征匹配算法的结果作为直接法的初始值,充分

利用了特征匹配算法鲁棒性好和直接法精度高的优点,有效地去除了级联算法所带来的累积误差,并减少了匹配搜索的时间。为了避免光照等拍摄条件对视频序列各帧色度和亮度的影响,算法运用自适应中值滤波绘制出全景图。实验结果表明,该算法能被很好地运用于常见视频序列的全景图绘制中。

参考文献:

[1] Capel D P. Image mosaicing and super-resolution [D]. UK: Department of Computing Science, University of Oxford, 2001:47-78.  
[2] Chan S C, Shum Heung-Yeung, Ng King-To. Im-

age-based rendering and synthesis[J]. Signal Processing Magazine, 2007, 24(6):22-33.

[3] 王嵩,李宏,朱霞,等. 基于全景图像和状态匹配的視頻修复算法[J]. 数据采集与处理,2009,24(5):694-698.

Wang Song, Li Hong, Zhu Xia, et al. Restoration algorithm for digital video based on panoramic image and state matching[J]. Journal of Data Acquisition and Processing, 2009,24(5):694-698.

[4] Stojmenovic I, Nayak A. Handbook of applied algorithms: solving scientific, engineering, and practical problems [M]. Canada: University of Ottawa, 2007: 317-346.

[5] Harris C, Stephens M. A combined corner and edge detector [C]//The Fourth Alvey Vision Conference. Manchester:[s. n. ],1988: 147-151.

[6] Farin D, de With P H N. Evaluation of a feature-based global motion estimation system[C]// SPIE Visual Communications and Image Processing. Beijing, China:[s. n. ],2005: 1331-1342.

[7] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6):381-395.

[8] Shih F. Image processing and pattern recognition: Fundamentals and Techniques[M]. USA:[s. n. ], 2010: 179-218.

[9] Candy J. Model-based signal processing[M]. USA: [s. n. ], 2005: 135-173.

[10] Qi Li, Sato I, Murakami Y. Interpolation effects on accuracy of mutual information based image registration [C]// The IEEE International Conference on Geoscience and Remote Sensing Symposium. Denver, Colorado: IEEE,2006:180-183.

[11] Dirk Sven Farin. Automatic video segmentation employing object/camera modeling techniques[D]. The Netherlands: University of Eindhoven, 2005: 52-171.

[12] Szeliski R. Image mosaicing for tele-reality applications[C]// The IEEE Workshop on Applications of Computer Vision. Sarasota, Florida: IEEE Computer Society Press, 1994: 44-53.

[13] Gonzalez R, Woods E. Digital image processing [M].USA:Prentice Hall, 2002:175-223.

[14] Brown M, Lowe D G. Recognising panoramas [C]// 9th International Conference on Computer Vision. Nice, France: IEEE Computer Society Press, 2003, 2: 1218-1225.

**作者简介:**邹晓春(1980-),女,讲师,研究方向:计算机视觉,E-mail:xczou@nwpu.edu.com;何明一(1958-),男,教授,研究方向:信号、图像与信息处理的理论与技术;赵歆波(1970-),男,教授,研究方向:眼动技术及其应用。