

基于视频片断的全景图拼接

季 诚, 欧阳宁, 张 彤

(桂林电子科技大学 图像信息研究所, 广西 桂林 541004)

摘 要: 视频拼接技术作为计算机视觉和计算机图形学两个学科上的交叉研究领域, 在近年来得到了极大的关注, 成为这两个互补领域的研究焦点之一。视频拼接技术广泛应用于虚拟现实、电影动画或特技、图像分辨率的提高, 及视频压缩、视频检索等方面, 有着广泛的研究价值和应用领域。利用八参数透射模型解决视频中相同场景图像之间的坐标变换问题, 通过光流法求得模型的八个参数并利用求得的参数模型有效地生成全景图。实验结果显示, 算法在保持较好拼接效果的同时能降低时间复杂度, 适合运用在对实时性要求较高的场合。

关键词: 图像拼接; 视频全景图; 透射模型; 光流

中图分类号: TP751

文献标识码: A

文章编号: 1673-629X(2009)09-0239-03

Panorama Stitching in Video Sequences

JI Cheng, OUYANG Ning, ZHANG Tong

(Image Information Institute, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: Video mosaic has been active area of research in recent years. It becomes a major focus of attention combining two complementary fields: computer vision and computer graphics. Video mosaic has wide applications, such as virtual reality, image resolution enhancement, video compression, video indexing and so on. Using transmitting model, resolve the problem of coordinate transform between two images of the same scene in a video. By optical flow, figure out the 8 parameters of the model and use them to generate the panorama effectively. The experimental results show that the algorithm can reduce the rate of time complexity and maintain a relatively good stitching impression at the same time, which is fit for real-time processing.

Key words: image stitching; video panorama; transmitting model; optical flow

0 引 言

在研究中人们发现, 视频数据往往包含大量的冗余信息, 特别是在某些静止场景的场合, 视频序列往往是由于摄像机作一定的运动生成的, 如果可以用一定的数学模型来表示摄像机的对应运动, 然后用一定的方法求出这种模型的参数, 就可以用一幅图像来表示这个场景, 然后利用几幅信息量大的图像拼接成全景图像, 从而有效地消除视频中的冗余信息。

1 变换模型的选择

这里所谓变换模型问题就是指找到两帧相同场景图像之间的坐标变换, 并且把它用相应的变换模型(函

数)表示出来。给定视频帧之间的变换模型, 寻找视频帧之间的坐标变换的问题就转化成确定该模型的参数问题。所以说, 精确地选定一个变换模型对于进行视频全景图拼接起到了一个关键的作用。在参考文献[1~6]中, 分析了各种变换模型, 发现八参数透射模型是最为精确反映特定条件下视频帧之间坐标变换的模型, 在如下两种情况下可以精确地反映由于摄像机镜头作运动而产生的坐标变换:

- a: 静止 3D 场景, 镜头在固定位置任意运动;
- b: 静止平坦场景, 镜头任意运动。

在试验中, 使镜头做平移运动, 并利用八参数透射模型作为变换模型。下面给出了透射模型的变换函数及参数矩阵:

$$X' = \frac{b_{11}X + b_{12}Y + b_{13}}{b_{31}X + b_{32}Y + 1} \quad Y' = \frac{b_{21}X + b_{22}Y + b_{23}}{b_{31}X + b_{32}Y + 1}$$
$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & 1 \end{bmatrix}$$

综上所述, 如何得到一种快速的透射模型八个参

收稿日期: 2009-01-05; 修回日期: 2009-03-20

基金项目: 广西教育科研计划项目(桂教科研[2006]26号文107);
广西科技厅资助项目(桂科基 0731020)

作者简介: 季 诚(1983-), 男, 江苏响水人, 硕士研究生, 研究方向为图像处理; 欧阳宁, 硕士, 副教授, 硕士研究生导师, 研究方向为图像处理和识别等。

数的求解法成为问题的关键。Shum^[7]用特征点对应的方法求解透射模型。在下一节里采用一种无特征点对应的,用光流的办法来求解透射模型。

2 视频的全景图拼接算法

2.1 算法流程

确定了八参数的透射模型,问题的关键就变成如何快速而且准确地来求出模型的八个参数,以及求得参数后如何利用模型有效地生成最后的全景图。

需要指出的是这里视频拼接算法的输入是已经经过处理的视频,具体地说,就是已经利用一定的镜头分割算法对视频进行时间分割,并且用一定的手段对于分割后的镜头进行分类,只对于那些适合作视频拼接的镜头才作为视频拼接算法的输入。所谓适合进行视频拼接的镜头,就是指那些静止场景,或者只有极小量运动物体,由于摄像机作运动而产生的视频序列,这些序列在一些风光片里面大量存在,在其他的一些领域,诸如新闻、娱乐等领域也为数不少。

下面给出算法的流程:第一步,输入视频帧序列 $\{I_i | i = 0 \cdots N-1\}$,运用光流场估计从视频帧 I_i 到视频帧 I_{i-1} 对应的透射变换参数矩阵序列 $\{B_i | i = 0 \cdots N-1\}$;第二步,利用上一步求得的参数矩阵序列 B_i ,求出从第 i 帧 I_i 到初始帧 I_0 的投射变换参数矩阵序列 $\{P_i | i = 1 \cdots N-1\}$;最后,对所有帧图像 $\{I_i | i = 0 \cdots N-1\}$ 按照透射变换 P_i 变换到初始帧 I_0 ,在初始帧 I_0 的坐标系里面,对每一点取均值,输出结果图。

2.2 用光流估计透射模型的参数矩阵 B

光流场的概念最早是由 Horn 和 Schunck^[8]提出来的,它适合解决光照一定的情况下运动较小的两帧图像之间的运动估计问题。

设图像上的点 (X, Y) 在 t 时刻的光照强度为 $I(X, Y, t)$,经过时间间隔 Δt 以后,对应点的光照强度为 $I(X, Y, t + \Delta t)$,当 $\Delta t \rightarrow 0$ 时,可以认为光照强度不变,于是就有:

$$I(X, Y, t) = I(X + \Delta X, Y + \Delta Y, t + \Delta t)$$

将上式右边 Taylor 展开,有

$$I(X + \Delta X, Y + \Delta Y, t + \Delta t) = I(X, Y, t) +$$

$$\frac{\partial f}{\partial X} \Delta X + \frac{\partial f}{\partial Y} \Delta Y + \frac{\partial f}{\partial t} \Delta t + \epsilon$$

忽略二阶无穷小量,有

$$\frac{\partial I}{\partial X} \Delta X + \frac{\partial I}{\partial Y} \Delta Y + \frac{\partial I}{\partial t} \Delta t = 0$$

记做光流方程

$$uI_x + vI_y + I_t = 0$$

其中, $u = \Delta X, v = \Delta Y, I_x = \frac{\partial I}{\partial X}, I_y = \frac{\partial I}{\partial Y}, I_t = \frac{\partial I}{\partial t}$,

(u, v) 即为该点的光流。可以利用有限差分的方法来估算基本光流方程 $uI_x + vI_y + I_t = 0$ 中的偏导数。这里限于篇幅,不再具体介绍。

根据透射变换模型,有

$$u = X' - X = \frac{b_{11}X + b_{12}Y + b_{13}}{b_{31}X + b_{32}Y + 1} - X$$

$$v = Y' - Y = \frac{b_{21}X + b_{22}Y + b_{23}}{b_{31}X + b_{32}Y + 1} - Y$$

将上式代入光流方程,可得

$$\left(\frac{b_{11}X + b_{12}Y + b_{13}}{b_{31}X + b_{32}Y + 1} - X \right) * I_x + \left(\frac{b_{21}X + b_{22}Y + b_{23}}{b_{31}X + b_{32}Y + 1} - Y \right) * I_y + I_t = 0$$

可以对上式求最小均方误差来估计参数矩阵 B

$$\epsilon = \sum \left[\left(\frac{b_{11}X + b_{12}Y + b_{13}}{b_{31}X + b_{32}Y + 1} - X \right) * I_x + \left(\frac{b_{21}X + b_{22}Y + b_{23}}{b_{31}X + b_{32}Y + 1} - Y \right) * I_y + I_t \right]^2$$

由于目标函数是非线性的,利用上式很难求得参数矩阵 B 。在 2.3 节的实验里面,采用了下面的目标函数取得了较好的实验结果。

$$\epsilon = \sum \left[(b_{11}X + b_{12}Y + b_{13} - (b_{31}X + b_{32}Y + 1) * X) * I_x + (b_{21}X + b_{22}Y + b_{23} - (b_{31}X + b_{32}Y + 1) * Y) * I_y + I_t \right]$$

将上式展开,用最小二乘法获得如下的方程组:

$$\left(\sum \phi \phi^T \right) \cdot \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{21} \\ b_{22} \\ b_{23} \\ b_{31} \\ b_{32} \end{bmatrix} = \sum [(xI_x + yI_y - I_t) \phi]$$

解上面的八元一次方程组就可以得到所求的参数矩阵 B 。

2.3 实验与结果

当所有帧都在参考坐标系里面对齐以后,理论上可以选用以下的方法来合成最后的视频拼接结果图。

- 1) 当前位置所有点灰度值的均值(average);
- 2) 当前位置所有点灰度值的中值(median);
- 3) 均值或者中值的加权,权的大小与该点离图像中心位置的距离成反比;
- 4) 当前位置所有点灰度值的最大值;
- 5) 当前位置所有点灰度值的最小值。

方法 3 的依据是,透射变换中远离图像中心的点更容易受到噪声的干扰。方法 4 和方法 5 由于其对于



图1 输入第一帧



图2 输入第二帧

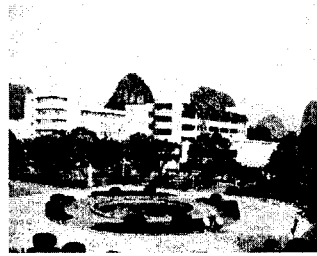


图3 输入第三帧



图4 输入第四帧



图5 输入第五帧



图6 全景拼接图

噪声的敏感性,在实际应用中很少采用。在文中的试验里采用了方法1。

图1~5是用于视频全景图拼接的输入帧图像,图6是最后的拼接图。从图中可以看出,拼接的效果还是基本令人满意的。

3 结束语

由于视频信息中存在大量的冗余,对静止场景的视频采用视频拼接的方法可以极为有效地消除这种冗余。但是,由于进行视频拼接的前提是静止场景,这就使得视频拼接的应用领域大为受限。所以,下一步研究的目标就是要扩张视频拼接的应用领域,对存在目标运动的视频进行透射模型参数估计,在获得视频全局运动参数的前提下,分离前景和背景,甚至分割出视频对象,最终将之用于实现视频检索

笔者创新点:运用光流法求解模型参数,从而在不影响拼接效果的基础上大大提高了全景图的拼接速度。

参考文献:

- [1] Peleg S, Herman J. Panoramic Mosaics by Manifold Projection[C]// IEEE Computer Society Conference. [s. l.]: [s. n.], 1997: 338-343.
- [2] 蔡丽欢,廖英豪,郭东辉. 图像拼接及其关键技术研究[J]. 计算机技术与发展, 2008, 18(3): 1-4.
- [3] 方青,王博亮. 一种改进的基于比值模板匹配的显微图像拼接[J]. 计算机工程, 2005, 31(15): 159-160.
- [4] 冯林,颜世鹏,孙焱. 图像配准中的一种特定区域轮廓提取算法[J]. 计算机技术与发展, 2006, 16(3): 11-13.
- [5] 朱远平,夏利民. 一种适用于图像拼接的自适应模板匹配算法[J]. 计算机工程与应用, 2003, 39(31): 109-111.
- [6] 侯舒维,郭宝龙. 一种图像自动拼接的快速算法[J]. 计算机工程, 2005, 31(24): 70-72.
- [7] Shum H-Y, Szeliski R. Panoramic Image Mosaics[R]. US: Microsoft Research, 1997.
- [8] Horn B K P, Schunck B G. Determining Optical Flow[J]. Artificial Intelligence, 1981, 17: 185-203.

(上接第238页)

- [5] MODBUS. ORG. MODBUS Application Protocol Specification[EB/OL]. 2008. www.modbus.org.
- [6] 赵素林. 利用多线程实现串口数据的实时图形化显示[J]. 计算机技术与发展, 2006, 16(6): 124-126.
- [7] 江峰,刘高嵩. 串口通讯中系统资源分配问题的研究[J]. 计算机技术与发展, 2006, 16(11): 64-66.
- [8] 郎锐,罗发根. VC++网络通信程序开发指南[M]. 北京:机械工业出版社, 2004.
- [9] Hu Sideng, Zhao Zhengming, Zhang Yingchao. A novel Modbus RTU-based communication system for adjustable speed drives[C]//Vehicle Power and Propulsion Conference, VPPC '08. [s. l.]: [s. n.], 2008: 1-5.
- [11] Peng Dao-gang, Zhang Hao, Yang Li, et al. Design and Realization of Modbus Protocol Based on Embedded Linux System[C]//Embedded Software and Systems Symposia, ICES Symposium '08. [s. l.]: [s. n.], 2008: 275-280.
- [10] Dutertre B. Formal Modeling and Analysis of the Modbus Protocol[C]//IFIP. Critical Infrastructure Protection. Boston: Springer, 2007: 189-204.