

由未标定手持相机拍摄的图片构造全景图

陈 辉 龙爱群 彭玉华

(山东大学信息科学与工程学院 济南 250100)

摘 要 提出了一种利用普通未经标定的相机,在近似固定位置拍摄的图片,来构造水平方向 360° 全景图的方法. 一般来说这种图片可能存在透视变形较大、相邻图片重叠部分较小、亮度差异明显以及相机的倾斜和旋转等问题. 这就使此类图像的配准和全景图拼接问题比采用经标定的相机或专用设备所拍摄的图片较为困难. 因不宜直接采用简单的图像配准技术,文中特设计一种含梯度方向极稳健高效的特征匹配方法来对此类图像进行配准;同时提出了可靠的初始参数估计方法,并将求取相邻图像间透视变换参数的非线性优化过程简化为线性步骤,提高了运算速度和稳定性. 实验结果表明,即使重叠区域很小(16%),该方法仍可获得较为满意的拼接效果.

关键词 图像配准;全景图;图像拼接;特征匹配;虚拟现实

中图法分类号 TP391

DOI号: 10.3724/SP.J.1016.2009.00328

Building Panoramas from Photographs Taken with An Uncalibrated Hand-Held Camera

CHEN Hui LONG Ai-Qun PENG Yu-Hua

(School of Information Science and Engineering, Shandong University, Jinan 250100)

Abstract This paper presents a method for building a full view cylindrical panorama from uncalibrated photos taken naturally with an ordinary hand-held camera held at an approximately fixed location. Such photos usually have large perspective distortion, small overlap, brightness differences, and camera rotations such as tilt and roll. These characteristics make both image alignment and panorama building more difficult than using photos taken by cameras calibrated by special equipment. To align such images, simple registration techniques cannot be applied directly. This paper shows how these images can be registered by a particular robust feature matching scheme, which contains robust information of gradient direction; this paper also proposes a reliable initial parameter estimation method as well as an iterative approach with linear steps to perspective transformation parameter optimization for pairs of adjacent images. Experiments show that the method yields good results even when the overlap region of adjacent photos is as small as 16%.

Keywords image registration; panorama; image mosaic; feature matching; virtual reality

1 引 言

由照片来构造全景图是基于图像虚拟现实系统

(IBVR)的关键建模技术,也在其他可视化领域有着广泛应用^[1-2]. 这些照片可以由专用设备获取,比如全景图摄像机、鱼眼镜头、线性推扫相机^[3],也可以由较便宜的普通照相机或摄像机来获得^[4]. 文献[5]对

图像全景图技术作了较全面的综述。

拍摄全景图的专用设备价格较贵,不易得到。若采用普通照相机,获取的输入是一组部分重叠照片,其首要问题是如何自动配准相邻接图片。为简化问题,通常的做法是用特殊设备将相机固定在特定视点,横向转动相机拍照一系列图片,如此严格控制下拍摄的图片可较容易地进行机械拼接生成 360° 环带状全景图。若用手持相机,则不能保证绝对固定的视点位置和相机转角,配准图像通常采用搜索或优化方法。近来流行的做法是应用经典的光流场运动估计方法求取图像间的透视变换来实现对齐,见文献[4],其变换参数由非线性优化过程求解。该方法对大多数灰度一致的图像序列很有效,但受限与光流场方法不适用图片间灰度差异较大的情况;并且需要用其他方法提供初始变换参数。国内有封靖波等人[6]的工作,主要针对克服亮度差异,利用图像中的曲线梯度特征进行搜索对齐。该方法严格限制相机只有水平方向环绕运动,且依赖于图像含有明显曲线特征。还有基于FFT配准的全景图算法[7]和Brown[8]基于SIFT特征匹配的全景图算法。FFT配准算法又称为相位相关法、梅林变换法等,其最大特点就是速度快,但精度不高,并且因为基于相同图像区域的频谱相关性最大,要求两相邻图像重叠区域必须足够大(相同部分为主);Brown使用了基于尺度放缩不变技术的SIFT[9]算子进行图像配准,算法完全自动完成且效果较好,但是SIFT算法需要在各个尺度上计算梯度方向直方图 and 主方向来构建特征向量,其时间复杂度相对较高。

本文介绍一种稳健高效特征点配准方案以解决上述问题,其中匹配算法基于特征点一阶梯度方向的预筛选和有向梯度窗的相关性,可在较小重叠区域通过少量对应特征点快速求取初始变换参数和优化参数,鲁棒性极强;并介绍如何简化非线性优化过程进一步提高速度和稳定性。另外,本文还将介绍生成最终全景图的一些误差纠正措施。

本研究的目的是使用户可以方便地由未经标定的普通手持相机拍摄图片来自动构造水平方向 360° 全景图,而不必使用复杂的相机标定设备或旋转台等;并允许较宽松的拍摄条件,如图片间较大亮度差异及相机位置微小移动等。文中称此类图片为自然图片。本文假设拍照时尽量保持相机在一个近似固定视点。允许镜头上下倾斜和旋转,但必须保持在一个较小的范围内。拍摄相邻图片时,相机可以有较大水平环绕角,但要保证图片间有一定重叠。在此本文要求重叠部分不小于整幅图像的16%。另外拍

摄所有图片时应保持焦距不变。作者相信这些要求对于拍照者是不难做到的。由于使用特殊设备拍摄的图像的配准方法不适用于自然图片的配准,故在此专门针对自然图片的配准进行研究。

由于相邻图片拍摄于近似同一视点,两幅图片中的对应像素点可近似认为遵循透视变换。因此,我们首先在相邻两幅图片中寻找对应特征点,然后采用一种优化技术求取透视变换来最佳匹配这两幅图片。本文提出的方法包括3步:(1)特征点提取与匹配;(2)由已经匹配的特征点计算透视变换;(3)在所有相邻图片经以上两步配准拼接之后,投影到圆柱形模型上生成全景图。

本文第2节讨论特征匹配;第3节介绍如何求取透视变换;全景图构造和缝隙闭合技术将在第4节介绍;实验和讨论在第5节;第6节总结全文。

2 特征匹配

为将一组自然图片拼接生成全景图,首先需要配准每对相邻图片。本研究的配准目标是求取透视变换矩阵,即求联系两幅图片相似度最大对应点的空间变换。相对于理想最佳配准,该变换应使相邻两幅图片重叠区域的空间形状差异和像素亮度差异达到最小。

2.1 相关工作

图像配准技术已有几十年的研究历史[10]。方法基本分为两类,第一类包括那些最小化图像灰度差异的方法(光流场)[4,11-12],这类方法依赖于待配准图片间像素亮度的一致连续性。其优点是不需要进行特征提取和特征匹配,且可利用所有像素信息。但由于灰度法(光流场法)的局部性,无法适用于图片间亮度差异较大以及像素发生较大偏移的情况。

第二类是基于特征的图像配准方法[8,10,13]。首先从两幅图像中提取特征点。再根据一定度量标准对特征点进行匹配,得出一组对应点集。然后采用优化算法从对应点集合计算空间变换,该变换是两幅图像对应特征点之间的映射。特征点可以是角点,闭轮廓的质心点和区分性良好的纹理点等。特征相似性的度量可以是亮度分布、傅立叶频谱、小波系数、直方图、还可以是基于所提取物体的结构信息等。这类图像配准方法可适用于像素偏移及亮度差异较大的情况,其关键是特征匹配方法的稳健性和效率。

本文采用的方法属于第二类。给定一对有部分重叠的待拼接图片,我们采取下列步骤来确定其对

应特征点: Canny 算子提取边缘, 边缘细化, 检测提取高曲率特征点(曲率大的点, 也即角点), 特征点匹配, 全局聚类匹配检验排除错误匹配的对对应点。

2.2 特征点的选取和匹配

现在描述如何选取图像特征点. 首先, 采用 Canny 边缘检测算子从图像中提取边缘轮廓信息. 在此偏差参数 σ 应取得足够大, 以确保仅较为显著的边缘被保留作后续处理, 而使那些不合需要的噪声、纹理以及透视变形效果最小化. 在取门限值对边缘图像进行二值化及细化之后, 采用高曲率检测器提取角点作为特征点. 本文中高曲率检测器参考文献[13]提出的方法, 改进后能更加精确地定位角点^[14].

$$NCC_{d_{xy}}(I_1, I_2) =$$

$$\frac{1}{2} \left[\frac{\sum_{i=1}^n (\nabla_x I_1(x_i, y_i) - u_{x_1})(\nabla_x I_2(x_i, y_i) - u_{x_2})}{\sqrt{\sum_{i=1}^n (\nabla_x I_1(x_i, y_i) - u_{x_1})^2} \sqrt{\sum_{i=1}^n (\nabla_x I_2(x_i, y_i) - u_{x_2})^2}} + \frac{\sum_{i=1}^n (\nabla_y I_1(x_i, y_i) - u_{y_1})(\nabla_y I_2(x_i, y_i) - u_{y_2})}{\sqrt{\sum_{i=1}^n (\nabla_y I_1(x_i, y_i) - u_{y_1})^2} \sqrt{\sum_{i=1}^n (\nabla_y I_2(x_i, y_i) - u_{y_2})^2}} \right]$$

理想情况下, 当两窗口 I_1 和 I_2 完全相同时其 $NCC_{d_{xy}}$ 具有最大值 1. 对预选后待对应点对 (P_1, P_2) 计算 $NCC_{d_{xy}}$. 当一对对应点对的 $NCC_{d_{xy}}$ 超过门限值 T , 则被暂且认为符合匹配要求. 如果某一特征点出现多个候选对应点, 则取 $NCC_{d_{xy}}$ 值最大的点. 经实验选定 $T=0.6$. 较之直接灰度图像上的相似度量度, $NCC_{d_{xy}}$ 是方向梯度图像上的归一化相似度匹配. 该量度包含灰度变化方向信息, 鲁棒性极强, 且对重叠区域两幅图像之间的亮度变化不敏感. 另外, 梯度对噪声敏感的问题被预处理时的大偏差参数 σ 消除. 考虑到相邻图像经常由于光照或曝光不同而引起亮度差异, 该特征匹配方法显著优越稳健. 大量图像实验也验证如此.

现在候选对应点经过了梯度模板筛选. 然而, 由于模板匹配只依赖于图像特征点邻域的局部信息, 一个好的匹配结果并不一定是正确匹配(比如相邻两幅图片含有大量重复性特征). 为解决该问题, 本文采用一种全局聚类分析算法来检测所有候选对应点是否能产生共同一致的透视变换, 如下.

相邻图片的每一对对应点大致决定了联系两幅图片的环绕角和倾斜角. 这两个角度可表示为 2D 聚类空间上的一个参数点. 将所有对应点都表示到 2D 聚类空间, 只有那些位于最大聚类的点被确认为正确匹配, 其他对应点则被去掉. 这步完成后, 就确定了最终对应特征点集. 下面通过这些对应特征点利用误差最小化来计算相邻两幅图片的透视变换关系.

下面进行特征点匹配. 当搜索某一特征点的另一对应点时, 首先比较特征点与其候选对应点在各自灰度图像中一阶偏导数的正负号. 这步操作可迅速有效去除大量无法匹配的特征点. 因为两个可能匹配特征点的亮度梯度方向应该是一致的, 否则就不可能成为对应点. 预筛选之后, 剩余的候选对应点要采用较复杂的梯度模板作进一步筛选.

令 I_1 和 I_2 分别为以特征点 P_1 和 P_2 为中心的两个窗口图像. 并令 $\nabla_x I_i$ 和 $\nabla_y I_i$ 为窗口 I_i 的差分偏导数(方向梯度), $i=1, 2$. 设 u_{x_i} 和 u_{y_i} 分别为 $\nabla_x I_i$ 和 $\nabla_y I_i$ 的窗口平均值. 我们定义 I_1 和 I_2 的梯度模板量度如下:

3 透视变换

本节描述如何计算映射矩阵 \mathbf{M} . 因假定相机近似位于一固定点进行拍摄, \mathbf{M} 可由透视变换矩阵来逼近. 透视变换的计算基于对应特征点信息, 整个过程包括初始参数估计和参数优化.

假设 (x_i, y_i) 为图像 $I(x_i, y_i)$ 中的一特征点, 相邻接图片 $I'(x_i, y_i)$ 中的对应点为 (x'_i, y'_i) . (x_i, y_i) 与 (x'_i, y'_i) 来自 3D 场景中同一个点, 其映射关系为透视变换

$$(x'_i, y'_i, 1)^T \propto \mathbf{M}(x_i, y_i, 1)^T.$$

该变换的通用模型是 8 参数模型. 而对于本文应用可采用参数较少的旋转模型^[4,14], 如下所述.

设 $\theta_y, \theta_x, \theta_z$ 分别为相机的环绕角(pan)、倾斜角(tilt)和旋转角(roll). 环绕角指相机绕一竖直轴左右水平方向的转动; 倾斜角指相机绕一水平轴上下方向的转动; 旋转角指相机绕垂直于成像平面的轴心作的转动. 设 f 为相机焦距. 则待求透视变换矩阵 \mathbf{M} 可表示如下

$$\mathbf{M} = \mathbf{V} \cdot \mathbf{R}(\theta_y, \theta_x, \theta_z) \cdot \mathbf{V}^{-1} \quad (1)$$

其中, \mathbf{R} 定义为

$$\mathbf{R}(\theta_y, \theta_x, \theta_z) = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix} \cdot \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

V 定义为

$$V = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

在矩阵 M 中有 4 个未知量: $\theta_y, \theta_x, \theta_z$ 和 f . 令一图像点 (x_i, y_i) 经矩阵 M 映射到相邻图像点 $(\tilde{x}_i, \tilde{y}_i)$, 而其对对应点实际位于 (x'_i, y'_i) , 则总匹配误差 E 为

$$\begin{aligned} E &= \sum_{i=0}^n \left\| (x'_i, y'_i)^T - (\tilde{x}_i, \tilde{y}_i)^T \right\|^2 \\ &= \sum_{i=0}^n \left\| (x'_i, y'_i)^T - M \cdot (x_i, y_i)^T \right\|^2. \end{aligned}$$

目标求解未知参数使匹配误差 E 最小化, 是无闭合解的非线性优化问题.

3.1 初始参数估计

首先估算相机焦距 f . 采用文献[4]的方法, 取 4 对互不共线的对应点, 求解由 8 参数通用模型导出的线性方程组, 得 8 参数透视变换. 因含有偏离旋转模型的多余参数, 该通用变换仅用来估算焦距 f , 而不用作优化起始矩阵. 令 $\theta_{y_0}, \theta_{x_0}, \theta_{z_0}$ 为初始估计矩阵 M_0 的环绕角、倾斜角和旋转角. 可利用对应点的位置信息来估算这些角度的初值如下.

假定 $\theta_{z_0} = 0$ (因为在拍照时可以方便地找到水平参照物, 所以用户旋转相机的可能性较小). 环绕角和倾斜角则由图片中所有对应特征点转动和倾斜的平均距离来确定:

$$\begin{aligned} \theta_{x_0} &= \sum_{i=0}^n \left(\arctan \frac{y'_i}{f} - \arctan \frac{y_i}{f} \right) / n, \\ \theta_{y_0} &= \sum_{i=0}^n \left(\arctan \frac{x'_i}{f} - \arctan \frac{x_i}{f} \right) / n, \end{aligned}$$

参见图 1. 至此得出 4 个初始参数: $f, \theta_{y_0}, \theta_{x_0}, \theta_{z_0}$, 进而由旋转模型式(1)求得初始变换 M_0 作为优化起始矩阵.

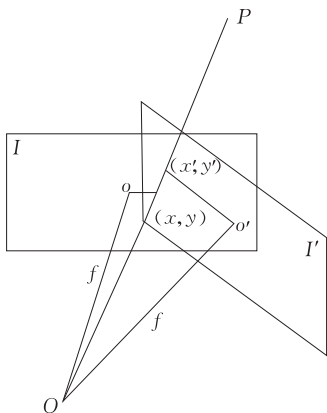


图 1 初始角度估计

3.2 参数优化

由初始变换 M_0 出发采用增量更新法求更精确的透视变换以减小误差 E , 即在 M_0 的基础上进行迭代优化以最终确定 M . 令 M_k 为当前迭代的估计结果, 假设更新量 ΔM_k 可由误差 E 最小化来求得. 其中

$$\Delta M_k = V \cdot \Delta R_k \cdot V^{-1} \quad (3)$$

又根据式(2)

$$\Delta R_k = R(\Delta \theta_y, \Delta \theta_x, \Delta \theta_z).$$

令 $d_x = \Delta \theta_x, d_y = \Delta \theta_y, d_z = \Delta \theta_z$, 因为增量角度 θ 很小时, 可近似认为 $\sin \theta \approx \theta, \cos \theta \approx 1$, 故式(3)展开为

$$\Delta M_k \approx \begin{bmatrix} 1 & -d_z & d_y f \\ d_z & 1 & -d_x f \\ -d_y/f & d_x/f & 1 \end{bmatrix} \quad (4)$$

设点 (x^k, y^k) 由 ΔM_k 映射到点 (x^{k+1}, y^{k+1}) , 则有

$$(x^{k+1}, y^{k+1}, 1)^T = \Delta M_k \cdot (x^k, y^k, 1)^T.$$

因焦距 f 的值远大于增量 d_y 和 d_x , 故可近似认为 ΔM_k 的最后一行为 $(0, 0, 1)$. 经 ΔM_k 迭代更新后, 匹配误差变为

$$\begin{aligned} E &= \sum_{i=0}^n \left\| (x'_i, y'_i) - (x_i^{k+1}, y_i^{k+1}) \right\|^2 \\ &\approx \sum_{i=0}^n \left\| \begin{bmatrix} x'_i \\ y'_i \end{bmatrix} - \begin{bmatrix} 1 & -d_z & f \cdot d_y \\ d_z & 1 & -f \cdot d_x \end{bmatrix} \times \begin{bmatrix} x_i^k \\ y_i^k \\ 1 \end{bmatrix} \right\|^2 \\ &= \sum_{i=0}^n \left\| \begin{bmatrix} \Delta x_i^k \\ \Delta y_i^k \end{bmatrix} - \begin{bmatrix} 0 & -d_z & f \cdot d_y \\ d_z & 0 & -f \cdot d_x \end{bmatrix} \times \begin{bmatrix} x_i^k \\ y_i^k \\ 1 \end{bmatrix} \right\|^2, \end{aligned}$$

这里 $\Delta x_i^k = x'_i - x_i^k, \Delta y_i^k = y'_i - y_i^k$.

由此, 匹配误差最小化问题被简化成了包含 3 个参数 d_x, d_y 和 d_z 的线性优化问题. 将 E 相对于 d_x, d_y 和 d_z 的偏导数分别置零, 得线性方程组:

$$A \cdot (d_x, d_y, d_z)^T = B,$$

其中

$$A = \begin{bmatrix} nf & 0 & -\sum x_i^k \\ 0 & nf & -\sum y_i^k \\ \sum x_i^k f & \sum y_i^k f & -\sum ((y_i^k)^2 + (x_i^k)^2) \end{bmatrix}$$

及

$$B = \left(-\sum \Delta y_i^k, \sum \Delta x_i^k, \sum (\Delta x_i^k y_i^k - \Delta y_i^k x_i^k) \right)^T.$$

由这些方程组可以方便地求解出增量角度参数 d_x, d_y 和 d_z . 注意每次迭代过程中增量矩阵 ΔM_k 应

直接由旋转模型式(1)和(2)计算,以避免用近似式(4)带来累积误差.本算法还避免了非线性优化易陷入局部极值的问题,提高了稳定性.实验证明,一般经过 4、5 次左右迭代, M_k 可以很快收敛到使误差 E 满足要求的值.参数焦距 f 的优化问题将在下节介绍.

4 全景图构造

本节描述构造圆柱形全景图.首先将已配准的图片投影到圆柱体的切平面上,然后再投影到圆柱体表面.为了平滑相邻两幅图像的亮度差异,本文在两幅图片的重叠区域,根据各像素点到重叠区域边界的距离对灰度值进行线性插值.另外,由于相机焦距参数估计误差的存在,当实现全景图闭合拼接时经常有错缝现象出现.竖直方向的缝隙可以通过调整初始倾斜角度来减小;起始时假定第一幅图片的倾斜角为零,如果实际中第一幅图片不是零倾斜,最终拼接结果将是一个倾斜的图;从其斜度可以估算出起始图片的倾斜角,然后使用该初始倾斜角来重

新构造一遍全景图.

对于水平方向的缝隙,将用下面优化方法来弥补.令 L 为全景图的周长.在理想情况下,圆柱体的半径应该为 $f'=L/2\pi$,而实际估算的焦距 f 与 f' 不一定相同.若将 f 纠正为 f' ,相邻图片的环绕角 $\Phi(\text{pan})$ 也要随之进行更新^[14](以保证重叠区域保持不变也即维持配准结果);然后更新各相邻图像间的变换矩阵,再由更新的透视变换矩阵序列计算新的周长 L 、焦距 f' 及环绕角 Φ ,如此迭代数次缝隙即可闭合.

5 实验和讨论

本文所提出方法经过了大量图片序列的测试,图像包括来自胶片相机的、数码相机的以及计算机生成的.图 2 是两幅照片.这两幅自然图片既有明显亮度差异也有显著透视变形(注意图片四角色调发暗是廉价镜头广角拍摄时常见的现象).提取的对应特征点用‘+’标出,如图 3 所示.经过配准后投影到圆柱形表面的效果图如图 4.



图 2 含亮度差别和透视变形的两幅图片

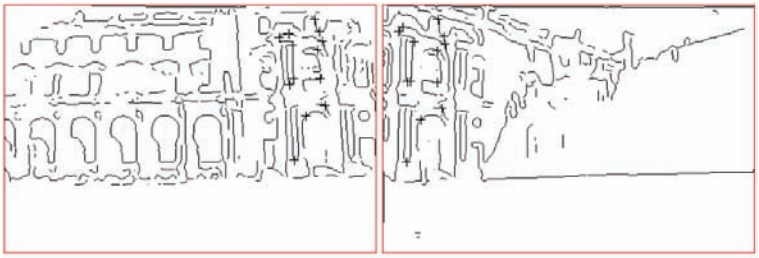


图 3 提取的特征点(十字叉点)



图 4 配准后的图片及图像重叠区域中的偏差

另外本文提出的方法也适用于那些纹理复杂、含重复性信息较多的自然图片的全景图拼接。如



图 5 含大量纹理信息的图片

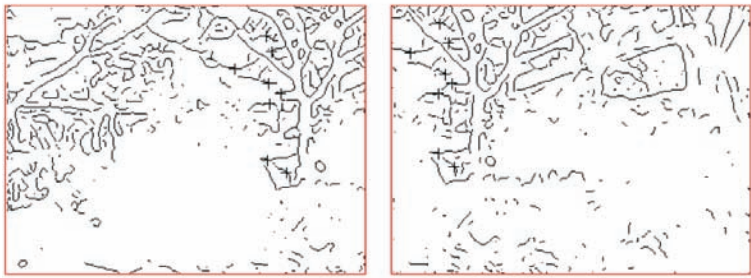


图 6 提取的特征点



图 7 拼接后投影到柱形表面的效果图片

在本文应用中,参数只有 3 个旋转模型包含了相机所有的转动约束条件,因此少量对应特征点就可以达到整体上较为理想的配准效果,从而减小了匹配特征和优化参数的开销.特征匹配中一阶偏导数的正负号匹配迅速去除了大部分非对应点,大大提高了效率. Canny 边缘检测步骤占用整个过程的大部分时间,特征匹配中的一阶偏导图像(方向梯度)直接采用 Canny 边缘检测的中间结果.实验表明本文方法稳定、可靠、速度很快.在奔腾 III 500MHz 微机上,配准一对文中的相邻图片需 2~3s,其中 Canny 边缘检测占大约 1.5s. 这些图片的尺寸(像素点数),宽在范围 320~390,高在范围 240~260 之内.

实验显示本文特征匹配方法十分稳健. Canny 边缘检测时取了较大偏差参数减小了透视变形、噪

图 5 是两幅待拼接的数码相机图片,含有大量纹理细节;图 6 是提取的特征点;图 7 是拼接成的图片.

声和纹理的影响;在梯度图像上的相似度匹配使算法抗亮度差异稳健性强,见例图 3~5,若直接在灰度图像上进行相似度匹配或用光流场方法都不能实现配准. 在全景图应用中通常相机焦距不变也不特别旋转镜头,即无图像放缩变化和大幅旋转,而放缩和旋转不变是 SIFT 算子高计算开销的优势所在,本文算法简洁高效地利用了梯度方向的稳健性.

全景图拼接构造的效果图示于图 8(采用扫描照片)和图 9(采用数码相机拍摄的数字照片),图片序列有明显上下倾斜角. 图 10 是缝隙闭合:将图 8 投影到圆柱体上面时会产生缝隙,如图 10(a);图 10(b)是对缝隙进行了闭合操作,图 10(c)为最终在圆柱体上面显示的结果,为便于观察首尾衔接处未做融合处理.



图 8 展开的圆柱形全景图(Louvre)



图 9 展开的圆柱形全景图(HKU)



(a) 缝隙



(b) 闭合



(c) 在圆柱体上进行闭合

图 10 缝隙闭合

处理终端缝隙时,简便做法是将焦距 f 纠正为 f' 进行强制闭合(拉缩环带图),而忽略相邻图像环绕角的更正.当 f 与 f' 比较接近时,视觉效果上差别不明显.本文采用的技术同时更正焦距和环绕角,可获得物理上正确且效果良好的缝隙闭合.

6 结 论

本文提出了一种利用未经标定的手持相机拍摄图片来构造全景图的方法.此方法对拍照者的要求不高,整个拼接过程是全自动实现的.本方法的关键首先是针对自然图片利用梯度方向的稳健性实现了高效可靠的特征点匹配,并设计了旋转模型初始估计方法,将透视变换 3 参数模型的非线性优化问题简化为线性迭代过程,快速精确.在柱形全景图的构造过程中,修正了倾斜、弥合了接缝.本文方法经过了大量各类图片的拼接测试,结果证明效果令人满意,技术非常实用.由于相机未经标定,图片仍存在非线性几何变形未纠正和四角发暗等问题有待进一步解决.

参 考 文 献

[1] Chen S C. Quicktime VR—An image-based approach to virtual environment navigation//Proceedings of the Computer

Graphics(SIGGRAPH'95). New York, USA, 1995: 29-38

[2] Zhu Z, Hanson A R, Riseman E M. Generalized parallel-perspective stereo mosaics from airborne video. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(2): 226-237

[3] Peleg S, Rousso B, Rav-Acha A, Zomet A. Mosaicing on adaptive manifolds. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(10): 1144-1154

[4] Szeliski R, Shum H-Y. Creating full view panoramic image mosaics and environment maps//Proceedings of the Computer Graphics(SIGGRAPH'97). Los Angeles, USA, 1997: 251-258

[5] Xie Kai, Guo Heng-Ye, Zhang Tian-Wen. A survey of image Mosaics technology. Journal of Electronics, 2004, 32(4): 630-634(in Chinese)

(解凯, 郭恒业, 张田文. 图像 Mosaics 技术综述. 电子学报, 2004, 32(4): 630-634)

[6] Feng Jing-Bo, Su Zhi-Xun, Liu Xiu-Ping. An similar-curve based automatic Mosaic algorithm of panoramic image. Chinese Journal of Computers, 2003, 26(11): 1604-1608(in Chinese)

(封靖波, 苏志勋, 刘秀平. 一种基于相似曲线的全景图自动拼接算法. 计算机学报, 2003, 26(11): 1604-1608)

[7] Li Xiao-Ming, Zhao Xun-Po, Zheng Lian. An image registration technique based on Fourier-Mellin transform and its extended applications. Chinese Journal of Computers, 2006, 29(3): 466-422(in Chinese)

(李晓明, 赵训坡, 郑链. 基于 Fourier-Mellin 变换的图像配准

方法及其应用拓展. 计算机学报, 2006, 29(3): 466-422)

- [8] Brown M, Lowe D G. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 2007, 74(1): 59-73
- [9] Lowe D G. Distinctive image features from scale invariant key points. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [10] Brown G. A survey of image registration techniques. *ACM Computing Surveys*, 1992, 24(4): 325-376
- [11] Bergen J R, Anandan P, Hanna K J, Hingorani R. Hierarchical model-based motion estimation//*Proceedings of the*

European Conference on Computer Vision (ECCV92). Santa Margarita Ligure, Italy, 1992: 237-252

- [12] Chen Hui. Gradient-based approach for fine registration of panorama images. *Journal of Computer Science and Technology*, 2004, 19(5): 691-697
- [13] Li H, Manjunath B S, Mitra S K. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 1995, 4(3): 320-334
- [14] Chen Hui. Building panoramas from photographs taken with a hand-held camera [Ph. D. dissertation]. University of Hong Kong, Hong Kong, 2002: 47-49



CHEN Hui, born in 1963, Ph. D., professor. Her research interests include computer vision and image processing.

LONG Ai-Qun, born in 1963, M. S., lecturer. Her research interests include optics and image processing.

PEN Yu-Hua, born in 1966, Ph. D., professor, Ph. D. supervisor. Her research interests include multi-resolution analysis and image processing.

Background

Constructing a full view panorama of a 3D scene from a sequence of partially overlapping photos is a critical modeling task in building an image-based virtual reality system^[1-2]. Such photos will also be referred to as images, views, or mosaics in the following. Photos can be captured by special equipment^[3], such as a panoramic camera or a fish eye lens camera with a large field of view, or by less expensive equipment such as an ordinary camera or video-camera^[4]. Literature^[5] gives an comprehensive review of the panorama technique.

If using ordinary camera, the problem is to stitch together adjacent views of a scene into an image having a larger field of view. Recently, a popular approach is to apply the classic optical motion estimation to find the perspective alignment parameters, see Szeliski and Shum^[4], the parameters are resolved by nonlinear optimization. The method is effective to most of the image sequences with identical lighting condition. However, with the limitation of the optical method itself, it is not applicable to images with large intensity changes, and it also needs other method to provide initial parameters. The authors also see other panorama literatures published while our paper is in reviewing; they are FFT-based registration method^[7] and Brown's^[8] SIFT-based method. The FFT-based method runs fast but it has low resolution and requires large overlap; Brown et al apply the invariant SIFT feature matching method for registration of adjacent image sequence. It is fully automatic and robust, but it

requires to compute the histogram of gradient direction at each scale to construct the feature vector, which has relatively much higher complexity and very time consuming. Here, the authors present a simpler approach for feature matching that makes use of the robust information of gradient direction as well, but in a more efficient way.

In this paper the authors assume that the photographer uses reasonable efforts to keep the camera at a fixed location while taking a series of natural photos. Tilting and rolling are allowed but should be kept to a minimum. Panning angles can be relatively large, but must be small enough to allow some overlap between adjacent views. They require an overlap of at least 16% of the width of each photo. The focal length of the lens should be kept the same for all photos. They believe that these requirements are not overly stringent or inconvenient to a photographer. The authors present here a new method for compositing natural photos, as most methods developed for compositing photos taken with special calibration devices are not applicable to natural photos.

Our method consists of three main steps;

(1) automatic identification of corresponding feature points;

(2) computing a perspective transformation for image alignment from corresponding feature points;

(3) mapping the panoramic mosaic onto a cylindrical model, after all adjacent images have been stitched together by applying steps (1) and (2).