

基于对数极坐标映射的图像拼接方法

李忠新 茅耀斌 王执钊

(南京理工大学自动化系, 南京 210094)

摘要 图像拼接在基于图像的绘制、视频检索以及景物匹配等领域有着广泛的应用, 为了获取大画面宽视场的场景表示, 针对存在旋转及缩放变化的图像, 提出了一种基于对数极坐标映射的图像拼接方法, 该方法先将图像从笛卡儿坐标空间转换到对数极坐标空间, 使得笛卡儿坐标空间中图像的旋转和缩放转换为对数极坐标空间中图像的二维平移, 这样可直接利用相位相关法来估算出图像间的旋转角度和缩放因子, 然后以此作为初值, 再采用非线性最小化优化算法进一步求精图像间的运动参数来实现图像的配准, 最后通过图像融合来实现图像的拼接。实验结果表明, 该方法是有效的。

关键词 对数极坐标映射 图像拼接 图像配准

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2005)01-0059-05

A Method of Image Mosaicing Using Log-polar Coordinate Mapping

LI Zhong-xin, MAO Yao-bing, WANG Zhi-quan

(Department of Automation, Nanjing University of Science & Technology, Nanjing 210094)

Abstract Image mosaicing has many applications, such as image based rendering, video retrieval, scene matching, etc. Aiming at images with rotation and scale, a method of image mosaicing based on log-polar coordinate mapping is put forward to obtain scene representation of wide view field. In this method, firstly each image is transformed from Cartesian space to log-polar coordinate space, which causes scale and rotation in Cartesian space to be manifested translation in log-polar coordinate space, and the translation between the pair of input images is calculated, then phase correlation method is used to estimate rotation angle and scale factor between the pair of input images, and then these motion parameters are refined by nonlinear least square iterative optimization algorithm, lastly image blending is imposed to implement image mosaicing. The experiments show that it is an available method of image mosaicing.

Keywords log-polar coordinate mapping, image mosaicing, image registration

1 引言

所谓图像拼接就是将一个场景的多个相重叠子画面配准, 并拼接成一个较大画面的表示方法, 其可用于生成全景图及宽视场的场景表示, 在基于图像的绘制^[1]、视频检索^[2]及景物匹配^[3]等领域有着广泛的应用。

图像拼接技术的关键问题是图像配准, 即寻找将两幅图像间相重叠的部分对齐所需的变换。这其中涉及到的问题就是寻找图像间的对应匹配。目前

图像匹配的一般解决方法是基于图像间的相关性, 但这些方法大都采用寻找局部对应或抽取重要特征的方法来进行匹配, 如 Szeliski 提出的基于运动 (Motion-based) 的方法^[4], Kyung 提出的等距匹配法^[5], Peleg 和 Herman 提出的多重投影法 (Manifold-projection)^[6], 其缺点是缺乏统一的匹配策略, 而且还易受到特征附近噪声的影响; 另外, 这些方法只能得到局部最优解。文献^[7]提出了先将图像进行傅里叶变换, 然后利用图像的傅里叶谱来计算图像间的相对位移, 文献^[8]将该方法进一步扩展, 即先计算出图像间的缩放因子、相对位移和旋转角度, 但

基金项目: 高等学校博士学科点专项科研基金项目(20020288024)

收稿日期: 2003-07-29; 改回日期: 2004-06-25

第一作者简介: 李忠新(1976~), 男, 2000年获南京理工大学硕士学位, 现为南京理工大学自动化系博士研究生。主要从事图像处理、虚拟现实等方面的研究。E-mail: lzxj@sohu.com

是要求图像间的缩放因子和旋转角度不能过大,否则得不到正确的结果,这类方法虽然可以得到全局解,但由于傅里叶变换的空间局部性,致使其所得到的往往是接近于最优解的结果。综上所述,本文提出了新的图像拼接算法,即首先将图像从笛卡儿坐标空间转换到对数极坐标空间,并根据对数极坐标映射的性质,直接利用文献[7]提出的相位相关法来得到接近于全局最优解的图像运动参数,然后通过迭代优化算法进一步求精图像间的运动参数来实现图像的拼接。实验结果表明,该方法可以获得比较理想的拼接效果。

2 对数极坐标映射

众所周知,一般图像中像素的位置可以用笛卡儿坐标 (x, y) 来表示,也可以用极坐标 (r, θ) 来表示,对于选定的坐标原点 (x_0, y_0) ,它们之间满足如下关系:

$$\begin{cases} r = \sqrt{(x - x_0)^2 + (y - y_0)^2} \\ \theta = \arctan\left(\frac{y - y_0}{x - x_0}\right) \end{cases} \quad (1)$$

这里取坐标原点 (x_0, y_0) 为 $(0, 0)$,其用复数 z 可以表示为

$$z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta} \quad (2)$$

令 $\omega = \ln z = p(z) + iq(z) = \ln r + i\theta$,则笛卡儿坐标转换为对数极坐标的映射方程为

$$p(r, \theta) = \ln r, q(r, \theta) = \theta \quad (3)$$

由此可见,当笛卡儿空间中的图像相对于坐标原点发生了缩放和旋转变化,例如图像放大了 r_0 倍,旋转了 θ_0 角度,即变化后新的相应极坐标为 $(r_0 r, \theta + \theta_0)$ 时,则其取对数极坐标映射之后有

$$\omega = \ln z = \ln r_0 r e^{i(\theta + \theta_0)} \quad (4)$$

即 $p(r, \theta) = \ln r + \ln r_0, q(r, \theta) = \theta + \theta_0$ 。以上说明笛卡儿坐标空间中图像的缩放变化相当于对数极坐标空间图像在水平方向发生位移,而旋转变化则相当于垂直方向发生位移,这样通过将图像从笛卡儿坐标空间转换到对数极坐标空间,就可以采用文献[7]提出的相位相关法直接估算出图像间的缩放因子和旋转角度,然后以此作为图像间运动参数的初值。

3 图像的拼接

3.1 运动参数的初值估计

对于相邻图像间存在的平移和旋转运动,可以

采用如下刚体运动模型来描述:

$$\begin{cases} \hat{x} = x \cos t_3 - y \sin t_3 + t_1 \\ \hat{y} = x \sin t_3 + y \cos t_3 + t_2 \end{cases} \quad (5)$$

其中 t_1, t_2 分别是水平和垂直方向的平移量, t_3 是图像的旋转角度。本文先利用第2节介绍的对数极坐标映射方法将图像由笛卡儿坐标空间转换到对数极坐标空间,再利用相位相关法估计出运动参数 $\{t_1, t_2, t_3\}$ 的初值,然后通过非线性最小方差优化方法进一步求精这些运动参数,最后根据求解出来的参数进行图像拼接。在第2节中,假定了对数极坐标映射时两幅图像的坐标原点(图像的中心)均为 $(0, 0)$,但在实际情况下,由于图像之间发生了平移和旋转,坐标原点的位置也相应发生了变化,除非已知图像间的相对位移,才可以利用对数极坐标映射的性质来计算图像的运动参数,然而图像间的相对位移恰恰是需要计算的,为了解决这个问题,可采用如下算法:

①从图像 I_1 中取出子块图像 \tilde{I}_1 ,为了保证子块图像完全包含在图像 I_2 中,可取图像 I_1 的中心区域;

②将子块图像 \tilde{I}_1 的转换为对数极坐标空间上的图像为 \tilde{I}_1^p (上标 p 表示对数极坐标);

③以图像 I_2 中的每一个像素 (x, y) 为中心点,截取与 \tilde{I}_1 同大小的子块图像 \tilde{I}_2 ,而将其转换到对数极坐标空间上的图像为 \tilde{I}_2^p ;

④利用相位相关法计算出图像 $\tilde{I}_1^p, \tilde{I}_2^p$ 间的位移量 $(\Delta x, \Delta y)$,并记录出现最大相关峰值的位置 (x, y) 以及对应的位移 $(\Delta x, \Delta y)$;则两幅图像间的缩放因子 $s = \Delta x$,旋转角度 $t_3 = \Delta y$,平移量 $(t_1, t_2) = (x, y)$ 。

由于在算法中采用了全局搜索的方法来确定图像间的位移量,导致了算法的计算量增加,故本文引入由粗到精的图像金字塔算法,即首先在低分辨率的图像上进行运动参数初值的估计,然后进一步地在高分辨率的图像上,根据上一层图像的估计结果缩小搜索范围,以提高计算效率。

此外,由于图像初始运动参数的估计精度主要受对数极坐标映射误差的影响,例如将笛卡儿坐标空间的 128×128 窗口映射为对数极坐标空间的 256×256 的窗口,并以窗口中心为变换原点,则 p 轴的分辨率为 $\ln 64\sqrt{2}/256 = 0.0176$,对数极坐标空间 $p = 255$ 对应笛卡儿坐标空间约2个像素距离, q 轴的分辨率为 $360^\circ/256 = 1.4063^\circ$,所以在利用对数极坐标法估计出图像间的运动参数的初值后,还

需利用优化算法进一步地求精图像间的运动参数。

3.2 运动参数的迭代求精

根据上一节计算出来的运动参数初值,本节采用最小化图像间灰度差 e_i 的平方和来进一步求精运动参数,即

$$E = \sum_i (I_2(\hat{x}_i, \hat{y}_i) - I_1(x_i, y_i))^2 = \sum_i e_i^2 \quad (6)$$

也就是求解如下的非线性最小二乘问题:

$$\min \sum_i (I_2(\hat{x}_i, \hat{y}_i) - I_1(x_i, y_i))^2 = \min \sum_i e_i^2 \quad (7)$$

(x_i, y_i) 、 (\hat{x}_i, \hat{y}_i) 为图像 I_1 、 I_2 间的对应点,若 I_1 中某一点在 I_2 中的对应点落在 I_2 的边界之外,则该点对上式没有贡献。若 (\hat{x}_i, \hat{y}_i) 不是整数点,则采用双线性插值方法计算它在 I_2 中的灰度值。这里使用了 Levenberg-Marquardt 非线性优化算法来求解上述问题。该算法首先要计算出 e_i 分别关于各运动参数 $\{t_1, t_2, t_3\}$ 的偏导数,即

$$\begin{aligned} \frac{\partial e_i}{\partial t_1} &= \frac{\partial I_2(\hat{x}, \hat{y})}{\partial \hat{x}} \\ \frac{\partial e_i}{\partial t_2} &= \frac{\partial I_2(\hat{x}, \hat{y})}{\partial \hat{y}} \\ \frac{\partial e_i}{\partial t_3} &= (-x \sin t_3 - y \cos t_3) \frac{\partial I_2(\hat{x}, \hat{y})}{\partial \hat{x}} + \\ &\quad (x \cos t_3 - y \sin t_3) \frac{\partial I_2(\hat{x}, \hat{y})}{\partial \hat{y}} \end{aligned} \quad (8)$$

其中 $\left(\frac{\partial I_2(\hat{x}, \hat{y})}{\partial \hat{x}}, \frac{\partial I_2(\hat{x}, \hat{y})}{\partial \hat{y}}\right)$ 为图像 I_2 在 (\hat{x}, \hat{y}) 点的灰度梯度。根据上面计算出来的偏导数就可以构造出近似的 Hessian 矩阵 A 和加权梯度矢量 b :

$$\begin{aligned} A &= [a_{kl}] = \left[\sum_i \frac{\partial e_i}{\partial t_k} \cdot \frac{\partial e_i}{\partial t_l} \right], \\ b &= [b_k] = \left[- \sum_i e_i \frac{\partial e_i}{\partial t_k} \right], \\ &\quad (k, l = 1, 2, 3) \end{aligned} \quad (9)$$

然后就可通过增量 $[\Delta t_1, \Delta t_2, \Delta t_3]^T = (A + \lambda I)^{-1} b$ 来不断地更新各运动参数 $\{t_1, t_2, t_3\}$, 直到误差 E 小于给定的阈值或者达到了给定的迭代次数为止,其中 λ 为迭代因子, I 为 3×3 单位阵。具体算法可以参见文献[9]。

4 图像融合

求出图像间的运动参数之后,如果将图像的重叠区域直接平均叠加,那么由于图像的边缘和图像

间可能存在亮度差,因而会在拼接处产生明显的接缝,为了有效地消除拼接痕迹,必须将图像重叠区域的像素值按一定的权值相加来合成新的图像,其重叠图像区域的像素值可表示为

$$I(x, y) = \frac{\sum_i w_i(x, y) I_i(x, y)}{\sum_i w_i(x, y)} \quad (10)$$

式中,权值 $w_i(x, y) = \left(1 - \left|\frac{x - W/2}{W/2}\right|\right) \left(1 - \left|\frac{y - H/2}{H/2}\right|\right)$, W 为第 i 幅图像的宽度, H 为第 i 幅图像的高度,由于这样可以使得相邻图像能够平滑地过渡,从而大大地减轻了图像间的拼接痕迹。

5 实验结果

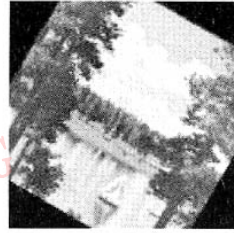
图1说明了图像的对数极坐标变换的性质,图1(a)为 128×128 的原始图像,图1(b)为与图1(a)对应的大小为 128×128 的对数极坐标映射图像(以图1(a)的中心位置作为变换的原点,以水平方向为对数极坐标轴,以下图像的坐标映射均采用相同的设定),图1(c)为图1(a)顺时针旋转 30° 的结果,图1(d)为图1(c)在对数极坐标空间上对应的图像。图1(e)为图1(a)放大后的图像,缩放因子 $s = 1.5$,图1(f)为图1(e)转换到对数极坐标空间上的图像,图1(g)为图1(a)经放大和旋转后的图像,缩放因子 $s = 1.5$,旋转角度 $\theta = 30^\circ$,图1(h)为图1(g)转换到对数极坐标空间上的图像。表1列出了对图1中对数极坐标空间上的图像采用相位相关法计算出的运动参数的估计值和实际值。从图1及表1可以看到,在笛卡儿坐标空间中,图像的旋转表现为对数极坐标空间图像沿着 θ 轴(垂直方向)的平移;反之,图像在对数极坐标空间中沿着 θ 轴的平移表现为图像在笛卡儿坐标空间上的旋转;图像在笛卡儿坐标空间中图像的缩放变化表现为图像在对数极坐标空间中水平方向的平移;反之,图像在对数极坐标空间中水平方向的平移表现为图像在笛卡儿坐标空间上的缩放变化。图2显示了采用本文提出的算法获得的图像拼接结果,图2(a)、图2(b)为原始图像(拍摄时相机距景物较远,可以假定相机焦距未发生变化,故不考虑图像间的缩放变化),实验时,首先采用3.1节提出的算法计算出两幅图像间的相对位移和旋转角度,然后以此为初值,再通过



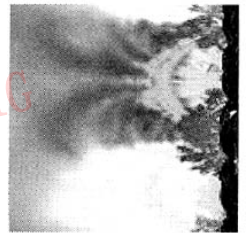
(a) 原始图像



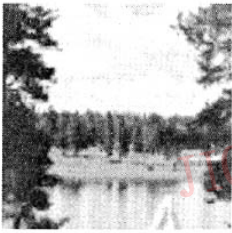
(b) 图 1(a) 转换到对数极坐标系上的图像



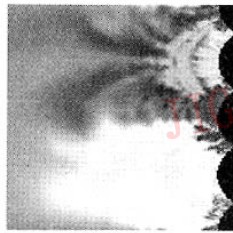
(c) 图 1(a) 沿顺时针方向旋转 30° 后的图像



(d) 图 1(c) 转换到对数极坐标系上的图像



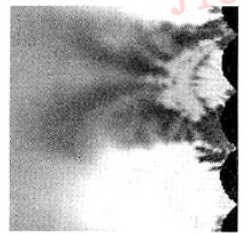
(e) 图 1(a) 放大 1.5 倍后的图像



(f) 图 1(e) 转换到对数极坐标系上的图像



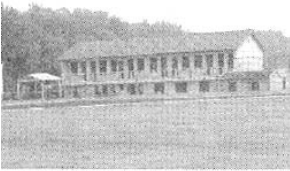
(g) 图 1(a) 放大 1.5 倍并旋转 30° 后的图像



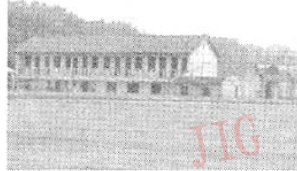
(h) 图 1(g) 转换到对数极坐标系上的图像

图 1 对数极坐标映射

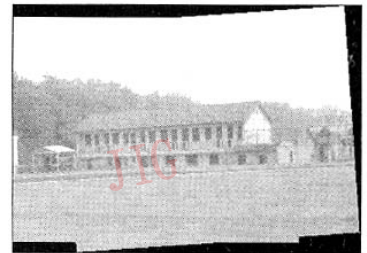
Fig. 1 Log-polar coordinate mapping



(a) 图像 1



(b) 图像 2



(c) 拼接结果

图 2 图像拼接

Fig. 2 Result of image mosaicing

表 1 图像间运动参数的估计值和实际值比较

Table 1 The estimated and real motion parameters between images

运动参数	估计值	实际值	相关峰值
旋转角度(°)	30.94	30	0.4454
缩放倍数	1.48	1.5	0.2157
旋转角度(°), 缩放倍数	30.94, 1.48	30, 1.5	0.1849

表 2 图像间运动参数的初值和求精值

Table 2 The initial and refined motion parameters between images

运动参数	初始值	求精值
平移	(60, 3)	(59.85, 3.79)
旋转角度(°)	3.97	3.99

3.2 节描述的非线性最小化算法进一步求精, 最后采用第 3 节的方法进行图像融合, 图 2(c) 为图像拼接结果。表 2 列出了图 2 图像间运动参数的初始值和求精的结果。

6 结论

针对图像间存在平移、旋转以及缩放变化的情形, 本文利用对数极坐标映射的性质, 首先将笛卡儿

坐标空间中图像的旋转和缩放转换为对数极坐标空间中图像的二维平移,然后通过相位相关法来得到接近于全局最优解的图像的运动参数,最后通过非线性迭代优化算法进一步求精这些参数来实现图像的拼接。实验结果表明,该方法可获得较为满意的拼接效果。由于本文提出的方法只考虑了在图像间存在刚体运动的情形下拼接,而在实际拍摄的图像中还存在噪声、畸变以及更复杂的运动的情形,故对上述情形下的图像进行鲁棒的拼接,尚需要作进一步的研究。

参考文献(Reference)

- 1 Chen S E. QuickTime VR—An image-based approach to virtual environment navigation[A]. In: Proceedings of the ACM SIGGRAPH Conference on Computer Graphics [C], Los Angeles, CA, USA, 1995:29 ~ 38.
- 2 Irani M, Anandan P. Video indexing based on mosaic representations [J]. Proceedings of the IEEE, 1998, **86**(5):905 ~ 921.
- 3 Plesea L, Jacob J. Building large scale mosaics from landsat data [A]. In: Proceedings of the ACM Workshop on Advances in Geographic Information Systems[C], Washington, DC, USA, 2000: 180 ~ 182.
- 4 Szeliski R. Image mosaicing for telereality applications [A]. In: Proceedings of the IEEE Workshop on Applications of Computer Vision [C], Sarasota, FL, USA, 1994:44 ~ 53.
- 5 Kyung H J, Soon K J. Constructing cylindrical panoramic image using equidistant matching [J]. Electronics Letters, 1999, **35** (20): 1715 ~ 1716.
- 6 Peleg S, Herman J. Panoramic mosaics by manifold projection[A]. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C], San Juan, PR, USA, 1997: 338 ~ 343.
- 7 Kuglin C D, Hines D C. The phase correlation image alignment method [A]. In: Proceedings of the IEEE International Conference on Cybernetics and Society [C], San Francisco, CA, USA, 1975: 163 ~ 165.
- 8 Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation, and scale invariant image registration [J]. IEEE Transactions. on Image Processing, 1996, **5**(8): 1266 ~ 1271.
- 9 Szeliski R. Video mosaics for virtual environments [J]. IEEE Computer Graphics and Applications, 1996, **16**(2): 22 ~ 30.