

基于相位相关的全景图像拼接

胡社教, 葛西旺, 苏 菱

(合肥工业大学 计算机与信息学院, 安徽 合肥 230009)

摘 要:全景图拼接是实现虚拟场景构建的一种重要方法。文章采用基于傅氏功率谱的频域相关技术来寻找相邻图像的重叠部分,首先将相机拍摄的实景照片投影到圆柱面上,以柱面图像的形式存储;然后精确定位相邻两张照片间的重叠位置并融合相邻图像的亮度,通过图像拼接形成完整的 360°环境视景;采用互功率谱中的相位信息进行图像配准,对图像间的亮度变化不敏感,可以有效避免局部最优的现象;实验结果证明,该算法具有一定的抗干扰能力和较高的匹配精度。

关键词:全景图; 频域配准; 柱面投影; 图像拼接

中图分类号: TP751.1

文献标识码: A

文章编号: 1003-5060(2007)01-0029-04

Panoramic image mosaic based on domain correlation

HU She-jiao, GE Xi-wang, SU Ling

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: In this paper, the overlapping part of adjacent images is sought by adopting the method which is based on the frequency domain in the Fourier power technology. First, the digital photos are projected to the cylindrical surface and registered, then the two adjacent images are mosaiced and the lightness is blended. By piecing together the images in sequence, the panoramic image of 360 degrees can be gained. The presented registration algorithm which uses phase information of domain correlation is non-sensitive to the lightness change, so that local optimum can be avoided effectively. Experiment results prove that this algorithm is robust and has certain capability of interference rejection and its matching accuracy is much higher.

Key words: panoramic image; frequency domain registration; cylindrical projection; image mosaic

0 引 言

基于图像绘制 IBR (Image-based rendering) 是近年发展起来的一种生成具有照片真实感虚拟场景的绘制方法。IBR 技术采用计算机图形学的思想构造虚拟场景,又使用计算机视觉技术从已知的图像中生成新的图像。传统意义上的虚拟场景构建是基于几何的,即用数学意义上的曲线、曲面等数学模型预先定义好虚拟场景的几何轮廓,再采取纹理映射、光照等数学模型加以渲染^[1]。现有的几何造型技术可以将极复杂的环境构造出来,存在的问题是极为繁琐,而且在真实感程度、

实时输出等方面有着难以跨越的鸿沟。与传统的图形学意义上的场景意义不同,IBR 技术具有以下优点:

- (1) 计算量独立于场景复杂程度,因为大量的图形细节可以直接从图像中获取;
- (2) 造型数据可利用视觉技术获得,节省了劳动;
- (3) 对物理特性(如光照)的设计要求降低,因为图像中已经包含此类信息。

基于图像的绘制技术的最大优势在于其采样样本是真实世界中的图像,具有建模时间更短、绘制速度更快以及真实感更强等特点。

收稿日期:2005-12-09;修改日期:2006-03-01

作者简介:胡社教(1964-),男,安徽歙县人,合肥工业大学副教授,硕士生导师。

1 全景图

全景图(Panorama)技术是指通过图像处理技术形成一个全视角、高分辨率的 360° 图像。它是观察者环视一圈所能看到景色的整合重现,能较好地表达完整的周围信息。全景图的获取方法可分为 2 类: 使用全景照相机等专用设备直接拍摄环境图像,生成全景图。该方法具有速度快、鲁棒性好等特点,但这样的设备一般比较昂贵,难以普及。采用普通数码相机拍摄的序列照片,通过图像拼接的方法来获得宽视角、高分辨率的全景图。为了拼接出的全景图能够保持透视效果以及视觉上的一致性,需要将拍摄的图像序列投影到一个统一的表示面上。常用的表示面有柱面、立方体和球面。其中,柱面全景图因其数据存储空间结构简单而成为目前研究的热点^[2]。

在全景图的生成过程中一般要求求解图像间的三维旋转量和单应性矩阵,为了简化这种求解图像间多个对准系数的复杂过程可对相机的运动作些限制,拍摄时保持相机的镜头水平绕光心旋转。图像间只有围绕 y 轴的旋转角度,不存在垂直方向上的倾斜角度。把相互重叠的图像映射到统一的约束空间上,只要计算出相邻图像间 2 个方向上的位移量,就可以将图像对准整合,再通过柱面投影,即可形成完整的柱面全景图^[3]。本文使用的拼接柱面全景图方法可分为 3 个部分:

(1) 利用频域相位相关性构造函数,该函数对 2 幅输入图像分别进行二维 Fourier 变换,返回相邻两图像之间偏移位置。

(2) 计算一组 360° 实景图片的相机焦距,对图像序列分别做柱面投影。

(3) 逐次调用第(1)部分中函数对投影后图像进行拼接,并对光照进行处理,完成柱面全景图的生成。

2 投影变换及焦距估计

在进行柱面投影时,只要找到柱面坐标和直角坐标直接的对应关系,就可以进行投影变换了。柱面投影原理图,如图 1 所示。

设投影前图像 I_1 宽为 width,高为 height,令

$$\text{width}_0 = \text{round}(\text{width}/2)$$

$$\text{height}_0 = \text{round}(\text{height}/2) \quad (1)$$

则投影后图像的坐标 (x, y) 同 I_1 的坐标 (w, h) 有如下对应关系,即

$$x = \text{round}(\text{sqrt}(f^2 / ((w - \text{width}_0)^2 + f^2)) \times$$

$$(w - \text{width}_0)) + \text{width}_0$$

$$y = \text{round}(\text{sqrt}(f^2 / ((w - \text{width}_0)^2 + f^2)) \times (h - \text{height}_0)) + \text{height}_0 \quad (2)$$

其中, $w = 1 - \text{width}$, $h = 1 - \text{height}$ 。

根据(2)式投影结果,如图 2 所示。

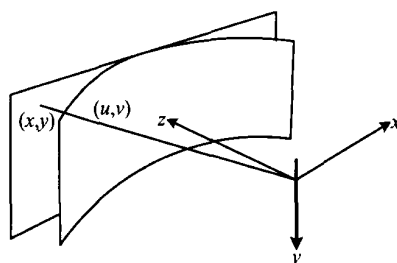


图 1 柱面投影原理图

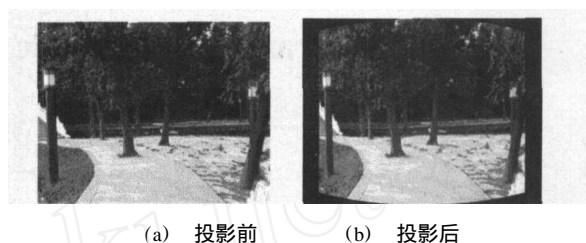


图 2 柱面投影结果

应用柱面投影公式对图像进行投影变换时,焦距 f 是重要参数。设投影前照片序列两两之间的平移量分别设为 C_1, C_2, \dots, C_{12} , 其中 C_k 表示第 k 幅照片与第 $k+1$ 幅照片之间的水平位移,则可计算出焦距的初始值 f_0 为

$$f_0 = \sum_{k=1}^{12} C_k / 2 \quad (3)$$

为了尽可能地逼近实际拍摄焦距,也可以采用迭代算法来减小误差:

(1) 由(3)式初步算出焦距 f_0 , 令 $f = f_0$ 。

(2) 以 f 作为柱面投影参数将该组照片投影到统一圆柱面上。

(3) 对投影后图片 I_1, I_2, \dots, I_{12} 进行处理后再次使用相位相关法,得出相对位移,据此重计算焦距 f_0 。

(4) 如果 $|f - f_0| > \epsilon$, 即精度不满足要求,则令 $f = f_0$, 程序转入(3)。

(5) 直至 $|f - f_0| < \epsilon$, 得到精确的 f 值。

3 基于相位相关的图像配准

图像拼接是根据相邻 2 幅图像重叠区域的相似性来确定的,包括图像配准和图像融合 2 个关键环节。图像配准有基于区域、基于特征和基于

相位的方法;基于区域的方法通过比较 2 幅图像亮度(或颜色)差异,并使之最小化来寻找最佳匹配点,只要时间足够长,就可以拼接得足够精确。基于特征的拼接方法采用寻找局部对应和抽取图像之间的对应特征点的方法进行匹配,如 Harris 角点算法和 SUSAN 角点算法,该算法的优点是计算量比较小,速度快,但是不一定能找到合适的图像特征,结果也可能是局部最优解^[4]。

本文采取的是一种基于频域的方法,通过求得图像在频域上的相位相关特点来找到特征位置,从而进行图像拼合。

相位相关法最早于 20 世纪 70 年代被提出来,是基于傅氏功率谱的相关技术。该方法仅利用互功率谱中的相位信息进行图像配准,对图像间的亮度变化不敏感,而且所获得的相关峰尖突出,具有一定的鲁棒性和较高的配准精度^[5]。

假设 $f(x, y)$ 表示尺寸为 $M \times N$ 的图像,该函数的二维离散傅里叶变换(DFT)为

$$F(u, v) = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)} \quad (4)$$

其中, $F(u, v)$ 是复变函数; u, v 是频率变量, $u = 0, 1, \dots, M-1, v = 0, 1, \dots, N-1$; x, y 是空间或图像变量。

二维离散傅里叶逆变换(IDFT)为

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)} \quad (5)$$

其中, $x = 0, 1, \dots, M-1; y = 0, 1, \dots, N-1$ 。

设 2 幅图像 I_1, I_2 的重叠位置为 (x_0, y_0) , 则图像 I_1, I_2 的互功率谱为

$$\frac{I_1(\cdot) \times I_2^*(\cdot)}{I_1(\cdot) \times I_2(\cdot)} = e^{-j2\pi(x_0 + y_0)} \quad (6)$$

其中, $*$ 为共轭符号,对(6)式两边进行傅里叶逆变换将在 (x_0, y_0) 处产生一个函数,因此只要检测(6)式傅里叶逆变换结果最大值的位置,就可以获得 2 幅图像间的平移量 (x_0, y_0) 。具体算法步骤和函数如下:

(1) 读入 2 幅照片 I_1, I_2 (函数输入),并转换为灰度图像。

(2) 分别对 I_1, I_2 作二维傅里叶变换,即

$$A = \text{fft}_2(I_1) \quad B = \text{fft}_2(I_2) \quad (7)$$

则通过 A、B 的简单的矩阵运算得到另一矩阵 C_3 ,即

$$C_3 = B^* \cdot \text{conj}(A) / \text{norm}(B^* \cdot \text{conj}(A), 1) \quad (8)$$

(3) 由(8)式可知, C_3 的二维傅里叶逆变换

阵 C 在 (x_0, y_0) 处取得最大,可通过遍历比较 $C(i, j)$ 大小即可找到该位置,并作为函数返回值。

4 图像拼接

采用三角函数进行加权融合的拼接效果图,如图 3 所示。

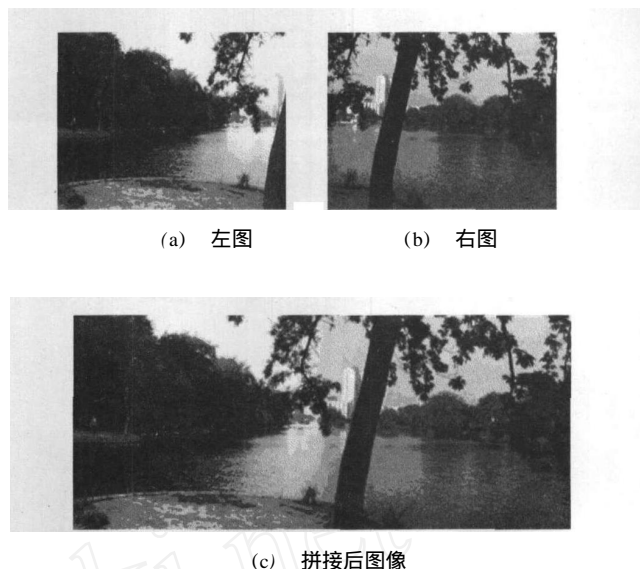


图 3 两幅图像拼接效果图

对于重叠区域的图像,如果仅进行简单地平均加权求和,会在拼接的图像中存在明显的拼痕带,主要原因是相邻图像间存在着亮度的差异和普通相机的拍摄会在采集到的图像中出现边缘失真的现象,而拼接使用的恰恰是图像的边缘区^[6]。本文采用函数加权法来处理重叠区域的图像,确保图像的无缝过渡。

设 $I_1(x, y), I_2(x, y)$ 分别为待融合图像区域 (x, y) 点的像素值,则融合后该点的像素值 $I(x, y)$ 为

$$I(x, y) = W_1(x, y) I_1 + W_2(x, y) I_2 \quad (9)$$

其中, $W_1(x, y), W_2(x, y)$ 分别为重叠像素点对应 I_1, I_2 的加权值。为使拼接后的图像在 2 个方向上都能平滑过渡,达到更好的融合效果^[7],本文中的加权系数 $W_1(x, y), W_2(x, y)$ 采用三角函数。

$$W = A \sin(\pi x / \text{width}) + T \quad (10)$$

其中, $x = i / \text{width}$ 。

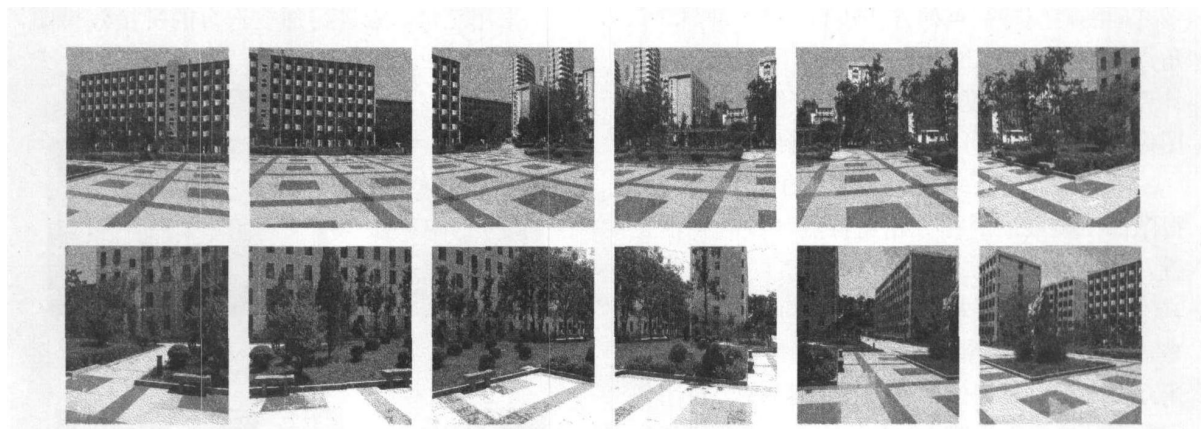
5 结果与分析

本文采用数码相机配三角架对校园内公寓景点进行了定点拍摄,选用 640×480 像素的照片,每隔 30 拍摄一次得到 12 幅序列图像。采用本文算法对序列图像进行了全景图的拼接^[8],如图 4

所示。其中,图 4a 按照每隔 30 拍摄一次得到的 12 幅序列图像,图 4b 采用本文算法得到的全景拼接效果图。

拼接前对拍摄的序列图像进行噪声消除、校

正等预处理工作,有效地改善了图像的频域相关性,并提高了图像的配准精度;在图像重叠区使用平滑拼接,对图像的亮度进行处理以消除拼痕带,无疑也提高了拼接的质量。



(a) 每隔 30 拍摄的 12 幅序列图像



(b) 12 幅图像拼接的全景图

图 4 基于相位和相关图像拼接效果图

参 考 文 献

- [1] Smith S M, Brady J M. SUSAN—a new approach to low level image processing[J]. International Journal of Computer Vision, 1997, 23 (1) : 45 - 78.
- [2] Tomasi C, Kanade T. Detection and tracking of point features[R]. Tech Rep CMU-CS-91-132, Carnegie Mellon University, 1991.
- [3] 李忠新, 张登峰, 茅耀斌, 等. 基于相位相关法的柱面全景图拼接技术[J]. 南京理工大学学报, 2002, (26) 增刊: 43 - 46.
- [4] 张 静, 欧宗瑛, 胡志萍. 基于轮廓相位相关的图像自动拼接[J]. 大连理工大学学报, 2005, 45 (1) : 68 - 74.
- [5] William S H. A subspace identification extension to the phase correlation method[J]. IEEE Transactions on Medical Imaging, 2003, 22 (2) : 277 - 280.
- [6] Chen S E. QuickTime VR—an image-based approach to virtual environment navigation [A]. Proceedings of SIGGRAPH95[C]. 1995 : 29 - 38.
- [7] Dornaika F, Chung R. Image mosaicing under arbitrary camera motion[A]. Asian Conference on Computer Vision, Taipei Taiwan, 2000[C]. 2000 : 20 - 26.
- [8] Hsu S, Harpreet S S, Kumar R. Automated mosaics via topology inference[J]. IEEE Computer Graphics and Applications, 2002, 22 (2) : 44 - 54.

(责任编辑 张秋娟)