

对环境光照鲁棒的全景图拼接

程 兵 郑南宁

(西安交通大学人工智能与机器人研究所, 西安 710049)

摘 要 提出了一种从光照变化的序列图象中拼接全景图的方法, 该方法首先将待拼接图象的重叠部分分解成水平集表示, 并且定义一个形态学距离用于测量水平集之间的相似度, 然后根据这个形态学距离, 对其中一幅图象的每一个水平集都在另一幅图象的水平集中找到对应, 从而得到一个单调转换函数, 用于表示两幅待拼接图象水平集之间的映射, 用这个转换函数调整其中一幅图象的对比度与另一幅图象相对应, 最后, 用基于灰度匹配的方法将两幅图象拼接, 图象两两拼接后经全局误差校正即可得到一幅正确拼接的全景图, 该方法可以广泛应用于基于图象的绘制、图象处理等领域。

关键词 计算机图象处理(520·6040) 全景图 水平集 图象拼接 环境光照 形态学距离

中图法分类号: TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2003)02-0135-05

Ambient Light Independent Panorama Mosaics

CHENG Bing, ZHENG Nan-ning

(Institute of artificial intelligence and robots, Xi'an Jiaotong University, Xi'an 710049)

Abstract This paper presents a novel method for constructing a panorama from sequences of images with light changing. First, the two images, which will be composited, are decomposed into level sets. A morphological distance is defined here to match the level sets in the overlapping parts of the images. Each level set in a part is matched with the level sets in another part, so the difference in contrast is found and a transform function is obtained. The contrast of an image is adjusted to fit the contrast of the other image using the transform function and then the two images are stitched together by minimizing the intensity error between two images. Every two images are composited together using the same way and a global alignment is implemented to eliminate the visible gap or overlap between the first and last images in a sequence. This work has relaxed the constraint, which needs the same ambient light conditions for all images in the sequence, in traditional panorama mosaic and has a wide variety of applications. Finally, the performance of our method is demonstrated with an experiment.

Keywords Computer image processing, Panorama, Level set, Image mosaic, Ambient light, Morphological distance

0 引 言

传统的计算机图形是用一系列三维图元(点、线、面、体素等)先建立物体的三维模型, 然后经过复杂的绘制过程生成输出的图象。近年来, 一种基于图象的图形绘制(IBR)方法被广泛研究^[1~3], 这种方法根据一组采样图象直接生成位置变化后的新视图, 它的计算量取决于输入和输出图象的分辨率, 与场景复杂度无关, 全景图拼接是基于图象的绘制中的

一个重要方法, 它是两维全光函数的特例, 一系列离散点的全景图又构成了五维全光函数^[2]。全景图是一种宽视角图象, 其允许使用者摇镜头观察, 且在观看时, 可拉近周围场景, 已经被广泛地应用在虚拟现实、飞行模拟、计算机游戏和电子商务等中。文献[4]提出了一种通用的使用手持照相机拍摄的图象拼接全景图的方法, 一些全景图的成像和拼接系统已经在商业上得到了广泛的应用, 例如, QuickTime VR^[3]和的 Real VR, 就是使用一个三脚架和 35mm 胶片的照相机拍摄一系列有重叠的图象, 然后将这

基金项目: 国家自然科学基金(60175006, 60024301); 863 项目(2001AA114202)

收稿日期: 2002-04-19; 改回日期: 2002-08-27

些图象拼接在一起,得到一幅全景图。全景图象也可以使用一个单独的相机和一组特殊的镜头或一些特殊的镜子^[6](圆锥形的、球形的、双曲形的和抛物形的等)通过这些镜头或镜子,照相机可以在一幅图象中拍摄下整个场景或者一个很大视角范围的视图。每一幅全景图都对应一个叫做“投影中心”的视点。全景图可以看成是整个场景通过这个视点到一圆柱面或球面的映射。

通常的全景图拼接方法,在图象配准时,都假设两幅图象拍摄时的光照条件没有变化,但是这个假设并不总是成立,在光照条件变化或照相机曝光参数变化时,这样的拼接算法可能会出现较大误差,为克服这个限制,提出了一种可以在环境光照变化的情况下正确拼接全景图的方法。

1 圆柱形全景图的生成

圆柱形全景图虽然在顶部和底部有很大的畸变,但它能满足大部分漫游系统和虚拟现实的需要而且易于创建和使用,因此得到了非常广泛的应用。先用照相机拍摄一系列图象,如果照相机的位置、焦距和视角已知,那么将每一个透视图象变换到圆柱坐标便可以得到一个圆柱形全景图。

全平面透视图的运动模型可以有許多自由度,这里只考虑摇镜头运动。将照相机放置在一个三脚架上,做水平摇镜头运动拍摄一个图象序列,通过调整摇移速度和拍照间隔,使得每两幅图象之间有一定的重叠部分。通过圆柱面透视投影可以将三维世界坐标映射到两维全景平面坐标^[1]。

$$\theta = \tan^{-1}\left(\frac{y}{x}\right), v = \frac{y}{\sqrt{x^2 + z^2}} \quad (1)$$

这里, θ 是摇移角, v 是扫描线。通过式(1),所有拍摄到的图象都被变换到圆柱面上。由于竖直方向上的抖动和光场的畸变,图象的重叠部分不能完全地对齐,因此需要先通过最小化两幅图象上对应像素的亮度误差来估计图象的水平运动和竖直运动增量 $\delta t = (\delta t_x, \delta t_y)$, 然后再将两幅图象对准。

$$E(\delta t) = \sum [I_{k+1}(X'_i + \delta t) - I_k(X_i)]^2 \quad (2)$$

这里, i 取遍所有像素点, I_k 表示第 k 幅图象, $X_i = (x_i, y_i)$ 和 $X'_i = (x'_i, y'_i) = (x_i + t_x, y_i + t_y)$ 是两幅图象中的对应点, $t = (t_x, t_y)$ 是整幅图象的全局运动场。

2 图象的水平集表示

通常用 $u(x)$ 表示在一个很宽波长范围内的光通量,这样就可以得到一幅灰度图象。灰度图象一般表示成定义在二维空间的实值函数 $u(x): \mathbf{R}^2 \rightarrow \mathbf{R}^+$, $x \in \Omega, \mathbf{R}^+$ 为灰度值的有效集,一般取实数的某一正的子集, Ω 为图象的支撑 ($\mathbf{R}^2$ 中的一个有界域)。图象的处理可以表示成图象算子 T 作用于该图象。

要定义图象的水平集,需要先定义函数的上半连续性。

定义 1 一个作用在 $\mathbf{R}^n$ 上的实函数 f 如果满足:对任意点 $x \in \mathbf{R}^n$ 以及任意实数 $a > f(x)$, 必定存在 x 的一个邻域 B_x , 使得对这个邻域中的任意点 $y \in B_x$, 不等式 $f(y) < a$ 总是成立, 那么则称 f 是 $\mathbf{R}^n$ 上的上半连续函数。

一般来说,在图象处理领域,一个函数是上半连续的这一假设是完全可以接受的,因为在感知上无法辨别一个任意函数和一个上半连续函数的区别。特别地,在数字空间中,所有函数均可看作是上半连续函数。

通过对图象灰度的阈值分解可将图象表示成水平集^[7]。

定义 2 如果一幅图象是上半连续的, 那么它在 λ 水平的上水平集定义为

$$\chi_\lambda u = X_\lambda = \{x \in \mathbf{R}^2 | u(x) \geq \lambda\}, \lambda \in \mathbf{R}^+ \quad (3)$$

其中, χ_λ 为上 λ 水平集算子。

同理,下水平集定义为

$$\chi_\lambda u = X_\lambda = \{x \in \mathbf{R}^2 | u(x) < \lambda\}, \lambda \in \mathbf{R}^+ \quad (4)$$

其中, χ_λ 为下 λ 水平集算子。

上、下水平集中的任何一个都可以完整地重构原图象,因此它们都是原图象的完全表示,并且是无冗余的。实际上,可以很容易地从图象的上水平集恢复图象本身

$$u(x) = \sup\{\lambda | x \in X_\lambda\} \quad (5)$$

与下水平集对应的图象重构公式为

$$u(x) = \inf\{\lambda | x \in X_\lambda\} \quad (6)$$

一个函数的水平集有以下 3 个性质

(1) $X_\lambda \subset X_\mu$, 如果 $\lambda > \mu$ 。

(2) $X_\lambda = \bigcap_{\mu < \lambda} X_\mu$ 。

(3) 如果 $(X_\lambda)_{\lambda \in \mathbf{R}}$ 是一组 $\mathbf{R}^n$ 中的子集, 并且满足性质(1)和性质(2), 那么就存在一个实值函数

$$u(x) = \inf\{\mu, x \in X_\mu\} \text{ 满足 } \chi_\lambda u = X_\lambda.$$

将图象转换成基于水平集表示后,图象处理就不再是对图象中具有灰度值的像素点直接处理,而是转化成图1的形式.

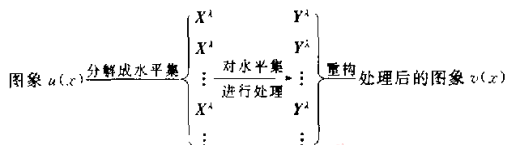


图1 图象通过水平集处理的过程

原图象 $u(x)$ 首先通过式(1)分解成水平集 X^k , 水平集是一组二值图象, 每一个水平集由若干连通分支组成, 用形态学算子对连通分支进行变换或者用曲线进化处理连通分支边界所组成的 Jordan 曲线, 从而得到新的水平集 Y^k , 然后用 Y^k 重构得到处理后的图象 $v(x)$. 这样就把对图象本身的处理转化成对图象水平集的处理, 这种方法具有保形、保对比度的特点^[7].

3 对光照变化图象的拼接

通常的全景图拼接方法都假定光照条件不变, 在一幅图象中某一像素点的亮度与另一幅图象中对应像素点的亮度要相等或几乎相等. 但是在有些情况下, 这个假设并不一定成立. 由于每一个像素点的亮度值取决于传感器属性、光照条件和物体的表面反射属性等条件, 而这些条件又是未知的, 因此对应点的亮度值并不一定总是相等. 如果这些条件有一个较大的变化, 那么使用式(1)和式(2)进行拼接就可能出现较大的误差. 为此, 提出了一种新的方法来克服这个缺陷. 将照相机安放在一个三脚架上作水平摇镜头运动, 每旋转一定角度拍摄一幅照片, 并使得每两幅照片有一定范围的重叠.

以两幅图象 u 和 v 的拼接为例, $\hat{u}$ 和 $\hat{v}$ 分别是图象 u 和 v 中的重叠部分. 对 $\hat{u}$ 和 $\hat{v}$ 用式(3)进行水平集分解. 水平集表示的一个显著特性是它具有全局保对比度不变性, 使用一个单调函数 g 作用于 u 得到 u' , u 和 u' 在全局上看, 水平集是一致的, 用公式描述为: 对于每一个实数 α' 必然存在另一个实数 β 使得 $\chi^\alpha u = \chi^{\alpha'} u'$ 成立, 反之亦然.

下一步的任务就是要寻找两个图象水平集表示之间的变换函数 g , 通过这个变换函数可以调整其中一幅图象的像素值, 使其与另一幅图象中对应点

的像素值相等或接近. 这里需要一个形态学距离, 用它来测量水平集之间的差异. Hausdorff 距离可以用来测量两个形状之间的差别, 但是它对噪声比较敏感, 尤其是对正的脉冲噪声和负的脉冲噪声不对称, 因为没有理由说明正脉冲和负脉冲对精度的影响不一致, 所以在实验中, 若使用 Hausdorff 距离, 则用该算法得到的结果就不是很稳定. 为此定义一个新的形态学距离测度.

定义3 给定一个集合 A 和一个元素 B , 集合 A 对应于 B 的尺寸为

$$M_B(A) = \sup\{\lambda | A \cap \lambda B \neq \emptyset\}$$

这里 $A \cap B$ 指集合 A 和集合 B 的开运算.

定义4 根据对集合尺寸的定义可以得到两个集合 X 和 Y 之间的形态学距离为

$$d(X, Y) = \max\{M_D(X - Y), M_D(Y - X)\}$$

这里 D 是预定义的一个小的形状元素, 可取半径为 1 的圆盘, 也可以根据图象的一些先验知识, 如纹理和形状等取为矩形、三角形和菱形等元素. 这里定义的测度不满足通常意义下距离测度应满足的三角不等式, 即: 对任意集合 Z 满足

$$d(X, Y) \leq d(X, Z) + d(Z, Y)$$

但是这并不影响它在形状匹配中的应用, 而且这个测度对噪声不敏感, 对正负脉冲的影响是一致的. 在每一个水平集上, 可以通过最小化两个图象水平集之间的误差 $E(\tau) = d(X_{\lambda+\tau}, X_{\lambda-\tau})$ 来得到一个增量 τ , 从而最终得到转换函数 g . 这里可以根据光照变化的大小来确定 τ 的搜索范围, 以减少运算量. 函数 g 得到后, 用它来调整图象 u 的对比度, 得到一幅新的图象 u' , 最后用图象 u' 和图象 v 进行拼接得到一幅更大的图象, 所有的图象两两拼接完成后, 在最后一幅图象和第 1 幅图象拼接时会有较大的累计误差, 需要采用全局误差校正^[8], 最终得到一幅完整的全景图.

4 实验结果

实验中, 通过调整照相机的曝光参数来模拟光照的变化. 图2显示了在不同光照条件下(即不同的曝光参数下)拍摄的两幅图象, 很明显, 图2(a)比图2(b)要亮一些. 通过两幅图象中重叠部分的水平集分解, 然后计算得到转换函数 g . 用转换函数 g 调整图2(a)的对比度, 得到一幅新的图象(图3(a)), 图3(a)的像素点的灰度值与图2(b)中对应点的灰

度值相等或相近,用式(1)和式(2)将图3(a)和图2(b)拼接起来得到一幅更大的图象(图3(b)),其余图象也用类似的方法两两拼接,然后经过全局

调整即得到一幅拼接完整的全景图(图4),为了显示方便,横向做了50%的压缩。



(a)



(b)

图2 不同光照条件下的两幅图象



(a)



(b)

图3 经光照调整后的图象(a)及拼接后的图象(b)

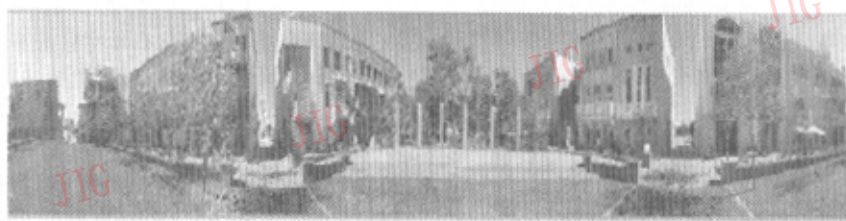


图4 拼接得到的全景图(水平方向上做了50%的压缩)

5 结 论

本文提出了一种简单实用,对环境光照变化鲁棒的全景图拼接方法,它可以广泛地应用于基于图象的绘制、图象处理等领域,这里只考虑了环境光变化引起的图象整体亮度的改变,如果场景中有大的局部光照的变化,那么此方法失效,所以需要引进对水平集局部连通分支的调整,这也将是下一步的研究工作。

参 考 文 献

- 1 Chai J X, Tong X, Chan S C *et al*. Plenoptic sampling[A]. In: Proceedings of ACM SIGGRAPH2000[C], New Orleans, Louisiana, USA, July 2000.
- 2 McMullan L, Bishop G. Plenoptic modeling: An image-based

rendering system[A]. In: Proceedings of ACM SIGGRAPH95[C], Los Angeles, CA, USA, August 1995,39~46.

- 3 Chen S E. QuickTime VR—an image based approach to virtual environment navigation [A]. In: Proceedings of ACM SIGGRAPH95[C], Los Angeles, CA, USA, August 1995,29~38.
- 4 Szeliski R. Video mosaics for virtual environments[J]. IEEE Computer Graphics and Applications, 1996,16(2):22~30.
- 5 Nayar S. Catadioptric omnidirectional cameras [A]. In: Proceedings of IEEE CVPR97[C], San Juan, Puerto Rico, June 1997,482~488.
- 6 Szeliski R, Shum H Y. Creating full view panoramic image mosaics and environment maps [A]. In: Proceedings of SIGGRAPH[C], Los Angeles, CA, USA, August 1997,251~258.
- 7 Guichard F, Morel J M. Partial differential equations and image iterative filtering[A]. In: Tutorial Int. Conf. Image Processing [C], Washington, DC, 1995.

8. Shum H Y, Szaliski K. Panoramic image mosaics[R]. Microsoft Research Technical Report, MSR-TR-97-23, 1997.



程兵 1971年生,硕士,西安交通大学人工智能与机器人研究所博士研究生。目前主要研究方向为计算机视觉与图形学。



郑南宁 1952年生,博士,教授,中国工程院院士。主要研究方向为计算机视觉、并行与分布式计算、神经网络动力学及基于视觉信息的智能检测与识别系统。

2003 全国数字媒体与数字城市学术会议

征 文 通 知

2003 年全国数字媒体与数字城市学术会议定于 2003 年 5 月 24~26 日在武汉大学召开。由中国地理信息系统协会、中国图象图形学会虚拟现实专委会及中国测绘学会摄影测量与遥感专委会主办,武汉大学城建学院、遥感信息工程学院、测绘遥感信息工程国家重点实验室联合承办。会议将以“图象图形、3S 及各种工程数字化设计的理论、应用及发展”,开发广泛的交流和讨论。会议将特邀国内外著名学者和专家就此领域的最新动态和热点问题作专题讲演和报告,欢迎从事科研、教学、工程、开发及推广应用的广大地理信息、图象图形工作者积极投稿。

一、征文范围

影像处理	智能虚拟环境	GPS 理论及应用	数字水利
影像分析	图象图形理论及应用	数字化城市规划	数字电力
图象图形识别与理解	多媒体数据库及多媒体信息	数字化建筑设计	人工智能及专家系统
计算机仿真	遥感图象理论及应用	城市规划与建筑设计的新技术	互联网及网络技术
计算机可视化	GIS 理论及应用	数字管网	数据管理技术
虚拟现实及虚拟技术			

二、征文要求及有关事项

- 稿件应反映图象图形、3S 及各工程数字化设计的理论、技术或应用成果,并未在其他会议及国内外公开刊物上发表过。
- 经会议程序委员会评审录用的论文将在《武汉大学学报》(信息版 No. 3)和《工科版 No. 3》及《测绘信息与工程》No. 3、No. 4 等核心期刊上正式发表,部分文章将进入 EI 检索。
- 文稿请按 Word 格式排版,接受电子投稿和正式投稿(附软盘)。
- 重要日期:
征文截止日期:2003 年 2 月 20 日;录用通知日期:2003 年 3 月 1 日
- 来稿请寄:
(1) tshang@public.wh.hb.cn(电子版)
(2) 武汉市东湖南路 6 号 武汉大学城建学院 尚涛,孙宇宁,詹平(打印文稿,附软盘,一式三份,注明:2003 全国数字媒体与数字城市学术会议)

三、产品展示

会议期间将举办 3S 及图象图形产品展示会、新产品发布会和专题报告会,欢迎有关厂家、公司报名参展,请参展单位来函来人联系。

参展报名截止日期:2003 年 5 月 1 日

联系人:尚涛 武汉市东湖南路 8 号武汉大学城市建设学院

电话:(027)67803312,67802147 传真:(027)67803062,67802147

邮编:430072 E-mail: tshang@public.wh.hb.cn