

一种图像自动拼接的快速算法

侯舒维, 郭宝龙

(西安电子科技大学机电工程学院, 西安 710071)

摘 要: 针对现有基于灰度级相似的图像拼接方法的缺点, 提出了一种图像自动拼接的快速算法。该算法综合考虑了图像拼接的精度和速度, 在基准特征块的提取上, 采用简单的边缘信息阈值法, 实现了基准块的自主选取, 提高了图像拼接的精度; 在块搜索上, 采用金字塔式分层搜索策略, 提高了图像拼接的速度。实验证明, 该算法扩展了传统拼接算法的适用范围, 具有较好的性能。

关键词: 图像拼接; 配准; 特征块; 图像融合

A Fast Automatic Image Stitching Algorithm

HOU Shuwei, GUO Baolong

(School of Electromechanical Engineering, Xidian University, Xi'an 710071)

【Abstract】 By analyzing the available disadvantages of the intensity-based image stitching, a fast automatic image stitching algorithm is proposed. This algorithm takes into account the precision of image stitching as well as the speed of it, as to the pick-up of the base feature block, a simple method of edge information threshold is used, which realizes the automatic pick-up of base feature block, therefore improves the precision of image stitching. In the problem of searching matching block, a hierarchical search strategy is employed, which improves the speed of image stitching. The experiments show that the proposed algorithm greatly extends the scope of traditional image stitching algorithm and has good capabilities.

【Key words】 Image stitching; Registration; Feature block; Image composition

给定某个场景的一组互相有重叠的局部图像, 如何生成包含着这组局部图像的新的较大的视图, 称为图像的拼接(stitch)问题, 有时也称为图像镶嵌(mosaic)。图像拼接技术的研究是在现实应用的需求下逐渐发展起来的, 现已成为计算机视觉和计算机图形学的研究焦点。图像拼接技术广泛应用于数字视频、MPEG-4 编码、运动分析、虚拟现实技术、遥感图像处理、医学图像分析等领域。

图像拼接技术的关键和核心在于准确找到相邻图像间重叠区域的位置以及范围, 也就是图像配准问题。通常采用的方法是基于两幅图像重叠部分所对应 RGB 或 CMY 颜色系统中灰度级的相似性, 自动寻找图像的匹配位置。常用的配准算法有比值匹配法^[2], 它是从一幅图像的重叠区域中部分相邻的两列上取出部分像素, 然后以它们的比值作模板, 在另一幅图像中搜索最佳匹配。这种算法计算量较小, 但精度较低; 另外就是块匹配法^[3], 它是从一幅图像重叠区域中的一块作为模板, 在另一幅图像中搜索与此模板最相似的匹配块, 这种算法精度较高, 但计算量过大; 为了减小块匹配法的计算量, 后来出现了网格匹配法^[4], 它首先要进行粗匹配, 每次水平或垂直移动一个步长, 记录最佳匹配位置, 然后在此位置附近进行精确匹配, 每次步长减半, 然后循环此过程直至步长减为 0。这种算法较前两种运算量都有所减小, 但在实际应用中仍然偏大, 而且粗匹配时如果步长取得太大, 很可能会造成较大的粗匹配误差, 从而很难实现精确匹配。

本文采用的新算法就是在块匹配法的基础上, 找到一种精度既高运算量又较小的方法。这里的基于特征块匹配的图像拼接新算法, 就是在两幅待拼接图像中任选一幅图像, 选取其重叠区域中的一个特征块作为基准模板, 然后在第 2 幅图像中搜索与此基准模板相匹配的块, 从而确定两幅图像的

位移。在特征块的选取上, 采用简单的边缘信息阈值法, 实现了特征块的自主选取, 从而提高了匹配的精度; 在特征块的匹配上, 采用金字塔式分层搜索策略, 从而大大减少了运算量。实验证明, 此算法抗干扰能力较强, 速度较快, 拼接效果也较为理想, 具有一定的实用价值。

1 基准特征块的选取

进行特征块匹配的关键且首要的一步就是选取基准特征块, 而现有的很多模板法图像拼接技术都较少兼顾考虑到这一点, 从而造成拼接算法不能适应于待拼接图像的变化, 甚至有些图像会因此而导致拼接的失败。针对这一点, 本文提出了边缘信息阈值法进行基准特征块的选取, 下面具体介绍它的实现方法。

我们所要找的基准特征块应该含有图像的一些主体信息。从人的视觉方面来分析, 这个特征块应该是人眼较为集中的部分。而图像的边缘信息是图像的主要特征之一, 能引起人视觉集中的地方往往是边缘分布较密的地方, 因此以图像的边缘为特征寻找基准块。

首先需要对图像进行边缘检测, 这里使用 sobel 算子获得原图像的边缘图 $E(i, j)$, 然后以图像中每个像素点的邻域边缘量 (NEA) 来定义此位置的边缘信息大小:

$$NEA(i, j) = \sum_{m=-A}^A \sum_{n=-A}^A E(i+m, j+n) \quad (1)$$

其中, $E(i, j)$ 为原图像所对应的二值边缘图, A 为在点 (i, j)

基金项目: 国家自然科学基金资助项目 (69975015); 教育部优秀青年教师资助计划项目

作者简介: 侯舒维(1980—), 女, 硕士生, 主研方向: 图像处理, 计算机视觉和模式识别等; 郭宝龙, 博士、教授、博导

定稿日期: 2004-06-26 **E-mail:** hsw521@sina.com

处所取邻域的 1/2 边长。

以 $NEA(i, j)$ 值为依据, 找出最大值所对应的点 (i, j) 作为基准点, 然后以此点为中心点选择大小适中的块就可以找到基准特征块。一般来说, 图像的主体边缘分布是呈区域分布的, 如果在某点的邻域内它的边缘量很大, 则此位置所在的部分区域边缘都较为集中, 由于使用点的邻域边缘量来衡量此点所在位置的边缘信息, 因此一些孤立的边缘点可以通过这一步得到排除。

虽然按此原理可以找到图像的特征块, 但是在后续的匹配过程中, 可能会出现有多重匹配的情况, 这就是因为还不能保证所找的特征块是图像中独一无二的特征块, 为了避免这种情况发生, 再作如下的定义:

以每一点的邻域边缘量与整幅图像的平均边缘量的差值为依据来衡量以此点为中心的区域与其它区域的差异:

$$D(i, j) = |NEA(i, j) - evend| \quad (2)$$

其中 $evend$ 为整幅图像的平均边缘量。找出 D 值最大的点 (i, j) 即为所求的基准点。但是如果仅以式(2)为标准选择基准点, 很可能会选到邻域边缘量很小的点, 甚至有可能会选到孤立边缘点, 而这不是特征块应该具有的性质, 因此为了结合式(1)和式(2), 采用一种简单的边缘信息阈值法进行处理。

首先, 由式(1)计算出图像中每一个点的邻域边缘量, 然后取一阈值先去除掉邻域边缘量 (NEA) 较小的点, 这样做的好处有: (1) 在去除掉这些 NEA 值较小的点后, 为后续的选点大大减少了计算量; (2) 去除掉这些 NEA 值较小的点后, 在剩余的点中求平均值 $evend$, 这样可以减少图像中 NEA 值较小的点对主体 $evend$ 值的影响; (3) 去除掉这些 NEA 值较小的点后, 就可以先保证最后选到的点 NEA 值较大 (至少大于阈值), 因为对于一般的图像, 出现多重匹配的机率总是有限的, 我们最关心的是希望能够找到一个边缘信息丰富的特征块, 所以保证 NEA 值较大是首要的。其次, 在剩余的点中求邻域边缘量的平均值 $evend$, 然后按照式 (3) 找出 D 值最大的点即为最后所求的基准点 (i, j)

$$D(i, j) = i \times j \times |NEA(i, j) - evend| \quad (3)$$

(i, j) 为图像中像素点的坐标, 以图像的左上角为坐标原点。引入坐标值的目的是为了保证最后所选的点尽量位于图像的中心, 因为靠近图像边界的点对最后所选特征块的大小有所限制, 这样就很难实现基准特征块大小的自主选取。

对于阈值的选取, 第一要保证最后所选的点是图像中 NEA 值较大的点, 即阈值不能太小; 第二要保证有足够的剩余点可供选择, 否则点数太少会造成在后续的选点中陷入局部最优, 即阈值也不能太大。在本文的实验中, 取 $maxd/2$ 作为阈值, 其中 $maxd$ 是图像中邻域边缘量 (即 NEA) 的最大值。

确定了特征块的基准点之后, 以此点为中心, 取一大小适中的块就可满足要求, 特征块的大小采用自适应的方法进行自主调整, 它的计算如下:

$$r = r_0 \times \frac{evend}{\bar{n}} \quad c = c_0 \times \frac{evend}{\bar{n}} \quad (4)$$

其中, r 和 c 分别是特征块沿 x 方向和沿 y 方向的长度, r_0 和 c_0 是特征块的初始长度, 这里取

$$r_0 = rinput / 8, c_0 = cinput / 8$$

\bar{n} 是初始特征块的块内平均 NEA , 如图 1 所示。图中, $input$ 为待拼接图像, $rininput$ 为图像沿 x 方向的长度, $cininput$

为图像沿 y 方向的长度, w (图中阴影部分) 为所选的基准特征块。

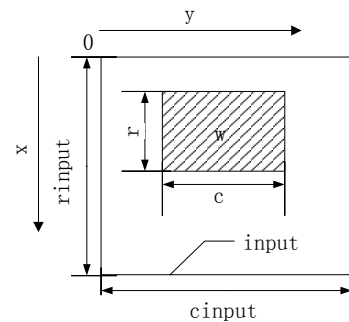


图 1 特征块选取示意图

2 图像配准

图像配准就是确定待拼接两幅图像的重叠区域以及重叠位置, 以便最终完成图像的拼接, 它是整个图像拼接算法中的核心部分。传统的配准算法由于在原图上直接进行块匹配, 计算量大, 因此无法满足实际需求。为提高拼接速度, 采用的方法是由粗到细的分层块匹配法。

2.1 配准则则

本文采用传统的 Richard Szeliski 方法, 以两幅图像重叠部分 (这里是指重叠块) 的灰度差的平方和为标准来衡量此区域是否配准, 设待拼接的两幅图像为 $I_1(x, y)$ 和 $I_2(x, y)$, 配准公式如下:

$$E = \sum_x \sum_y [s(x, y) - w(x, y)]^2 \quad (5)$$

其中, $s(x, y)$ 是图像 $I_1(x, y)$ 中的搜索块, $w(x, y)$ 是图像 $I_2(x, y)$ 中的基准特征块, 求和是指当 $w(x, y)$ 在 $I_1(x, y)$ 中滑动时, $w(x, y)$ 覆盖下的 $I_1(x, y)$ 内对应块 $s(x, y)$ 与 $w(x, y)$ 差值的叠加。最佳匹配块为 E 取最小值时所对应的 $I_1(x, y)$ 中的搜索块。将式 (5) 展开, 可得到下式:

$$E = \sum_x \sum_y [(s^2(x, y) - 2 * s(x, y) * w(x, y) + w^2(x, y))] \quad (6)$$

其中, 括号内的第一项是待拼接图像 $I_1(x, y)$ 的块能量, 它的值与 w 无关, 最后一项是基准特征块的能量, 当 w 选定后就是个常数, 而中间一项可以看成是衡量 $s(x, y)$ 与 $w(x, y)$ 的相似度的量, 由不等式的性质可以知道, 当 $s(x, y)$ 与 $w(x, y)$ 相同即完全匹配时, 此项的绝对值取最大, 因此可以将配准则则改为下式:

$$R(u, v) = \sum_x \sum_y s(x, y) * w(x - u, y - v) \quad (7)$$

式中, $R(u, v)$ 为 $s(x, y)$ 与 $w(x, y)$ 的相关系数, 在图像 $I_1(x, y)$ 中的每一个位置 (u, v) 处计算对应重叠区与 w 的卷积, 最大值所对应的位置 (u_0, v_0) 即为最佳匹配位置。式 (7) 有一个缺点, 它受 $s(x, y)$ 幅值和 $w(x, y)$ 幅值的影响较大, 因此再作以下的归一化处理:

$$R(u, v) = \frac{\sum_x \sum_y [(s(x, y) - \bar{s}) * (w(x - u, y - v) - \bar{w})]}{\left\{ \sum_x \sum_y (s(x, y) - \bar{s})^2 * \sum_x \sum_y (w(x - u, y - v) - \bar{w})^2 \right\}^{\frac{1}{2}}} \quad (8)$$

其中, \bar{s} 为图像 $I_1(x, y)$ 中对应搜索块 $s(x, y)$ 的灰度平均值, \bar{w} 为基准特征块的灰度平均值, 由柯西不等式可以知道, $R(u, v) \in [-1, 1]$ 。

2.2 搜索策略

本文采用金字塔式分层搜索策略,由粗到细,由上到下逐步搜索。首先要分别对两幅待拼接图像进行粗提取,即依次对每幅图像中 2×2 大小的块内像素灰度求平均得到一个粗像素点,假设对图像 I 和图像 J 分别做 n 次粗提取,则每一个图像都得到 n 个粗提取图像,构成如图 2 所示的金字塔分层结构,若原图像的大小为 $M \times N$,则第 n 层粗提取图像的大小为 $\frac{M}{2^n} \times \frac{N}{2^n}$,即减小为原图的 $\frac{1}{2^{2n}}$;然后进行分层搜索,搜索时从第 n 层开始进行,直到第 0 层(原图像)结束。为了避免误匹配,又兼顾考虑到运算量,在第 n 层的搜索中,要确定一些粗匹配点,转到下一层搜索时,只在这些粗匹配点中进行精细匹配,而最高层即使采用全局搜索,搜索点数也仅为原图像的 $\frac{1}{2^{2n}}$,因此通过这种方法,就可以大大减小搜索点数,提高匹配速度。

层数越高,所选择的粗匹配点应该越多,但图像所分的层数并不是越多越好,如果分层太多,不仅增加了计算量,而且容易造成误匹配,一般选 2~3 层效果较好。

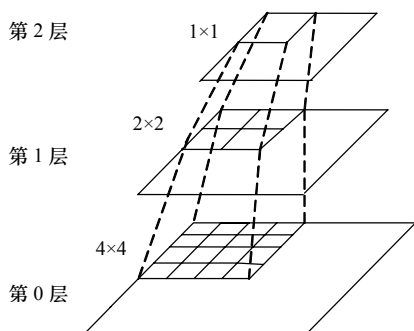


图2 金字塔分层结构

3 图像融合

由于在重叠区域直接取两幅图像中其中的一幅进行图像合成,很容易在重叠区域的边界处形成明显的拼缝,因此,需要采用图像融合技术合成全景拼接图。

本文采用渐入渐出法,将图像重叠区域的像素值按一渐变系数相加合成全景图像,以左右拼接的两幅图像 $I_1(i, j)$ 和 $I_2(i, j)$ 为例,则重叠区域图像的像素值 $I(i, j)$ 可表示为

$$I(i, j) = dI_1(i, j) + (1 - d)I_2(i, j)$$

其中, d 为一渐变系数,由上式可以看出,当 d 由 1 慢慢变化到 0 时,图像从 $I_1(i, j)$ 慢慢过渡到了 $I_2(i, j)$,这样就可以实现图像间的平滑过渡,从而消除了拼接的痕迹。

4 实验结果

为减小计算量,避免全局选取,在选择基准特征块时,预先限定一个区域。对于一般图像,在原图大小的 $50\% \pm 20\%$ 范围内选择,具体情况按重叠区域大小而定。我们用 CCD 采集图像对本文的算法做了拼接实验,相邻图像间有 $30\% \sim 50\%$ 的重叠区域,且保持焦距不变,表 1 给出了 65 对图像拼接的实验结果。

表1 实验结果

实验图像 (对)	拼接成功的图像 (对)	拼接失败的图像 (对)	成功率
65	64	1	98.46%

图 3 给出连续 4 幅图像的室内场景拼接效果图,图像均为 CCD 采集图像,相邻图像间约有 50% 的重叠区域。

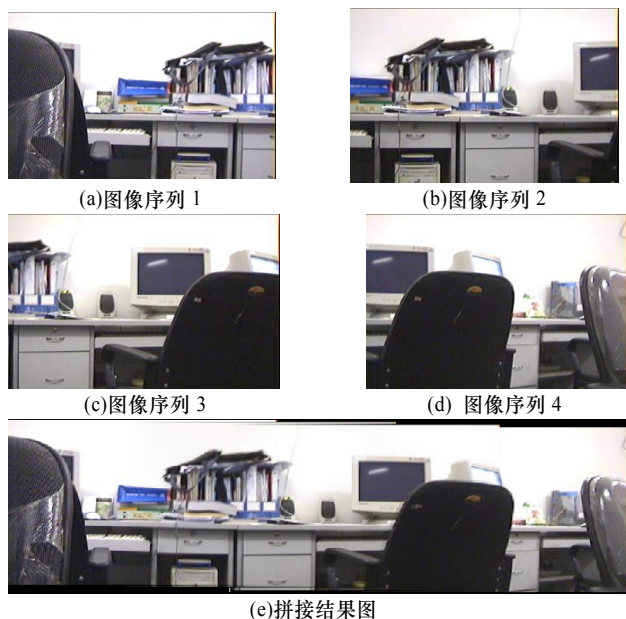


图3 室内场景拼接效果图

图 4 给出连续 3 幅图像的室外场景拼接效果图,相邻图像间约有 30% 的重叠区域。

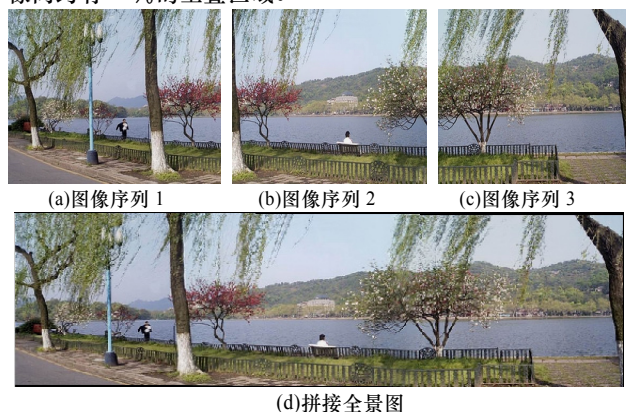


图4 室外场景拼接效果图

可以看出,本文的算法能够实现多幅图像的无缝拼接,效果令人满意。

5 结论

图像拼接是一个日益流行的研究领域,它可以提供具有真实感的全景图,是虚拟现实场景创建和虚拟漫游的基础。本文的算法采用自主选择基准特征块的思想,实现了图像的精确拼接,同时采用金字塔式分层匹配的思想,进而实现了图像的快速拼接。因此,本文的算法既具有良好的鲁棒性又具有较好的实用性。

参考文献

- 1 Szeliski R. Video Mosaics for Virtual Environments. IEEE Computer Graphics and Applications (C). IEEE, 1996, 16(2): 22-30
- 2 钟力, 胡晓峰. 重叠图像拼接算法[J]. 中国图像图形学报, 1998, (5): 365-369
- 3 汪成为. 灵境(虚拟现实)技术的理论、实现及应用[M]. 南宁: 广西科学技术出版社, 1996
- 4 李志刚. 边界重叠图像的一种快速拼接算法[J]. 计算机工程, 2000, 26(5): 37-38
- 5 陈果, 左洪福. 显微图像的一种快速拼接算法. 仪器仪表学报, 2003, 24(4): 343-347
- 6 章毓晋. 图像处理和分析[M]. 北京: 清华大学出版社, 1999