

图像拼接技术

王俊杰 刘家茂 胡运发 于 玉

(复旦大学计算机系 上海200433)

Research and Development of Image Mosaics

WANG Jun-Jie LIU Jia-Mao HU Yun-Fa YU Yu

(Department of Computer and Information Technology, Fudan University, Shanghai 200433)

Abstract Image mosaics have been an active area of research in the fields of computer vision, image processing and computer graphics in recent years. The automatic fast construction of unlimited field of view, high-resolution image mosaics is a main research task of this area. According to the procedure of image mosaics, the paper introduces and discusses image acquisition, geometric corrections, image register and image blending in detail. In the last part of the paper, we make a discussion on some problems of research and point out the future research directions.

Keywords Image mosaics, Image acquisition, Geometric corrections, Image registration, Panoramic image mosaics, Image blending

1. 引言

图像拼接(Image Mosaics)是一个日益流行的研究领域,它已经成为照相绘图学、计算机视觉、图像处理和计算机图形学研究中的热点。图像拼接解决的问题一般是,通过对齐一系列空间重叠的图像,构造一个无缝的、高清晰的图像,它具有比单个图像更高的分辨率和更大的视野。

早期的图像拼接研究一直用于照相绘图学,主要是对大量航拍或卫星的图像的整合。近年来随着图像拼接技术的研究和发展,它使基于图像的绘制(IBR)成为结合两个互补领域——计算机视觉和计算机图形学的研究焦点。在计算机视觉领域中,图像拼接成为对可视化场景描述(Visual Scene Representations)的主要研究方法^[1];在计算机图形学中,现实世界的图像过去一直用于环境贴图,即合成静态的背景和增加合成物体真实感的贴图。图像拼接可以使 IBR 从一系列真实图像中快速绘制具有真实感的新视图^[2]。

目前,一个特别流行图像拼接技术的应用是全景图的拼接,它是基于图像绘制虚拟现实场景创建和虚拟漫游的基础。由于图像是独立拍摄的,在光滑表面上(如圆柱面^[3-5]和球面^[6-9])进行图像拼接对清晰度没有约束,并且可以避免不连续现象。全景图提供一种在虚拟场景交互式浏览中良好的感觉,使用节点合成多个场景可以让用户在场景之间切换漫游^[3,8],利用计算机视觉的方法,能够从两个节点之间产生新的中间视点图像^[4]。与几何模型绘制真实场景相反,可以从这些节点重建场景的3D几何模型^[10]。

图像拼接的其它一些应用有图像的稳定和变化检测^[11],图像分辨率增强^[12],视频处理^[13],包括视频压缩^[14]和视频索引^[15]。

图像拼接过程一般分为四个步骤:

(1)拍摄获取图像序列。希望拼接的结果决定了图像获取方式。

(2)镜头引起几何变形的校正。利用图像序列数据或照相机模型进行校正。

(3)对几何校正后的图像进行对齐。

(4)图像的融合,消除对齐图像的不连续和缝隙。

很多情况下,步骤(2)(3)是结合在一起进行的。通常的图像拼接的技术也适用于全景图像拼接,因此本文将一同讨论全景图像拼接。

本文首先介绍了图像获取(拍摄)的方法,接着讨论了图像几何变形的校正,然后对典型的图像对齐模型和对齐后的图像合成进行了详细的论述和比较分析。最后指出了研究中存在的问题,以及今后的研究方向。

2. 图像获取(Image Acquisition)

图像获取方法的不同导致取得输入图像的不同,最终拼接结果也不同。图像获取由照相机拍摄时运动状态决定。一般有三种情况^[16]:(1)照相机固定在三角架上,旋转照相机拍摄。(2)照相机放置于一个滑轨上,平行移动照相机进行拍摄。(3)是一种普通的情况,人手持照相机,站在原地拍摄四周,或者沿着照相机的光轴垂直方向走动拍摄。

2.1 旋转照相机拍摄

在这种情况下,放置照相机的三脚架在拍摄过程中一直在同一位置。拍摄时,照相机绕垂直轴旋转,每旋转一定的角度,拍摄一张照片。理想的情况下,照相机不绕照相机光轴旋转。

拍摄得到一系列照片中相邻两张必须有部分重叠。重叠区域大小是图像拼接最重要的影响因素。文[3]建议相邻图像之间重叠比例达到50%。重叠比例越大,拼接就越容易,但是需要的照片越多。

旋转照相机拍摄由于照相机固定,不需要恢复过多参数,较容易实现。但是,拍摄图像不在一个平面上,需要投影到同一个平面上,这将会导致图像质量下降。一个解决方法是使用

王俊杰 硕士研究生,研究方向:图像绘制,计算机视觉。刘家茂 硕士研究生,研究方向:图像处理。胡运发 博士生导师,研究领域:人工智能,计算机视觉。于 玉 教授。

短焦距,即广角镜头。

2.2 平移照相机拍摄

平移照相机指的是照相机在一个平行于成像平面的方向上平移。在固定焦距的情况下,照相机放置在一个滑轨上移动拍摄。物体和照相机的距离远近,或者拍摄物体的大小的变化,都会影响到最后的拼接结果。

这种情况的缺点:拍摄的相片在一个平面上,全景图的三维感觉不如旋转拍摄的。

2.3 手持照相机拍摄

这种方法比较容易做到,手持照相机原地旋转拍摄,或按一定的路线平行于对象拍摄。但是,拼接手持照相机拍摄的照片是很困难的,因为在拍摄过程中,照相机的运动非常复杂。原地旋转拍摄类似于固定照相机旋转拍摄,但是角度控制,旋转控制都很差。沿一定路线移动时,类似于平移拍摄,控制距离和保持相同的成像平面很困难。为了减少这些影响,可以增加重叠比例,使照相机旋转角度、平移减小,因而减小相邻图像之间的不连续程度。

2.4 存在的问题

最常见的问题就是相邻图像之间光强的变化较大。理想情况下,相同的区域应该有相同的光强,但是因为光源变化或者照相机运动和光源平角的变化,导致光强的差异。

另外一个和光条件相关的问题是反光区域,例如,镜子和闪亮金属。高亮光将会降低相应区域的对比度。

场景中物体移动和拍摄时透镜引起的图像变形也将给对齐拼接带来困难。

上述的方法一般限制了照相机的运动,但是实现中拍摄的图像存在小视差,不同比例的缩放和大角度旋转,这些都增加了对齐拼接的难度。因此很多文献^[4,5,17]要求照相机以最小运动视差旋转拍摄。

3 几何校正(Geometric Corrections)

几何变换通过重新映射,改变图像像素点之间空间关系。几何变换包括全局和局部的变换^[18]。全局的变换通常有一个等式定义可以对整幅图像进行变换。局部变换只用于图像的部分,难以精确表达。几何变换通常用一些参数控制整个图像像素的运动,大量的像素信息对参数的估计很有效。

3.1 照相机运动投影模型

照相机运动投影模型包括2D的仿射变换和透视变换。仿射变换可以表示为:

$$f(x)=Ax+b$$

其中 A 是变换矩阵, b 是平移矢量。矩阵 A 控制仿射的缩放、旋转和剪切效果。

透视变换是中心投影的射影变换,用非齐次射影坐标表达时,透视变换用8参数单应矩阵(homography) A 表示,改变矩阵 A 进行透视校正,透视变换也表示为平面的分式线性变换:

$$\begin{bmatrix} X' \\ Y' \\ W \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$x' = \frac{m_{11}x + m_{12}y + m_{13}}{m_{31}x + m_{32}y + 1}$$

$$y' = \frac{m_{21}x + m_{22}y + m_{23}}{m_{31}x + m_{32}y + 1}$$

其中 W 为缩放参数。变换后的新图像坐标: $x' = X'/W$, $y' = Y'/W$ 。

仅当图像之间没有运动视差时,相应图像的仿射变换才对所有像素有效。运动参数可以通过迭代求解^[5,6,12,19]。

8参数单应矩阵能够精确描述照相机沿固定点旋转拍摄的不同视图间透视变换。具有代表性的透视变换有:投影变换,立体平面(Stereographic)变换和等距投影变换。这些变换有各自的特点:投影变换保持了线的形状,而立体平面变换保持了圆形状的特征。立体平面变换可以映射球面全部区域到一个投影平面上。等距投影变换可以看作球面的展开,它映射视图全部区域而不再是近似的情形。

3.2 柱面、球面投影

照相机绕固定点旋转拍摄获取图像序列的描述和合成一般用柱面和球面投影表示。在柱面、球面上合成的360度的图像称为全景图。全景函数^[4](plenoptic function)描述了一个已知场景可能环境集合,全景图是从固定观察位置全景函数的一个采样。

描述全景图通常采用柱面映射^[3~5],不考虑底面和顶面,在柱面坐标系统中用一致的采样合成全景图。球面映射^[6~8]也作为描述全景图的环境。但球面全景图在极点处的异常使其构造显得复杂,极点附近的数值误差会导致自动拼接时评价误差。在极点区域用鱼眼透镜获取图像和相对小尺寸图像会减轻上述异常的出现。在文[6,8]中,相邻图像间在映射到球面前的相对旋转运动可以避免图像对齐时的异常。

3.3 照相机条扫描运动

平面照相机在平移运动情况下,拍摄存在的视差会使几何校正变得复杂,一维照相机(如“pushbroom camera”)扫描式的场景拍摄能够避免这种情况。这种方法能够用传统的照相机模拟,把二维图像序列中扫描条合并为一系列相邻区域。一维照相机可以旋转得到柱面映射图像或者平移得到正交映射图像^[20]。

文[21]中,二维图像的扫描条垂直于图像流。这些扫描条簇可以处理更多的运动,包括推进和光学缩放。文[22]提出了更复杂运动情况的描述。扫描条的图像获取记录了图像的范围和照相机的运动路径,这对场景的3D完全重建显得非常有效^[23]。

3.4 任意几何校正的多项式变换

多项式函数可以拟合图像和变换理想结果之间的对应关系,就像橡皮筋一样变换图像直到符合理想的结果。多项式变换通过一系列控制点的位移来定义图像空间变换,控制点将图像分成许多多边形区域,每个多边形区域使用双线性插值来填充非控制点。

4 典型的图像对齐(Image Register)模型

图像对齐是图像拼接过程的核心工作。其目标是找出对齐的两幅或多幅重叠图像之间的运动情况。早期,通过标定照相机或预先记录照相机的参数可以进行自动对齐,用户交互对图像对齐很有效。近年来,通过一些复杂的模型方法,实现了不需要用户交互的自动对齐算法。

图像自动对齐方法一般分类如下:(1)在频域上操作的算法,例如基于FFT的相位相关度法。(2)变换优化法,即迭代优化调整照相机参数,使评价误差最小。(3)基于图像特征的算法。

4.1 相位相关度法(phase correlation)

相位相关度法是一种简单的对齐方法,也是最有前途的自动对齐方法。它基于二维傅立叶变换的性质:空间域上的平

移等价于频域相位的平移。相位相关度法最早在1975年由Kuglin 和 Hines 提出,具有场景无关性,能够将纯粹二维平移的图像精确地对齐。De Castro 和 Morandi 发现用傅立叶变换确定旋转对齐,就象平移对齐一样。Reddy 和 Chatterji 改进了 De Castro 的算法,大大减少了需要转换的数量^[24]。两幅图像的平移矢量可以通过它们互功率谱(Cross Power Spectrum)的相位直接计算出来,也就是说,相位相关是基于互功率谱的相位估计的。

设两幅离散图像 $f_1(x, y)$ 和 $f_2(x, y)$ 在空间域简单的平移相关:

$$f_1(x, y) = f_2(x - x_0, y - y_0)$$

相应的傅立叶变换相关:

$$F_2(\xi, \eta) = e^{-j(\xi x_0 + \eta y_0)} F_1(\xi, \eta)$$

采用两个图像规格化互功率谱时,相位相关度的公式如下:

$$\frac{F_1(\xi, \eta) F_2^*(\xi, \eta)}{|F_1(\xi, \eta) F_2^*(\xi, \eta)|} = e^{-j(\xi x_0 + \eta y_0) \frac{IFT}{\delta}} \delta(x - x_0, y - y_0)$$

规格化互功率谱的结果是简单复指数,即相位差。相位差的傅立叶反变换是在平移运动坐标上的脉冲,搜索最大值的位置就是两幅图像的对齐点。

相位相关度法即使在有噪音情况下,如有重复的模式时,结果仍旧可靠和精确。为了加快计算速度,一般用快速傅立叶变换(FFT)计算。

相位相关度法可以扩展到估计平面的旋转和缩放,利用傅里叶变换的旋转和缩放比例的性质,利用 log-polar 坐标,将 $abs(F(\xi, \eta))$ 从直角坐标转换为极坐标 $(\log(\rho), \theta)$, 可以发现旋转和缩放在频域上相应的平移^[24]。文[25]用 IDL (Interactive Data Language) 实现了基于快速傅立叶变换的自动图像对齐算法,把直角坐标转换为 log-polar 坐标,能够象平移一样描述旋转和缩放,从而确定旋转和缩放以对齐图像。文[26]中描述了一种基于 Zernike moments 的快速有效的算法,这个算法允许更大程度地旋转、平移和缩放。

相位相关度法一般需要比较大的重叠比例,通常要求对齐图像之间有50%重叠比例。文[27]中的实验在重叠比例为2.9%仍可以对齐时,讨论了重叠比例减少增加了位移脉冲函数的噪音,出现了许多对齐点相似于正确的对齐点。

4.2 变换优化法(Transform Optimisation)

变换优化法是一种直观的图像拼接方法,这种方法潜在假设了两个空间图像通过一致性变换互相相关,通常情况下,变换矩阵的描述为3.1中的仿射变换或透视变换,变换优化是两个相关图像之间变换矩阵逐步求精的优化过程^[5],可以运用 Levenberg-Marquardt 最小化算法计算光强差平方和的最小值。

设两幅图像 I_1, I_2 , 变换优化矩阵 M , 变换等式为: $X' = MX$

最小化目标评价函数如下:

$$\sum_{x,y} (I_1(x, y) - I_2(x', y'))^2 = \sum_{x,y} e_{x,y}^2$$

目标评价函数不能区分两幅图小区域重叠,如果没有区域重叠,目标函数最小值可为0。为了纠正这个问题,文[27]重新定义了目标评价函数:

$$\frac{\sum_{x,y} (I_1(x, y) - I_2(x', y'))^2}{N} = \frac{\sum_{x,y} e_{x,y}^2}{N}$$

这个评价函数是基于重叠区域像素数量 N 的规格化。

对于变换优化算法,很好的初始估计对于过程的收敛是非常重要的。通常有两个方法,相位相关度法(4.1)和高斯金字塔(分级匹配法)。

金字塔分级估计采用分级图像,分级求精的方案。先在低的(粗的)级别图像上计算确定粗略的结果,然后在高的级别图像计算求得精确的结果。分级估计处理包括:(1)金字塔构造;(2)运动估计;(3)图像变形;(4)逐步求精。

但是实际情况中高斯金字塔初始估计可能不很理想^[27],迭代速度太慢,甚至得不到对齐结果。而相位相关度法却是一个非常好的图像对齐初始化估计方法。

4.3 基于几何特征的图像对齐

几何特征分为低级的特征,如边和角,和高级特征,如物体的识别和特征之间关系。基于低级特征对齐算法一般分为三步:首先过滤图像提取特征集,用这些特征集对两幅图像大概对齐,最后对转换迭代求精。文[28]通过二维高斯模糊过滤可以得到一些低级特征模型,如边模型,角模型和顶点模型。因为角模型提供了比坐标点更多的信息,文[29]中利用几何角模型提出了图像对齐算法。文[30]中基于几何点特征优化匹配和文[31]中利用小波变换提取保留边(dege-preserving)的视觉模型进行图像以齐。

基于高级特征的图像对齐利用低级特征之间的关系或者通过识别出的物体实现对齐,文[32]利用特征图像关系图进行图像对齐。

基于特征的图像对齐依赖特征的检测提取,如果特征不是很明显,可能会导致对齐失败。

4.3.1 角模型对齐(Corner Model) 角模型用参数矢量 $(x_0, y_0, \theta, \beta, \sigma, A, B)$ 描述。其中, x_0, y_0 : 角的位置; θ : 角对称轴的方向; β : 孔径角(aperture angle); σ : 模糊程度参数; A, B : 角内、角外灰度。

角模型利用 Harris 点和窗口尺寸初始化,得到在这个窗口内最适合这个模型的参数矢量和一种评测拟合质量方法。这种拟合方法是用图像灰度和模型内灰度平均最小平方差,取得平方差最小的值的角。这个模型拟合的角称为 D. B. 角,利用两个 D. B. 角中通过参数 $(x_0, y_0, \theta, \beta)$ 定义的两条线计算变换优化的单应矩阵。这种方法比单纯的点特征计算单应矩阵减少了算法复杂性。

在很少的 D. B. 角时,拟合角模型的步骤占了很大运算时间比例,为了提高算法的速度,使用了金字塔结构的方法由粗求精。在粗糙的等级上,相应的是用小的窗口进行拟合角。

角模型对齐算法可以是对光轴的任意旋转和相当大的缩放,但是要求图像重叠的大约是50%,计算速度一般。

4.3.2 基于特征图像关系图的全局对齐 由于对齐过程中会产生对齐误差累加,因此减小累加误差需要全局对齐。文[32]采用一种图描述图像序列时间空间相邻关系的2D 拓扑结构,使用格点(grid point)而不是像素点组成拼接图像的镶嵌块,通过搜索图的最优路径进行全局对齐。

图像关系图描述了图像序列之间的关系,节点表示全局对齐过程中的单个图像。如果两个节点对应的图像在空间或时间上是相邻的,那么这两个节点之间用一条边相连。空间相邻由两幅图像是否重叠决定。对于图像关系图中每一个节点,定义格点为特征点,它是图像对齐的根据。每个节点有一个格点链表,每个格点在其它节点链表中都有对应格点。

图的拓扑描述可以搜索最优路径,这条路径把序列中每幅图像和参考图像连接起来。搜索是基于图中每条边的代价

函数。边的代价表示相邻两图像之间基于图像相关和格点变形的对齐程度。在最短路径上,累计的几何偏离和变形最小。

5 图像融合(Image Blending)

对齐的图像可能会存在强度或颜色的不连续和几何变形留下的缝隙。这可能是因为几何校正、动态的场景或光照条件的变化引起的,甚至是自动照相机,扫描设备的问题。这些问题可以通过图像融合来减小或消除。

图像融合的方法很多:最简单的有光强平均和加权平均融合,复杂的有图像 Vono 权重法^[33]和高斯样条(spline)插值法。其中文^[5,34]的图像融合方法最为显著。

5.1 采样权重函数

文^[5]提出了一种简单有效的方法,对合成结果的每个图像采样的权重分布这样定义:离图像中心越近的像素对最终合成结果贡献越大。即图像中心的权重大,边缘的小。权重分布函数表现为三角形,像一顶帽子,又称帽子函数(hat function)。

利用帽子函数加权平均算法^[35]:

$$C(x, y) = \frac{\sum_k w(x, y) I_k(x, y)}{\sum_k w(x, y, k)}$$
$$w(x, y, k) = \left(1 - \left| \frac{x}{width_k} - \frac{1}{2} \right| \right) \times \left(1 - \left| \frac{y}{height_k} - \frac{1}{2} \right| \right)$$

$w(x, y, k)$ 为帽子函数。 $width$ 和 $height$ 是第 k 个图像的宽和高。

这种方法可以完全消除边缘手工的痕迹,但是边缘低频的斑纹仍旧可见。

5.2 基于欧氏距离的有效权重^[34]

图像的每个像素都分配权重,这个权重与到边缘(或最近不可见点)的距离成比例。在拼接时,主要目的是减少离边缘近的像素点的光强贡献。融合算法中计算距离映射 $d(x, y)$,利用块距离和欧氏距离,计算到最近的透明点($a=0$)或边的距离。融合所有变形图像公式:

$$C(x, y) = \frac{\sum_k w(d(x, y)) I_k(x, y)}{\sum_k w(d(x, y))}$$

w 是单调函数, $w(x)=x$ 。 I_k 是第 k 幅变形图像的光强函数。 d 的计算很简单,取离矩形四条边距离的最小值。

文^[34]算法边缘像素对拼接图像没有贡献,随着向图像中心的逐渐增加。文^[5]算法中,边缘的点对拼接图像的贡献是50%。因此前者的算法减少了合成图像时光强颜色的不连续性。

结论和未来的研究 在本文中,我们论述和分析比较了图像拼接过程中每个步骤(图像获取、几何变形校正、图像对齐和图像融合)的主要方法模型。尽管这些方法模型已经比较完善,但是实际研究中还存在一些更复杂的问题,这也正是我们今后的研究方向。

通常情况下,为了简化几何校正和对齐,图像获取时对照相机运动总是做了一定限制,但是这些限制简化不能描述实际中的照相机运动情况,如手持照相机的任意运动拍摄。不限制的照相机运动会增加几何校正和对齐的复杂性。

在图像拼接时,可能会找不到一种合适的几何变形校正模型。一些不知道的图像之间的运动关系很难选择一个有效

的校正模型,这会导致对齐结果中遗留不合适变形的痕迹,这个问题或许可以在融合中减弱或消除,但不是图像拼接的初衷。因此寻找一个合适的几何校正模型对齐和融合是有利的。对于照相机任意的运动和对三位视差的检测和校正,需要建立一个通用(或近似通用)的照相机运动模型。

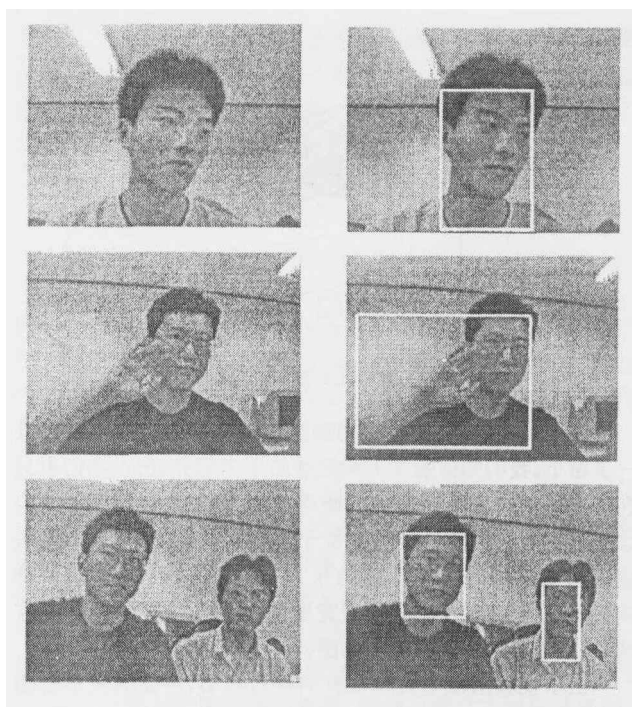
在图像对齐中,希望能够找到一种更好的输入图像之间优化变换算法,不必要考虑不利的因素,如图像间光强跳变,不理想的照相机运动和缺少明显的特征。图像对齐算法的计算量一般都很大,在商业应用中可能是无法忍受的,因此拼接效果和计算量的折衷是商业应用必须要考虑的。既考虑了拼接结果的质量,又能兼顾到效率,这对算法的要求很高。

通过本文的论述,我们希望能够提供给读者国际上图像拼接技术最新研究动态,并且对我国基于图像绘制技术的进一步研究有所启发和帮助。

参考文献

- 1 Anandan P, et al. IEEE Workshop on Representations of Visual Scenes. Cambridge, Massachusetts, IEEE Computer Society Press, 1995
- 2 Kang S B. A survey of image-based rendering techniques: [Technical Report 97/4]. Digital Equipment Corporation, Cambridge Research Lab, Aug. 1997
- 3 Chen S E. Quicktime VR-An image based approach to virtual environment navigation. In: Proc. of ACM SIGGRAPH, 1995
- 4 McMillan L, Bishop G. Plenoptic modeling: An image-based rendering system. In: Proc. of ACM SIGGRAPH, 1995. 39~46
- 5 Szeliski R. Video mosaics for virtual environments. IEEE Computer Graphics and Applications, March 1996. 22~30
- 6 Szeliski R, Shum H Y. Creating full view panoramic image mosaics and environment maps. In: Proc. of ACM SIGGRAPH, 1997
- 7 Xiong Y, Turkowski K. Creating image-based VR using a self-calibrating fisheye lens. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 1997
- 8 Coorg S, Master N, Teller S. Acquisition of a large pose-mosaic dataset. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 1998
- 9 Coorg S, Teller S. Spherical Mosaics with Quaternions and Dense Correlation. IJCV, 2000, 37(3): 259~273
- 10 Shum H Y, Han M, Szeliski R. Interactive construction of 3d models from panoramic mosaics. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, June 1998
- 11 Morimoto C H, Chellappa R. Fast 3d stabilization and mosaic construction. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 1997. 660~665
- 12 Capel D, Zisserman A. Automated mosaicing with super-resolution zoom. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 1998
- 13 Irani M, Anandan P, Bergen J, Kumar R, Hsu S. Mosaic representations of video sequences and their applications. Signal Processing: Image Communication, special issue on Image and Video Semantics: Processing, Analysis, and Application, 1996, 8(4)
- 14 Lee M-C, et al. A layered video object coding system using sprite and affine motion model. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(1): 130~145
- 15 Irani M, Anandan P. Video indexing based on mosaic representations. In: Proc. of the IEEE, 1998, 86(5): 905~921

(下转第150页)



(a)采集到的彩色图像 (b)检测到的人脸区域
图11 部分试验结果

用嘴唇这一特征进行特征匹配,找到最后的人脸区域。实验表

明,此方法执行速度快,鲁棒性好,检测结果正确率高,且对人的头部运动适应性好。此方法适合于视频实时图像中的人脸检测与跟踪,该方法的进一步改进还可用于一般物体的实时检测与跟踪。

参考文献

- 1 Yang G. Human face detection in a complex background. *Pattern Recognition*, 1994, 27(10): 53~63
- 2 Brunelli R, Poggio T. Face recognition: feature versus template [J]. *IEEE Trans PAMI*, 1993, 15(10): 1042~1052
- 3 Sung K, Poggio T. Example-based learning for view-based human face detection. *IEEE Trans Pattern Analysis and Machine Intelligence*, 1998, 18(1): 39~51
- 4 Rowley H A, Baluja S, Kanade T. Neural network-based face detection. *IEEE Trans Pattern Analysis and Machine Intelligence*, 1998, 20(1): 23~28
- 5 艾海舟, 梁路宏, 徐光佑, 张钊. 基于肤色和模板的人脸检测. *软件学报*, 2001, 12(12): 1784~1790
- 6 康学雷, 邵凌, 张立明. 一种基于肤色和模板的人脸检测方法. *红外与毫米波学报*, 2000, 6: 209~214
- 7 Chai D, Ngan K N. Face segmentation using skin-color in video-phone application. *IEEE Trans. CSVT*, 1999, 9(4): 550~566
- 8 姚鸿勋, 刘明宝, 高文. 基于彩色图像的色系坐标变换的面部定位与跟踪法. *计算机学报*, 2000, 2(2): 158~165

(上接第144页)

- 16 Chen C-Y, Klette R. Image Stitching-Comparisons and New Techniques. *CAIP*, 1999: 615~622
- 17 Kang S B, Weiss R. Characterization of errors in compositing panoramic images. In: *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'97)*, San Juan, Puerto Rico, 1997: 103~109
- 18 Gumustekin S. A Introduction to Image Mosaicing. <http://likya.iyte.edu.tr/>, July, 1999
- 19 Sawhney H S, Kumar R. True multi-image alignment and its application to mosaicing and lens distortion correction. In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1997: 450~456
- 20 Peleg S, Herman J. Panoramic mosaics by manifold projection. In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1997
- 21 Rouso B, Peleg S, Finci I. Mosaicing with generalized strips. In: *DARPA Image Understanding Workshop*, May 1997
- 22 Russo B, Peleg S, Finci I, Rav-Acha A. Universal mosaicing using pipe projection. In: *Proc. of IEEE Intl. Conf. on Computer Vision*, 1998
- 23 Rademacher P, Bishop G. Multiple-center-of-projection images. In: *Proc. of ACM SIGGRAPH*, 1998
- 24 Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Image Processing*, 1996, 5(8): 1266~1271
- 25 Xie Hongjie, et al. Automatic image registration based on FFT algorithm and IDL/ENVI. In: *Intl. Conf. on Remote Sensing and GIS/GPS-ICORG*, 2001
- 26 Badra F, Qumsieh A, Dudek G. Rotation and Zooming in Image Mosaicing. In: *IEEE WACV*, 1998: 50~55
- 27 Richmond J. Image Mosaicing <http://www.cs.ubc.ca/~jlrchmo/>, 1999
- 28 Blaszk T, Rachid Deriche Recovering and Characterizing Image Features Using An Efficient Model Based Approach: [Technical Report 2422]. INRIA, Nov. 1994
- 29 Zoghalmi I, Faugeras O, Deriche R. Using Geometric Corners to Build a 2D Mosaic from a Set of Images. In: *Proc. of the Intl. Conf. on Computer Vision and Pattern Recognition*, Puerto Rico, IEEE, 1997
- 30 Nielsen F. Adaptive Randomized Algorithms for Mosaicing Systems. *IEICE TRANS. INF. & SYST.*, VOL. E83 CD, 2000(7)
- 31 Bao P, Xu D. Complex Wavelet-based Image Mosaics using Edge-Preserving Visual Perception Modeling. *Computers and Graphics*, 1999, 23(3): 309~321
- 32 Kang E, Cohen I, Medioni G. A graph-based global registration for 2D mosaics. In: *Proc. of the 15th Intl. Conf. on Pattern Recognition*, Barcelona, Spain, 2000
- 33 Sawhney H S, Kumar R G, Bergen J, Dixon D, Paragano V. Videobrush™, Experiences with consumer Video Mosaicing. In: *IEEE WACV*, 1998: 56~62
- 34 Shum H-Y, Szeliski R. Panoramic Image Mosaics, Microsoft Research: [Technical Report MSR-TR-97-23]
- 35 WONG, Chun Kit Wilbur YANG, Ruiduo ZHANG, Yan. Image Mosaics under Simple Planar Scenes and its Application. A Final Report Submitted to The Course COMP641F "Introduction to Image Processing" at The Hong Kong University of Science and Technology, 2001