

监控视频关键帧的人体高度测量算法^①

魏 玮, 吕 艳, 石可红

¹(河北工业大学 计算机科学与软件学院, 天津 300401)

²(河北省数学研究中心, 石家庄 050024)

摘 要: 由于视频资料众多, 如何从海量的视频数据中找到需要的视频, 已经受到更高的重视, 因此关键帧的提取算法越来越成为研究的热门问题。实验首先对监控视频进行了关键帧提取, 然后利用摄像机标定以及灭点的方法, 对人体的高度进行了分析和计算。结果表明: 基于彩色直方图的语义距离, 比较适用于监控视频中人物关键帧的提取。且算法简单可行, 有较高的准确度, 因此具有一定的实用价值。

关键词: 彩色直方图; 动态规划; 关键帧提取; 人体高度测量

Body Height Measurement Algorithm in Key Frame from Surveillance Video

WEI Wei, LV Yan, SHI Ke-Hong

¹(School of Computer Science and Engineering, Hebei University of Technology, Tianjin, 300401)

²(Mathematics Research Center, Shijiazhuang 050024)

Abstract: It is important to find what we need from the massive video, thus the research of key frame extraction algorithm is becoming more and more hot. First, the paper present the algorithm about extraction key frames from the surveillance video. Then camera calibration and vanishing point is used to analyse and calculate the height of human body. The result showed that semantic distance based on color histogram is more suitable for extracting key frame from the surveillance video. Meanwhile this algorithm of human body height is very feasible and accurate. So the algorithm has practical value.

Key words: color histogram; dynamic programming; key frame extraction; body height measurement

随着现在社会的发展, 监控视频变得很普遍, 得到的视频信息越来越多, 想从视频中找到指定的人物信息, 也就越来越繁琐。其中人的身高是识别一个人很重要的信息。本文就是对监控视频中的人物的高度进行测量, 从而作为识别关键人物的一个重要的特征。

关键帧: 反映所在视频信息内容的一帧或若干帧图像。从较多的视频中取得关键帧, 这样能减少视频中的冗余信息。提取关键帧对自动的检索监控视频具有重要意义。

关键帧提取的算法中, 计算图像帧之间的语义距离对算法的准确度有着很重要的影响。

本文在计算图像帧之间的距离的时候, 采用的是基于彩色直方图的欧几里得距离, 因为直方图计算简

单, 并且对于视角的变换, 直方图具有很好的鲁棒性。本文在彩色直方图的基础上用参考文献[1]中的提取关键帧的算法, 文献中的算法运用了动态规划的方法, 使得计算的效率提高, 对于算法的适用性具有重要的意义。

对于关键帧中的目标人物的高度测量, 现有的图像高度测量方法很多, 很多是基于两幅或者多幅图像的, 但这些方法最大的也是最基本的问题: 图像的配准, 目前没有很好的解决方法。

从单幅图像中获得高度信息, 可以简化测量的工作。这种方法在建筑测量, 交通事故侦测现场还有侦探犯罪现场有很好的应用。因此基于单幅图像测量是以后计算机辅助测量的一个发展方向。

① 基金项目:河北省自然科学基金数学研究专项资助项目(08M006)

收稿时间:2011-09-15;收到修改稿时间:2011-10-31

1 关键帧的提取

在文献[1]中提到的关键帧的动态算法具有很好的计算效率,但是这些是基于假设已经得到很好的帧之间的语义距离(semantic distance)之上的。但在实际的实验中,由于各种因素,例如光照、运动等各种因素会影响实验的结果。因此选择一种好的语义距离的计算方法对于结果的准确性具有很重要的意义。本论文基于彩色直方图的优点,计算彩色直方图的语义距离,将彩色直方图的语义距离用于关键帧提取。并将多种关键帧提取的计算方法,进行了比较。

1.1 彩色直方图

彩色直方图(color histograms)是将每一个颜色通道分成不同的等级,这样整个颜色空间,就像是被分成很多小的小的“格子”(bin)。通过计算颜色落在每个小区间内的像素数量可以得到颜色直方图。直方图具有对光照、运动具有很好的鲁棒性,以及计算简单的特点。

在本文中,利用的是计算整个图像的彩色直方图,然后基于直方图分别计算帧之间语义距离。再利用文献[1]中的动态优化算法进行关键帧提取。

在监控视频中,可以根据人物的衣服的颜色特点,然后选择用彩色直方图,目的是根据衣服颜色信息特征,更好的识别人物。因为在监控视频中,一般摄像头是固定的,也就是场景是固定的,在彩色直方图中,同一个目标人物的在短时期同一场景的帧的彩色直方图很相似,但是不同目标人物的同一场景的帧的彩色直方图相对来说语义距离会大,因此采用这种方法,可以将关键帧提取出来。彩色直方图中的计算可以基于RGB空间也可以基于HSV等其他颜色空间,以基于RGB空间为例:首先在RGB三维空间中,将颜色量化(color quantization),也就是将颜色空间划分为若干个小的颜色空间。

颜色直方图分为小的“格子”(bin)。如图1所示。其中本算法中将每一个颜色分为16个等级,共得到4096个bin。

1.2 颜色直方图的语义距离的计算

计算颜色直方图语义距离的时候,可以选择用巴氏距离(Bhattacharyya coefficient)来确定两帧之间的语义距离,也可以采用的是欧几里得距离(Euclidean distance)表示语义距离,在本文中用的是欧几里得距离。

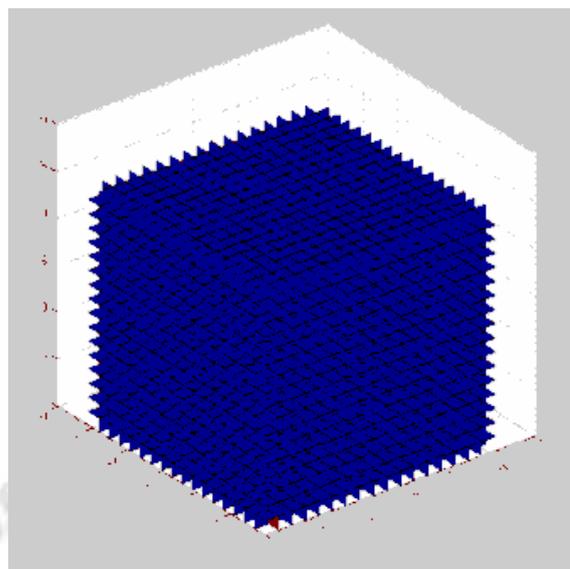


图1 颜色直方图的表示

Bhattacharyya coefficient 表示语义距离的公式:

$$d(R, Q) = - \sum_{u=1}^{4096} \sqrt{hist_R^u \cdot hist_Q^u} \quad (1)$$

用的是 Euclidean distance 表示帧的语义距离公式为:

$$d(R, Q) = - \sqrt{\sum_{u=1}^{4096} (hist_R^u - hist_Q^u)^2} \quad (2)$$

公式中的 $hist_R^u$ 表示的是图像 R 中的直方图第 u 个小方格中的像素的个数。

得到相邻帧之间的语义距离之后,利用文献[1]中的关键帧的提取方法,建立最小堆,利用动态规划的优化方法和贪心算法得到最能反映视频信息的关键帧(与相邻帧的语义距离最小的帧)的序列号。该算法的时间计算的复杂度为 $o(n \log n)$,空间复杂度为 $o(n)$,具有很好的适用价值。实验的结果如文中的第三部分所述。

2 图像中的人物的身高的测量

人的高度在一般的情况下指的是头部顶点到地面的高度,用单幅图片计算人体的高度,就是计算空间中的从头部顶点到地面的垂直于地面的垂线段的距离。这就需要找到空间中的头部顶点的图像点,以及垂直于垂线段在地面的垂足的图像点,然后根据投影

原理得到人的高度。

2.1 投影原理

在透视投影模型中,空间中的一点 $M(X,Y,Z)$,其次坐标表示为 $\tilde{M}(X,Y,Z,1)$,对应的图像点为 $m(x,y)$,其次坐标表示为 $\tilde{m}(x,y)$,二者之间的关系表示为

$$\begin{aligned}\lambda\tilde{m} &= PM = (p_1, p_2, p_3, p_4)\tilde{M} \\ &= K[R \quad T]\tilde{M} \\ &= K[r_1 \quad r_2 \quad r_3 \quad T]\tilde{M}\end{aligned}\quad (3)$$

其中 λ 为非零因子, P 为投影矩阵, K 是内参数矩阵, R 是旋转矩阵, T 是平移向量。

空间中的一个平面的点 $M_1(X,Y)$, 其次坐标 $\tilde{M}_1(X,Y)$ 和像平面的点 $m_1(x,y)$, 其次坐标 $\tilde{m}_1(x,y)$, 对应关系用公式表示为

$$\begin{aligned}\lambda\tilde{m}_1 &= H\tilde{M}_1 = (p_1, p_2, p_4)\tilde{M}_1 \\ &= (h_1, h_2, h_3)\tilde{M}_1 = K[r_1 \quad r_2 \quad T]\tilde{M}_1\end{aligned}\quad (4)$$

H 是单应矩阵。

2.2 提取特征点

应用张正友的摄像机的棋盘标定法,将地面作为参考平面,计算地面的单应矩阵和摄像机的内参数矩阵(内参数矩阵一般不会变化),根据文献[2]中的方法计算得到摄像机的投影矩阵。得到投影矩阵之后,需要确定人的头部顶点和足点对应的图像的特征点,通过公式(3)就可以得到人体的实际的三维坐标,最后可以确定人体的高度。

将地面作为参考平面,与参考平面垂直的无穷远点的像(消影点)为垂直消影点。首先计算垂直消影点 m_z , 利用公式

$$\lambda\tilde{m}_z = K(r_1 \times r_2) = p_3 \quad (5)$$

根据旋转矩阵的性质,可以得到

$$\begin{cases} r_3 = K^{-1}\lambda\tilde{m}_z \\ r_1^T \bullet r_3 = 0 \\ r_2^T \bullet r_3 = 0 \end{cases} \quad (6)$$

根据公式(6)可以计算垂直消影点。

在关键帧中,利用背景减除法得到目标人物的前景部分,即得到人的整体轮廓。由得到的垂直消影点,计算人物的头部顶点,以及脚的特征点。垂直消影点

y 方向如果为正值,则选取垂直消影点与前景轮廓的所有的点的距离的最大的点为头部顶点,否则就是选取最小的值。

从前景轮廓确定两脚的区域,分别对两脚的区域选择最低的点作为脚步特征点。特征点的实验显示结果如第三部分的图 3 所示。

得到了头部特征点和脚部特征点就可以得到人的整体的高度在图像中的特征,利用投影原理得到人的高度。

3 实验结果

试验中用的监控视频是人进出实验室的视频,这一段视频中包括 100 帧图像,实验中取得关键帧的帧数是 4,用本文中的基于彩色直方图的算法得到的关键帧序号依次为: 3,39,53,75。对应的图像分别如图 2 所示。



图 2 关键帧序号从上到下,从左到右依次为 3,39,53,75

当选择的关键帧数是 2 的时候,对应的帧如图 3 所示:



图 3 关键帧序号是: 18 和 75

从图片可以看出,实验得到的关键帧图像可以很好的反映视频信息,当选择的关键帧数是 4 时,得到的图像是首先是没有进入视频范围的图像,然后是人进入视频范围内,人的半身图像,以及包括整个人的

图像。当选择的关键帧数是 2 的时候得到的关键帧是人未进入视频范围的图像, 以及进入视频范围的图像, 很好的反映了视频信息。

利用帧平均法提取关键帧和直方图平均法提取关键帧, 实验的结果分别如图 4 和图 5 所示。帧平均法得到的都是未有人物进入视频范围的关键帧图像。直方图平均法得到的视频图像在计算后面的人体高度测量时不如彩色直方图得到的关键帧准确性高。



图 4 帧平均法得到的关键帧序号分别为: 1 和 23



图 5 直方图平均法得到的关键帧序号: 33 和 65

通过对比, 可知用彩色直方图的方法得到的关键帧可以较好的反映视频信息。因此可以得出基于彩色直方图的关键帧提取算法比较适用于监控视频的关键帧提取。

对于人体高度测量, 首先是图像的特征点的选取, 实验的结果如图 6 所示。

其中的#表示的是头部特征点, f1 和 f2 分别表示的两个脚步的特征点。根据投影矩阵以及公式(3)

得到人体的高度。试验中的人体的实际高度是 150cm, 得到的是值为 1.493421 和 1.523458。这个准确度对于后面的基于人体高度的人体的识别具有很好的意义和效果。



图 6 人体在图像特征点

参考文献

- 1 Tiecheng Liu, John R. Kender. Computational Approaches to Temporal Sampling of Video Sequences. *ACM Trans. on Multimedia Computing, Communications and Application*, 2007,3(2).
- 2 黄凤荣. 三维重建问题研究[博士学位论文]. 中国科学院自动化研究所, 2004.
- 3 王斌. 摄像机定标与单视测量技术研究[硕士学位论文]. 西安电子科技大学, 2008.
- 4 Jifeng Ning, Lei Zhang, David Zhang, Chengke Wu. Interactive Image Segmentation by Maximal Similarity Based Region Merging. *Pattern Recognition*, 2010,43:445-456.
- 5 董秋雷, 吴毅红, 胡占义. 基于视频的实时自动人体高度测量. *自动化学报*, 2009,35(2):137-144.
- 11 Wiskott L, Fellous R, Kruger N, et al. FACE recognition by elastic bunch graph matching. *IEEE Trans. PAMI*, 1997, 19(7):775-779.
- 12 Atonen T, Hadida, Pietikainen M. Face description with local binary patterns: application to face recognition. *IEEE Trans on PAMI*, 2006,28(12):2037-2041.
- 13 Ojala T, Pietikainen M, Harwood D, A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 1996,29(1):51-59.

(上接第 204 页)

2002,22(10):1090-1104.

- 7 李姜龙. 基于局部二值模式的 ASM 人脸特征定位方法研究. *闽西职业技术学院计算机系*, 2009,6(11)2.
- 8 Turk MA, Pentland AP. Face recognition using eigenfaces. *Journal of Cognitive Neuroscience*, 1991,3:71-86.
- 9 彭辉, 张长水, 荣钢等. 基于 K-L 变换的人脸自动识别方法. *清华大学学报(自然科学报)*, 1997,37(3):67-70.
- 10 边肇祺, 张学工, 阎平凡, 等. 模式识别. 北京: 清华大学出版社, 2000.