

基于区域深度特征的人头检测方法

何扬名, 杜建强, 肖贤波

(江西中医学院 计算机学院, 江西 南昌 330004)

摘 要: 根据摄像机自上而下拍摄检测人头的特点, 提出了以目标及其周围像素为候选匹配区域进行匹配来计算视差的人头检测方法. 该方法首先用左视图检测所有可能的人头目标, 然后选取目标及其周围像素作为匹配的整体, 根据一些匹配的约束条件, 用其来和右视图进行匹配, 算出以目标及其周围像素为整体的视差. 将目标及其周围像素作为整体, 克服了人头区域纹理比较弱缺点, 使得匹配的准确率有所提高, 从而提高了视差计算的准确度.

关键词: 区域深度; 人头检测; 立体匹配; 客流统计

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-7180(2013)11-0039-04

Detecting Human Head by Depth Characteristics of Regions

HE Yang-ming, DU Jian-qiang, XIAO Xian-bo

(School of Computer, Jiangxi University of Traditional Chinese Medicine, Nanchang 330004, China)

Abstract: According to the features of cameras shooting from up to down, a new method of detecting human head is put forward, which uses the pixels of the target and pixels around the target. First, detect all possible targets in the left image. Then chose the pixels of every target and the pixels around the target as an integrated object, and use them to match the right image according to some constraint conditions. Thus the disparity of the integrated object can be gotten. The method integrating the pixels of target and around the target overcomes the shortcoming of weak texture of human head, which would increase the accuracy of matching. So the accuracy of disparity can be increased, too. According to the disparity, some fake target can be excluded. And the accuracy of recognizing human head would be increased.

Key words: depth of regions; detecting human head; stereo matching; passenger flow statistical

1 引言

在利用计算机检测人的活动的系统中, 由于摄像机的安装条件限制, 常需要将摄像机安装在人头顶部的位. 由于人头顶部形状和特征相对固定, 所以对人头顶部的检测就成了检测人的活动的有效方法之一. 头顶部检测在许多领域得到了应用, 特别是乘客流量统计领域. 在客流统计中, 摄像机通常安装在人头上, 从上而下进行拍摄. 摄像机采集的实时图像作为输入, 客流统计系统对其进行处理, 检测出图像中是否存在头部或者头部的个数、位置等信息.

摄像机从上而下拍摄, 得到的人头顶部图像是

一个类圆形目标. 利用单目视觉的方法来检测人头目标已经发展出很多, 文献[1]中利用结合梯度与彩色直方图的方法来进行人头顶部的识别, 由于彩色图像的数据量很大, 导致系统的实时性不理想. 人头顶部的目标也可以采用检测圆的办法来进行检测^[2-3], 其中文献[2]中采用随机抽取四个边缘点的方法, 通过计算它们的距离来判断是否存在一个可能的圆形目标, 这种方法对噪声较大的图像, 由于随机抽取的点命中率很低, 该方法很难适用于复杂的环境. 模糊 C 均值聚类方法^[4]由于首先要确定聚类数目而很难在这种聚类数目很难确定的情况下使用. 检测类圆形目标的最常用的方法就是 Hough 变

换(HT)^[5]. 基于单目图像处理技术的人头检测方法的主要问题在于顶部相对于复杂的背景并没有明显特征,例如和顶部颜色相似的背包、提箱,拥挤的人群中的空隙以及深色着装的行人的背部、肩部等极易被系统误判为顶部,因此这些系统不能达到很高的准确率^[6-7].

摄像头自上而下拍摄,由于这种环境下,人头离摄像头的距离相对背景要近一些,所以利用双目立体视觉的深度感知特性识别人头,可以尽量避免虚假人头的出现. 基于双目视觉的目标检测方法一个很大的缺点就是实时性较差^[8-9],文献[6]中利用了目标视差的方法来检测人头目标,提高了实时性. 但是由于人头目标的纹理往往很弱,目标视差也常常难以提取. 而且由于人头纹理弱,利用灰度相关的匹配办法有时会使得目标匹配产生错误. 而且文献[6]利用了 Hough 变换来检测人头,对分辨率较大的图像进行检测时,其实时性不够理想.

本文提出基于区域深度特征的人头检测方法,并用实验验证其较好的实时性和较高的准确性.

2 双目视觉检测原理

双目立体视觉是基于视差的原理,如图 1 所示为双目立体成像原理图,两摄像机的投影中心连线的距离就是基线距 B . 两摄像机在同时采集空间物体的特征点 P ,分别在左右图像平面上成像,他们的图像坐标分别为 $P_l = (X_l, Y_l)$, $P_r = (X_r, Y_r)$. 通常两个摄像机的图像面在同一个平面上,则特征点 P 在 Y 方向的坐标 $Y_l = Y_r = Y$,则由几何关系可以得到:

$$\begin{cases} X_l = f \frac{x_c}{z_c} \\ X_r = f \frac{(x_c - B)}{z_c} \\ Y = f \frac{y_c}{z_c} \end{cases} \quad (1)$$

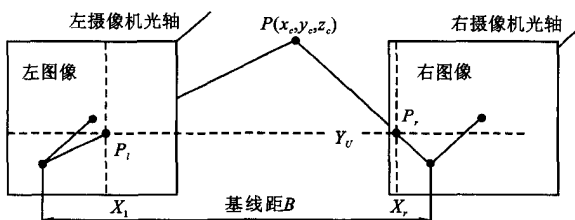


图 1 双目立体成像原理图

则水平方向的视差为 $D = X_l - X_r$. 于是就可以计算出点 P 在左摄像机坐标系下的三维坐标是:

$$\begin{cases} x_c = \frac{B \cdot X_l}{D} \\ y_c = \frac{B \cdot Y}{D} \\ z_c = \frac{B \cdot f}{D} \end{cases} \quad (2)$$

因此,只要将左摄像机图像上的像素点在右摄像机图像上找到对应的匹配点,就可以算出该像素点的视差,从而确定出该点的三维坐标. 这种方法是点对点的运算. 由式(2)可以看出视差越大的匹配点,其离摄像头的距离就越小. 在摄像机从上而下拍摄的场景,人头离摄像机是最近的,所以具有较大的视差.

3 基于区域深度特征的人头检测原理

3.1 区域的生成方法

由于摄像机自上而下进行拍摄,人头区域的像素灰度变化比较小,可以采用如下的区域生长算法来对图像进行区域划分.

将图像的像素划分成 $L \times H$ 小矩形(通常为 3×3 的小矩形),然后比较相邻小矩形的中心像素灰度的差值,如果其差值小于等于某一事先设定的阈值,则将相邻小矩形合并为一个区域;如果差值大于事先设定的阈值,则不属于同一个区域.

根据这种方法生成的区域,其形状不光滑,常常带有毛刺,有时当两个人头距离比较近的时候,两个人头区域之间被一个狭窄的区域连接起来. 由于人头具有类圆形的特征,所以对各区域采用具有圆形特征的结构元素进行开运算操作,这样既可以去除区域的毛刺,还可以使得每个区域的类圆形特征更加明显.

一副图像进行这种区域生长和开运算后,将得到很多区域,有时区域数量达到上百个. 如果对每个区域都在另一幅图像中进行匹配运算,其运算量将比较大. 而且有很多区域根本不具有人头的特征,将这种区域进行匹配,完全是多余的. 所以需要对区域进行筛选. 筛选的参数主要有三个:面积、灰度均值、区域圆度. 人头区域的面积在一定的范围内,如果该区域面积不在面积范围内,则该区域可以排除. 由于人头区域灰度偏低,所以该区域的灰度均值如果高于某个阈值,则该区域可以排除. 由于人头区域具有类圆形的特征,可以用如下的圆度参数来评价一个区域的类圆特征:

$$C = F / (\max^2 * \pi i),$$

式中, C 代表圆形度, F 表示区域的面积, $\max$ 表示

区域中离区域中心最远的距离。

C 的值在 $0 \sim 1$ 之间,当区域为圆形的时候, C 为 1。 C 的值越接近 1,说明该区域越圆。一般将 C 的阈值设为 0.5,当区域的 C 值小于 0.5 的时候,该区域被排除。

经过以上的处理后,剩下的区域即可计算其视差,视差的大小预示着区域深度的特征。

3.2 区域深度特征的提取

立体视觉算法的首要目的就是在同一场景的两幅或多幅图像对中搜索匹配点用以构成对应关系 (Correspondence)。利用匹配对应关系可以获取视差 (Disparity),然后根据三角几何关系则进一步获得场景的三维世界坐标或深度信息 (Depth)。整个算法中最复杂、是最困难也就是确定匹配对应关系。立体匹配方法主要可分为两大类:一是基于窗口的局部匹配方法;二是基于能量函数形式的全局匹配方法。匹配的相似性测度的方法很多,常用的是归一化灰度互相关 (Normalized Cross Correlation, NCC),计算如式(3):

$$NCC(u, v) = \frac{\sum_{(u, v)} I_R(u, v) \cdot I_T(u + d, v)}{\sqrt{\sum_{(u, v)} I_R^2(u, v) \cdot \sum_{(u, v)} I_T^2(u + d, v)}} \quad (3)$$

式中, u, v 是匹配窗口中的像素坐标, d 视差搜索范围, $I_T(u, v)$ 为待匹配图上像素 (u, v) 的灰度值, $I_R(u, v)$ 为参考图上像素 (u, v) 的灰度值。

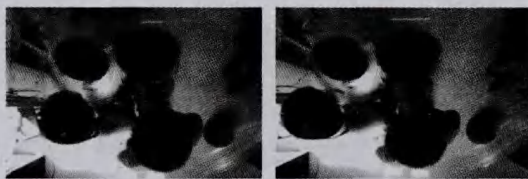
局部匹配和全局匹配都力求在整个图像范围内计算视差,但是在很多情况下所要检测的目标只占整个图像的小部分,这种情况下在全图像范围内计算视差将使得系统的实时性大大下降。本文提出一种以目标和周围像素为候选匹配区域,然后根据匹配的约束条件,在被匹配的图像中寻找该候选区域的匹配区域。从而计算出这个整体的视差,根据该区域的视差判断该目标是否满足深度要求,如果不满足,则认为该目标是虚假的人头目标,要排除该目标。这就首先需要利用单目视觉的算法提取人头目标,这时候提取的目标可能存在一些虚假的人头目标,在将检测到的各个人头目标逐个在另一幅图像中寻找匹配目标,计算相应的视差,根据视差排除虚假人头。

4 实验结果及分析

图 2 为实验室拍摄的两张图片,图 4 为两张原图的视差图,由于人头目标内部的纹理不是很明显,

所以检测到的人头中间部分的视差具有不确定性,而且整个视差的计算时间为 0.8 s 左右,这对于实时性要求很高的客流统计场合是完全不够的,而且由于视差图比较乱,从视差图中提取人头目标有一定难度。

图 3 所示为对左视图利用区域生长方法生成区域,并对其进行形态学开运算,并用区域的面积和圆度等判据将一些特征不符的区域排除之后得到的结果。从图中可以看出检测到一个虚假人头。再提取每个检测到的人头目标及其附近的像素灰度值(图中用矩形框框起来的区域),根据外极线约束以及视差范围的约束条件,在右视图进行匹配,然后计算出每块区域的总体视差,得到如表 1 所示的结果。从表 1 看,目标 5 的视差低于 10,远远低于其他目标的视差,所以可以排除目标 5 作为人头目标,因为人是站立的,离摄像头比较近,根据前面的视差理论,其视差也比较大,如果目标的视差小于某个阈值,则可排除。



(a) 左视图

(a) 右视图

图 2 双目视觉中的原始图像

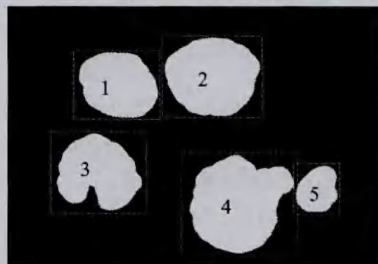


图 3 左视图检测到的目标



图 4 左右视图的视差图

用这种目标视差的方法程序的实时性要大大改善,本程序运行在 Pentium 3.0 GHz 的 CPU 上,该方法只需运行 0.08 s 左右,完全满足实时性要求。

表 1 目标的中心位置及其视差

	目标 1	目标 2	目标 3	目标 4	目标 5
左视图坐标	(153,186)	(135,354)	(333,140)	(403,391)	(372,531)
右视图坐标	(153,155)	(138,330)	(327,79)	(398,351)	(368,512)
目标视差值	31	24	61	40	9

5 结束语

本文根据摄像头自上而下拍摄的特点,在双目立体视觉的基础上提出了新的人头检测方法.该方法结合了单目视觉和双目视觉的各自优点,将目标及其周围像素作为参考匹配的对象进行匹配,利用视差阈值排除虚假人头目标,提高了识别的准确率.该方法由于只用目标及其周围像素作为整体进行匹配,避免了常规立体视觉的全图像计算视差的方式,所以其实时性有很大的改善.而且该方法在一个视图中检测完目标后,还进一步在另一视图中进行匹配处理,可以对匹配前的目标检测的参数阈值采用比较宽松的办法,从而可以尽量防止目标的漏检情况.实验也证明了该方法的有效性.

参考文献:

[1] 孙棣华,闫川,赵敏,等. 结合梯度与彩色直方图的人头识别[J]. 重庆工学院学报,2007,21(12):1-5.

[2] Chen Tehchuan, Chung Kuoliang. An efficient randomized algorithm for detecting circles [J]. Computer Vision and Image Understanding, 2001, 83 (2): 172-191.

[3] Chung Kuoliang, Huang Yonghuai, Shen Shiming, et al. Efficient sampling strategy and refinement strategy for randomized circle detection [J]. Pattern Recognition, 2012, 45(1):252-263.

[4] Ahmed M N, Yamany S M, Mohamed N. A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data[J]. IEEE Trans on Medical Imaging, 2002, 21(3):193-199.

[5] Walsh D, Raftery A E. Accurate and efficient curve detection in images: the importance sampling Hough transform [J]. Pattern Recognition, 2002, 35 (7):

1421-1431.

[6] 于海滨,刘济林. 基于区域视差提取的视觉客流统计方法[J]. 传感技术学报,2007,20(7):1546-1550.

[7] Huang Xiaoyu, Li Liyuan, SIM T. Stereo-based human head detection from crowd scenes[C]//Proceedings of International Conference on Image Processing. Singapore, 2004:1353-1356.

[8] Krotosky S, Cheng S, Trivedi M. Real time stereo based head detection using size, shape and disparity constraints[C]//Proceedings of IEEE Symposium on Intelligent Vehicles. Beijing, 2005:550-556.

[9] 尹传历,刘冬梅,宋建中. 改进的基于图像分割的立体匹配算法[J]. 计算机辅助设计与图形学学报,2008,20 (6):808-812.

作者简介:



何扬名 男,(1973-),博士,讲师. 研究方向为数字图像处理.



杜坚强 男,(1968-),博士,教授. 研究方向为数字图像处理和数据库.



肖贤波 男,(1978-),博士,副教授. 研究方向为计算机算法.