**文章编号**:1002-2082(2011)05-0904-05

# 基于积分图像的人体肢体检测算法

陈 华,胡春海

(燕山大学 河北省测试计量技术及仪器重点实验室,河北 秦皇岛 066004)

摘 要:为了实现视频序列图像中运动人体肢体的检测,提出了一种基于积分图像和类哈尔特征的检测方法,类哈尔特征(Haar-like feature)因其固有的特点,适用于检测矩形或类矩形的图像区域,积分图像的特点是可以利用类哈尔特征快速定位人体肢体在图像中的位置。积分图像用于表示视频中的原始图像,边缘检测模板与原始视频每一帧进行卷积之后,通过累加计算可以得到该图像的积分图像,然后利用类哈尔算子,根据设定的阈值,即可以准确定位人体肢体位置。实验结果表明:该方法较背景减除法检测人体肢体准确。

关键词:积分图像;类哈尔特征;肢体检测;视频图像

**中图分类号**:TN206;TG391.41 **文献标志码**:A

## Human body limbs detection algorithm based on integral image

CHEN Hua, HU Chun-hai

(Key Laboratory of Measurement Technology and Instrumentation of Hebei Province, Yanshan University, Qinhuangdao 066004, China)

**Abstract**: The method based on the integral image and the haar-like feature for detecting body limbs is proposed for video images. Haar-like feature is suitable for detecting rectangular or rectangular-like image areas because of its inherent characteristics, and the integral image can quickly locate the body by the haar-like feature in the image. The integral image is the original image for video. After edge detection template is convoluted with each frame of the original video, they are added to obtain the integral image. The body physical location can be calculated by the integral image and using haar-like feature according to the set threshold. The experiment results show the algorithm is better than the background subtraction.

**Key words**: integral image; Haar-like feature; limbs detection; video image

## 引言

在计算机视觉领域,运动人体检测、分析、理解是一个研究热点,在智能安全监控系统、体育运动分析、医疗诊断等领域具有广阔的应用前景。其中人体检测处于最基础和最关键的地位,检测的目的是在序列图像中将人体变化区域从背景图像中提取出来,运动区域的有效分割对于目标分类、跟踪和行为理解等后期处理非常重要,所以说人体检测既是人体跟踪系统后续处理的数据来源,也是能否实现实时跟踪与识别的关键预处理阶段。但是由于人体各肢体相对运动的复杂多变,所以在视频中准确预测人肢体的位置是一个难题,这就要求有快速准确的肢体检测算法。

现有的进行人体检测比较常用的方法有光流法、背景减除法和相邻帧差分法[1-4]。光流法是根据运动目标随光流变化的特性,通过计算位移向量光流场来初始化基于轮廓的算法,可以有效地提取和跟踪运动目标,但该方法计算实时性和实

**收稿日期**:2010-09-09; **修回日期**:2010-11-12

**作者简介**:陈华(1972—),女,黑龙江大庆人,博士,副教授,主要从事计算机视觉、图像处理研究工作。
E-mail:chenhua@ysu.edu.cn

用性较差;背景减除法是利用序列图像的连续性,将序列图像相邻帧的图像相减,可以检测到因人体肢体运动而出现的像素微小变化,同时可去除图像中的背景信息。其优点是计算简单,原理清楚,缺点是受光照变化影响太大,环境中光强的变化容易影响检测结果,同时不容易估计背景信息,且图像的相减比较耗费时间;相邻帧差分法是利用时间上连续的图像序列前后几个相邻帧,基于像素的时间差分,通过阈值化提取图像中的运动区域,但该方法提取的目标轮廓不清晰。

基于以上方法存在的问题,本文提出一种能快速检测视频序列中人体肢体的方法。该方法引入积分图像作为原始图像的中间表示形式,利用积分图像可以快速找到图像中人体肢体位置。边缘检测器与原始帧图像做卷积之后,计算该卷积图像的积分图像。获得积分图像之后,设定阈值,就可在积分图像的任何位置定位人体肢体。

## 1　基于积分图像的肢体检测

美国麻省理工学院的 Paul Voila 博士于 2001 年提出了一种基于类哈尔特征(Haar-like Feature)的快速检测算法[5],用于在图像中检测人脸[6]。该方法的最大优点是快速准确,在 1.1 G Hz笔记本上对 384×288 像素的动态图像检测,可以达到 12 帧/s 的检测速度,准确率超过 90%。该检测方法之所以快速的原因在于引进了类哈尔特征(Haar-like Feature)。类哈尔特征之所以能应用于检测人脸,是因为 Paul Voila 将脸部的不同部分的像素灰度特征建模为不同的矩形区域。在此之后,有不少专家对 Paul Voila 的方法进行诸多改进,并将类哈尔特征成功应用于步行者检测、人脸检测、车辆检测等不同领域。

将对人脸的检测推广到对人体各种肢体的检测,并引入积分图像,将其与类哈尔特征结合,以实现对人体各肢体的快速检测。要利用类哈尔特征实现检测,也需要对待检测的人体肢体做一个假设,因此将视频中的人体肢体如躯干、上肢、下肢等,建模为矩形,即人体是由不同尺度不同长度的矩形组成。基于上述的假设,就可以利用类哈尔特征和积分图像对肢体进行检测。

### 1.1　积分图像

在给定视频中跟踪人体时,人体肢体在图像上呈现为不同尺寸的矩形区域,需要跟踪的就是此类矩形组成的人体模型。为了快速在图像中定位人体肢体,于是引入积分图像作为原始图像。积分图像上$(x,y)$点的值是原始图像上该点左上方像素值之和[5-6],如(1)式所示:

$$ii(x,y)=\sum i(x',y') \tag{1}$$

式中:$ii(x,y)$为积分图像;$i(x',y')$为原始图像。

图 1 是在原始图像中计算积分图像的示意图。

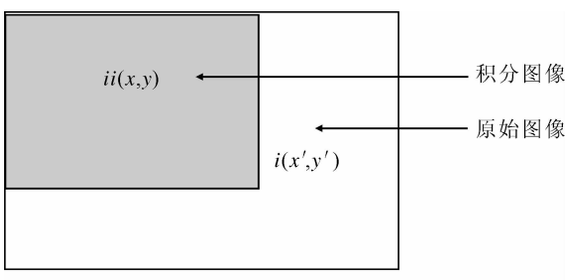



**图 1　积分图像示意图**

**Fig. 1　Integral image**

使用累加可在原始图像中计算积分图像。计算过程可利用(2)式与(3)式。

$$s(x,y)=s(x,y-1)+i(x,y) \tag{2}$$

$$ii(x,y)=ii(x-1,y)+s(x,y) \tag{3}$$

式中:$ii(x,y)$为积分图像;$s(x,y)$为累加值;$i(x',y')$为原始图像。

$s(x,y)$是每行的累加和,且有 $s(x,-1)=0$,并 $ii(-1,y)=0$,在原始图像中计算积分图像非常快,因为只是一个累加过程。

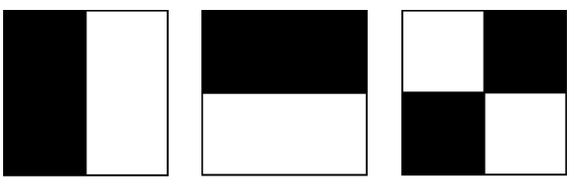
### 1.2　类哈尔特征

在肢体检测时,类哈尔特征可用于描述图像中的矩形区域,例如,可用类哈尔特征描述人体肢体所在区域与临近区域之间的差异,因而可以实现对人体肢体快速定位。另外一个更重要的使用类哈尔特征的原因是:基于特征的系统检测速度往往都比基于像素值计算的系统快得多,而且更能增加类间差异而减少类内差异,使类哈尔特征在检测与定位图像中的人体肢体时更有效率,使跟踪更为容易。

因此本文的检测方法利用类哈尔特征遍历积分图像以搜索人体肢体,假设人体各肢体的尺度已知,只需在积分图像上应用单一类哈尔算子搜索单一尺寸的人体肢体即可。由于积分图像固有的计算便利性,使得计算每一种类哈尔矩形特征

值都非常方便,从而有利于在图像上快速搜索并定位人体肢体。

人体各种肢体的检测过程都是基于类哈尔模板和简单的矩形特征,且均可使用积分图像高速计算类哈尔特征的特征值来实现。类哈尔特征与边缘检测器相似,但有着更多的形状,根据需要检测的人体肢体而定。类哈尔特征的特征值计算过程描述如下:每个类哈尔特征由两到三个黑白相邻的矩形组成,每个类哈尔特征的特征值可以计算如下:1) 计算黑色矩形区域像素值之和;2) 计算白色矩形区域像素值之和;3) 计算两者之差。图 2 所示为 3 种常见类哈尔特征。



(a)类型1　(b)类型2　(c)类型3

**图 2　类哈尔特征集**

**Fig. 2 Haar-like feature**

### 1.3　类哈尔特征计算

图 3 所示为在积分图像中计算类哈尔矩形特征的值[5-6]。

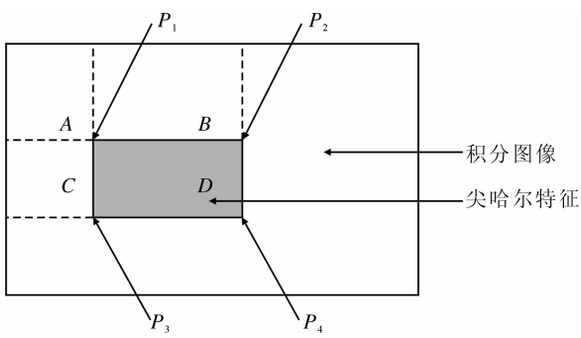



**图 3　类哈尔特征的计算示意图**

**Fig. 3　Calculation of haar-like feature**

在上图中要计算的是矩形区域 $D$ 内的像素值之和,可利用(4)式计算

$$p_1=A,\ p_2=A+B,\ p_3=A+C,\ p_4=A+B+C+D$$

$$p_1+p_4-p_2-p_3=A+A+B+C+D-A-B-A-C=D \tag{4}$$

## 2　肢体检测过程

利用积分图像实现人体肢体检测过程,如图 4 所示。

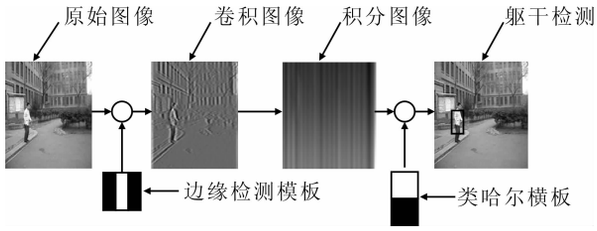
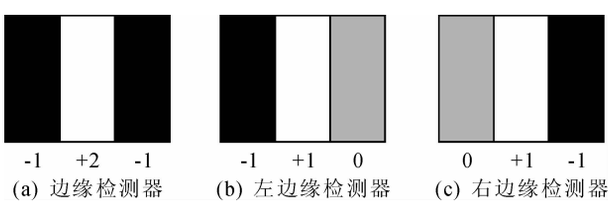



**图 4　利用积分图像检测人体肢体过程**

**Fig. 4　Limbs detection by the integral image**

首先将原始图像与边缘检测器做卷积得到卷积图像,目的是抑制图像中的噪声并突出要检测的人体肢体,类哈尔特征可作为人体肢体的检测器,该检测器是以黑色矩形为背景,而白色矩形为前景的矩形,如图 5 所示。为使特征在亮度均匀区域的响应为 0,将不同的权重赋予该模板的不同部分:白色前景为 2,黑色背景为-1,灰色背景为 0。用此模板作为检测器,与原始图像卷积,可以抑制大部分的噪声,检测出原始图像中的人体肢体所在的区域,但同时也会检测到因灰度值急剧变化产生的噪声;然后计算该卷积图像的积分图像,积分图像(integral image)作为原始图像的一种中间表示形式,其上每一点的数值都是原始图像对应点左上方区域像素值的累加;获得积分图像之后,设定阈值,将图 2 中的类哈尔特征用于遍历积分图像并计算该类哈尔特征的特征值,根据特征值数值大小就可以较快速度在积分图像的任何位置定位人体肢体。




(a) 边缘检测器　(b) 左边缘检测器　(c) 右边缘检测器

**图 5　边缘检测器**

**Fig. 5　Edge detector**

## 3　实验过程

在给定的原始视频图像中,给定的视频持续 10 s,一共 132 帧。从中抽取 3 帧加以检测,每帧大小为 240×320 像素,图 6 为在视频中抽取的 3 帧图像。然后对这 3 帧图像采用积分图像和类哈尔特征进行肢体检测。

图 6　视频中的原始图像

**Fig. 6　Original images in the video images**

从实验图像 7 可看到,因为阈值选择的原因,存在误检位置。太大的阈值会减少误检率,但有可能漏掉真正的人体肢体位置;太小的阈值会增加误检,从而加重后续的数据处理。为此采用更多的类哈尔特征以达到更好的检测效果,例如同时使用图 2 中的两种特征,检测效果要优于只用 1 种类哈尔特征。如图 8 所示,利用多种类哈尔特征改善了检测结果,降低了误检率。

图 7　躯干检测结果

**Fig. 7　Results of limbs detection**

图 8　使用 2 种特征得到的改善效果

**Fig. 8　Improved results by using two features**

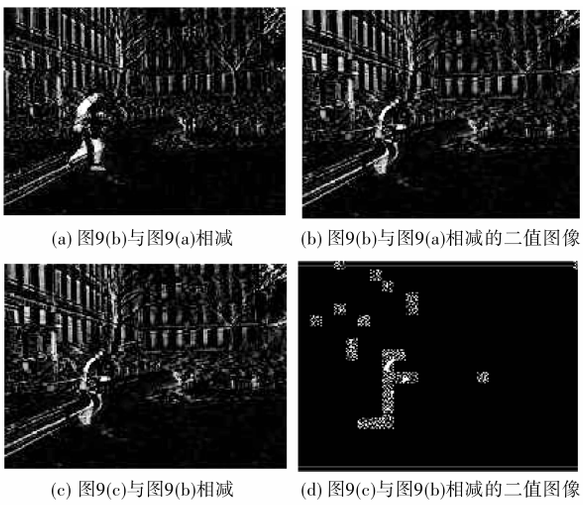
为了验证所提出方法检测的准确性与快速性,采用背景减除法对视频图像进行了肢体检测,图 9 为用背景减除法检测上臂的原始图像。

(a) 位置1　　(b) 位置2　　(c) 位置3

图 9　用背景减除法检测人体上臂的原始图像

**Fig. 9　Original images for detecting upper arm using background subtraction**

根据图 9 中 3 幅原始图像,利用图 9(b)图像与图 9(a)图像进行背景减除,图 9(c)图像与图 9(b)图像进行背景减除来检测人体上臂,检测结果如图 10 所示。



(a) 图9(b)与图9(a)相减　(b) 图9(b)与图9(a)相减的二值图像

(c) 图9(c)与图9(b)相减　(d) 图9(c)与图9(b)相减的二值图像

**图 10　用背景减除法检测人体上臂**

**Fig. 10　Detecting upper arm using background subtraction**

从图 10(a)可看出,利用背景减除法无法检测到人体上臂;从图 10(b)的二值图像可见,由于存在遮挡,人体上臂的像素由于相减原因已经为 0,无法在图像中定位人体肢体的位置。图 10(c)与图 10(d)也同样说明利用背景减除法无法检测到人体肢体。但是,由图 8 的实验可知,基于类哈尔特征与积分图像的肢体检测方法却可以快速检测得到人体各肢体。

## 4　结论

传统用于检测人体运动的方法如背景减除法等,因为鲁棒性差且相对比较费时,因此不适用于视频图像的人体检测系统。为此,将类哈尔特征的快速检测算法加以推广,同时引入了积分图像的概念,将两者结合用于在视频序列中检测运动人体的各种肢体。该方法将视频中人体肢体建模为矩形以便于利用类哈尔算子实现检测,经过对图像进行卷积之后,利用累加计算获得积分图像,然后利用与人体肢体尺度相同的类哈尔特征遍历积分图像,在积分图像中计算类哈尔特征的数值,并与预设阈值比较之后,获得人体肢体在图像中的位置,同时可以避免由于遮挡导致肢体无法检测的情况。

**参考文献:**


[1] 黄士科,陶琳,张天序. 一种改进的基于光流的运动目标检测方法[J]. 华中科技大学学报:自然科学版,2005,33 (5):39-41.
HUANG Shi-ke, TAO Lin, ZHANG Tian-xu. An improved algorithm of moving object detection based on optical flow[J]. Journal of Hua Zhong University of Science and Technology: Nature Science, 2005, 33(5):39-41. ( in Chinese with an English abstract)

[2] 王亮,胡卫明,谭铁牛. 人运动的视觉分析综述[J]. 计算机学报,2002,25 (3):225-237.
WANG Liang, HU We-iming, TAN Tie-niu. A survey of visual analysis of human motion[J]. Chinese Computers, 2002, 25(3): 225-237. (in Chinese with an English abstract)

[3] CHEUNG C S, KAMATH C. Robust techniques for background subtraction in urban traffic video[J]. SPIE, 2004, 5308: 881- 892.

[4] YU T, ZHANG C, COHEN M. Monocular video foreground /background segmentation by tracking spatial-color gaussian mixture models[C]. Austin: Proc of the IEEE Workshop on Motion and Vision Computing, 2007: 5-12.

[5] PAUL V, MICHAEL J. Robust real-time object detection[J]. Second international workshop on statistical and computational theories of vision-modeling, learning, computing, and sampling, 2001,7:1-25.

[6] PAUL V, MICHAEL J. Robust real-time face detection[J]. International Journal of Computer Vision, 2004,57(2) :137-154.