

中图分类号: TP301.4 文献标志码: A 文章编号: 1006-8961(2011)12-2217-06

论文索引信息: 何飞, 李亚利, 鹿文浩, 王生进, 丁晓青. 多方向线积分的梯度特征 [J]. 中国图象图形学报, 2011, 16(12): 2217-2222

多方向线积分的梯度特征

何飞, 李亚利, 鹿文浩, 王生进, 丁晓青

(清华大学电子工程系, 北京 100084)

摘要: 典型的梯度特征包括 HOG(梯度方向直方图)、Shapelet 及 Edgelet 等, 这些特征被广泛用于目标检测、目标识别、图像检索及场景分类等领域。针对 HOG 特征运算复杂度高的问题, 提出了一种新的多方向线积分的梯度特征(MDIG)。通过避免计算梯度方向并利用积分图, 该特征简化了计算过程, 提高了计算速度, 因而便于在 DSP 等硬件上实现。实验中新特征被应用于人体头肩检测。实验结果表明当使用 AdaBoost 算法训练分类器时, 该特征的描述能力与 HOG 相当, 同时其计算时间仅为 HOG 的 1/3, 整体性能优于 HOG。最后, 针对梯度特征的适用范围对其未来应用的发展方向进行了讨论。

关键词: 梯度特征; 梯度方向直方图; 多方向线积分; 头肩检测

Multi-directional integration of gradients

He Fei, Li Yali, Lu Wenhao, Wang Shengjin, Ding Xiaoqing

(Department of Electronic Engineering, Tsinghua University, Beijing 100084 China)

Abstract: Representative algorithms for extracting gradient features include HOG (histogram of oriented gradients), Shapelet and Edgelet. They have been widely used in the fields of object detection, object recognition, image retrieval, and scene classification. A new feature named Multi-Directional Integration of Gradients (MDIG) is proposed after analyzing the HOG and its fast version. By avoiding computing the gradient orientation and using an integral image, the new feature is easier to extract and can be hardware implemented on a DSP. The MDIG is tested on images of human head-shoulders. Using AdaBoost to train a classifier, the result is comparable with HOG, while the computation burden is only one third of that of the HOG. Finally, the use of gradient feature is discussed, and the future work is summarized.

Keywords: gradient feature; HOG; multi-directional integration; head-shoulder detection

0 引言

人体检测在安防监控、人机交互及智能交通等领域具有广泛的应用价值。这些应用系统往往需要首先检测到人体才能进行下一步操作。例如在安防监控中, 人体检测可以作为人流计数和人群密度估

计的基础; 而在人机交互中, 人体检测是姿势估计和行为识别的前提; 智能交通系统中的人体检测可以用于实现辅助驾驶等功能。这些系统中不仅要求人体检测的准确率要高, 虚警要低, 同时还要求算法复杂度低, 能够在小型的嵌入式系统中实现。目前的人体检测算法中, 性能较好的往往计算复杂, 并不能解决实际应用的需求。

收稿日期: 2011-05-17; 修回日期: 2011-09-06

基金项目: 国家重点基础研究发展计划(973)基金项目(2007CB311004); 国家自然科学基金项目(61071135); 教育部博士点基金项目(20090002110077)。

第一作者简介: 何飞(1987—), 男。清华大学电子工程系信息与通信工程在读博士研究生, 主要研究方向为图像处理和模式识别。E-mail: wgsgj@tsinghua.edu.cn。

针对这一瓶颈问题深入研究了相关算法,并在此基础上提出了一种新的多方向线积分梯度(MDIG)特征。为了尽可能全的保留图像的梯度信息,本文算法首先对梯度图像进行局部归一化,然后在多个方向上计算线积分,最后利用 AdaBoost^[1-2] 挑选弱特征并组成分类器。该特征计算过程简单,便于在 DSP 等硬件上实现。计算该特征所耗的计算时间仅仅是 HOG^[3] 的三分之一。由于 MDIG 特征有效地保留了图像的梯度信息,其描述能力接近甚至超过应用最广泛的 HOG。将 MDIG 特征应用到现有的系统中可以大幅降低系统的运算负荷,提高运算效率。

为了测试 MDIG 特征,选择人体头肩作为检测对象,基于的考虑主要包括以下 3 个方面:首先,相比于手臂、躯干等部位,头肩相对比较稳定;其次,在多数情况下我们都希望检测出带头部的人体,因而遮挡现象不如其他部位严重;第三,头肩可以作为人体的一个部件从而应用到基于部件的人体检测算法中。本文的训练样本和测试样本均选自 PASCAL,其中正样本 300 张,负样本(不含人体的样本)350 张。300 张正样本中包含 405 个人体头肩。为了扩大样本空间,这里对每个头肩样本进行平移、尺度和旋转的扰动,共生成 4 050 个头肩样本,其中 2 700 个用于训练,1 350 个用于测试。实验结果表明特征的描述能力与 HOG 相当,同时其计算时间仅为 HOG 的三分之一,整体性能优于 HOG。

1 梯度特征

在人体检测算法中,基于梯度的特征是应用最广泛的,包括文献[3-9]。HOG^[3] 特征即梯度方向直方图,最初被用于检测行人;文献[4]采用了 HOG 特征并将其应用到头肩检测,作者在 MIT^[10] 和 INRIA^[3] 两个库上进行了测试;文献[5]是文献[4]的拓展,其将头肩检测的结果作为行人跟踪的输入;文献[6]中,HOG 特征被用于描述人体部件,进而检测多姿态人体。文献[7]提出的 Shapelet 与 HOG 不太相同,它包括两层特征:底层特征来源于图像各个方向的梯度响应,中层特征则是相邻的底层特征的组合。Edgelet^[8] 又被称做边缘片段,也属于梯度特征,它用各种形状的边缘片段去描述图像局部的梯度响应,然后通过 AdaBoost 挑选最有效的弱特征组成分类器;文献[9]用 Edgelet 的方法检测人体头肩

并将其应用到虚拟会议中。文献[11]对比了 HOG 和 Shapelet 在行人检测中的效果,在“虚警/图像”的评价指标下,HOG 性能更优,因此本文选择 HOG 进行详细分析。

1.1 HOG

根据文献[3],HOG 特征的计算步骤主要包括以下一些特点:

- 1) 将图像划分成紧密排列的方格,相邻的方格组成一个块;
- 2) 相邻的块有重叠,可能共享几个方格;
- 3) 在方格中统计梯度方向直方图,根据梯度强度和像素与块中心的距离进行加权;
- 4) 每个像素根据其位置和梯度方向,在统计直方图时需要进行三线性插值;
- 5) 在块的范围内归一化直方图的 2-范数。

由于计算 HOG 特征时需要计算梯度方向,且要根据方向的值在直方图中进行线性插值,因而必须计算反正切函数。这样导致 HOG 特征的计算量非常大,且不容易在简单的硬件上实现。

1.2 HOG 的快速算法

通常牺牲性能往往能够减少计算量。文献[12]针对 HOG 特征提出了一种快速算法,相比于原始的 HOG 特征,快速算法具有如下特点:

- 1) 取消像素与块中心的距离权重,无论像素离块中心的距离是多少,其权重仅与像素的梯度强度有关;
- 2) 取消三线性插值,对梯度方向进行离散化,这样可以通过简单计算得到梯度方向;
- 3) 用 1-范数归一化代替 2-范数归一化,通过积分图加快计算速度。

通过这些简化,HOG 的快速算法避免了直接计算反正切函数,同时去掉了大部分浮点数乘法,这样使得计算时间大大缩短。

2 多方向线积分的梯度特征

2.1 计算步骤

如图 1 所示,MDIG 特征的计算方法步骤如下:

- 1) 将图像划分成紧密排列的块,相邻块可以有重叠;
- 2) 计算每一块图像的梯度强度;
- 3) 在块的范围内对梯度强度进行归一化;
- 4) 在多个方向上进行线积分,用积分结果作为

特征向量。

2.2 与 HOG 的对比

之所以 HOG 特征能够取得较好的效果,其原因在于它尽可能全的保留了图像的梯度信息。它通过梯度方向直方图的形式将梯度方向量化为若干个区间,保留了梯度方向的信息;又根据梯度强度计算各个梯度方向区间累加的权值,这样 HOG 又保留了梯度强度的信息;此外,通过对不同坐标的像素赋予不同的权重,HOG 特征还保留了梯度信息中的位置信息。

在 MDIG 特征中,我们舍去了梯度方向的信息,仅保留了梯度强度和位置信息。MDIG 特征中的线积分与 HOG 特征中各个梯度方向区间内梯度强度的累加作用相似,它保留了梯度的强度信息;在多个方向上进行线积分则最大限度的保留了梯度的位置信息,这种方式保留的位置信息比 HOG 中的更完全。事实上,当梯度的强度和位置确定以后,梯度的方向显得并不重要。例如已知在某一水平横线附近有较强的梯度,那么通常情况下这一区域的梯度方向都接近垂直。

除此以外,HOG 特征的一个重要措施是将图像划分成相互重叠的块,然后在块的范围内进行归一化。重叠的块能够保证每一个像素出现在不同的块中,这样每一个像素的梯度强度能够根据不同的邻域进行归一化,提高了特征的稳定性。MDIG 特征也采用了这样的方式。

综上所述,MDIG 特征用多个方向的线积分充分保留了梯度的位置信息,从而省略了计算梯度方向信息的步骤,进而简化了算法,加快了计算速度。

2.3 具体实现

2.3.1 块的划分

块的划分与 HOG 非常相似,如果定义块的大小为 S ,相邻块的距离为 t ,则当 $t < S$ 时相邻块重叠的区域面积为 $S(S - t)$ 。实验中 $S = 8, t = 4$ 。

2.3.2 梯度强度

文献[3]建议使用形如 $[-1, 0, 1]$ 的差分,即

$$\begin{cases} d_x(x, y) = I(x + 1, y) - I(x - 1, y) \\ d_y(x, y) = I(x, y + 1) - I(x, y - 1) \end{cases}$$

本文也采用了这样的方法。严格来讲,梯度强度的定义应该是:

$$m = \sqrt{d_x^2 + d_y^2}$$

但是开方运算比较耗费时间,在实验中同时采用了

m 和 m^2 两种形式并对比了其结果。

2.3.3 归一化

归一化主要有两种方式:1-范数归一化和 2-范数归一化。由于 MIDG 特征其归一化在计算线积分之前,因而无论采取哪种方式均可用积分图加快计算速度。这一点是 MDIG 特征优于 HOG 及其快速算法的一个方面。在 HOG 中,由于归一化是在累加得到直方图之后进行的,因而无法通过积分图这样的方式加快速度;HOG 快速算法利用了积分图,其代价是只能进行 1-范数归一化,而文献[3]中的实验表明,2-范数归一化效果优于 1-范数归一化。对于梯度强度小于 1 的块,实验时不进行归一化。

2.3.4 线积分的方向

如图 1 所示,具体实现时多方向线积分可以简化为 4 方向线积分,即 0° 、 90° 、斜向上 45° 和 135° 4 个方向。

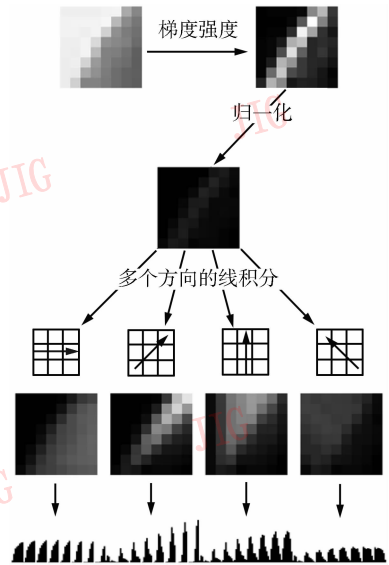


图 1 MDIG 特征的计算方法

Fig. 1 Steps for calculating MDIG

3 实验结果

实验样本的具体情况见引言。实验时所有图像均被归一化到 32×32 像素,这样一来每个样本将被划分为 49 个 8×8 像素的块(块与块之间的距离为 4 像素)。对于 HOG 而言,每个块分为 4 个方格,每个方格统计一个梯度方向直方图。直方图将 0 到 360 度划分为 8 个区间,每个块得到一个 $8 \times 4 = 32$ 维的特征向量。HOG 特征的总维数为 $32 \times 49 =$

1 568。MDIG 特征在每个块积分之后得到 4 幅线积分图,每幅图拉成一个特征向量后的维数是 $8 \times 8 = 64$,因而每个块得到一个 $64 \times 4 = 256$ 维的特征向量,总维数是 $256 \times 49 = 12\ 544$ 。

3.1 计算速度

测试计算速度时在同一环境下运行所有程序,测试对象是从样本中随机抽取的 100 000 个 32×32 像素的子图像。为了排除读入测试样本所需的时间,所有样本先读入内存中再一并提取特征。测试得到的时间如图 2 所示。图中横轴是各种测试方法,其中测试 1、2 采用 2-范数归一化,测试 3、4 采用 1-范数归一化,测试 1、3 对梯度强度进行积分,测试 2、4 对梯度强度的平方进行积分;纵轴是提取特征占用的时间,单位为秒。每组测试进行了 10 次取平均值。每一组柱状图左边是计算 HOG 特征的时间,右边是计算 MDIG 特征的时间。

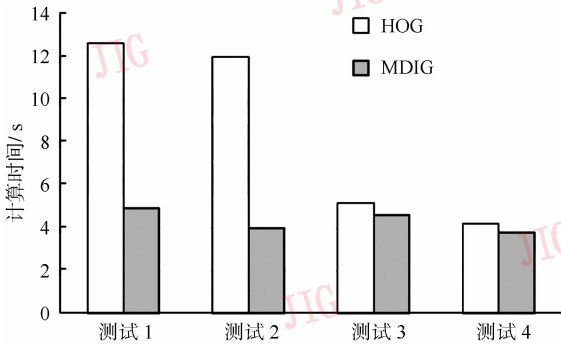


图 2 计算时间测试结果

Fig. 2 Comparisons of computing time

从图 2 中可以看出,计算 MDIG 特征所需的时间远小于计算 HOG 特征所需的时间,即便是利用快速算法简化 HOG 计算,其所耗时间也比 MDIG 特征长。图 3 更直观地显示了 MDIG 特征计算时间占 HOG 特征的百分比。在未简化的时候,MDIG 特征计算时间仅占 HOG 特征的不到 40%;而对于简化

后的 HOG 特征,MDIG 特征也能提高约 10% 的速度。归纳起来,MDIG 特征在计算速度上具有很大优势,原因在于 MDIG 特征具有以下优点:

- 1) 彻底避开了梯度方向的计算;
- 2) 采用先归一化后累积的方式,可以利用积分图加快速度;
- 3) 尽量避免浮点数乘法。

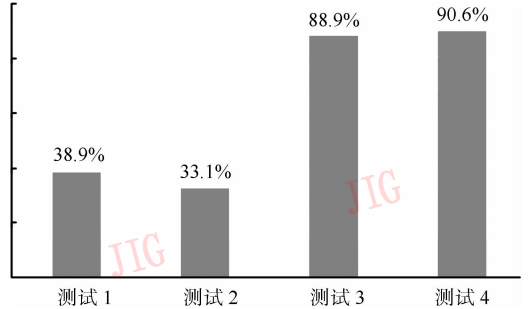


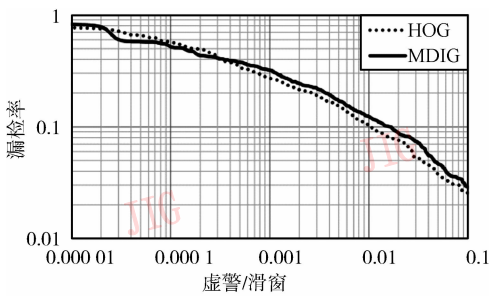
图 3 MDIG 特征计算时间占 HOG 特征的百分比

Fig. 3 Percentages of computing time

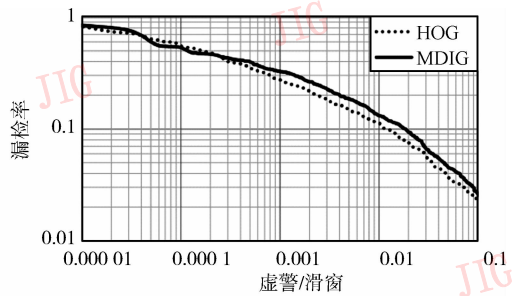
3.2 分类效果

分类器训练采用 Discrete AdaBoost 的方法,共选取占总量八分之一的弱特征组合成最终的分类器(没有级联)。测试样本包括 1 350 个正样本和 100 000 个负样本,负样本均从不包含人体的图像中随机抽取。最终得到的检测-错误权衡图如图 4 和图 5 所示。

图 4 和图 5 的两幅图分别代表利用梯度强度 m 和梯度强度平方 m^2 作为累加对象时各特征的分类效果。事实上,利用梯度强度和利用梯度强度的平方得到的结果差距很小,而计算梯度强度 m 时需要进行开方,一次开方运算占用的时间又远大于一次加减法占用的时间。从图 2 中的数据也能看出,用梯度强度的平方进行累加时可节约 10%~20% 的计算时间。因此在实际应用中可以就用梯度强度的平方。



(a) m 作为累加对象



(b) m^2 作为累加对象

图 4 MDIG 特征与 HOG 的分类效果对比

Fig. 4 Classification results of MDIG and HOG

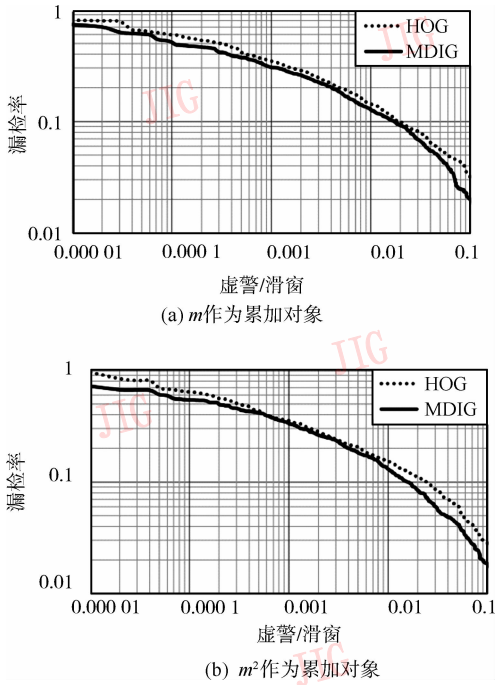


图 5 MDIG 特征与 HOG 快速算法的分类效果对比
Fig. 5 Classification results of MDIG and fast HOG

图 4 对比的是 MDIG 特征和 HOG 特征的分类效果。从图中可以看出,MDIG 特征与 HOG 特征的分类效果相当。正如前文分析的一样,MDIG 特征更加充分地利用了梯度的位置信息,达到了与 HOG 利用梯度方向信息接近的效果。总的来说,MDIG 特征其分类效果与 HOG 处于同一水平。

图 5 对比的是 MDIG 特征和 HOG 快速算法的分类效果。从图中可以看出,MDIG 特征与 HOG 快速算法的性能也非常接近,但是在虚警/滑窗非常小的时候 MDIG 特征的分类效果要优于 HOG 快速算法。一种可能的解释是,HOG 快速算法在取消距离权重以及三线性插值后,梯度的位置信息没有得到有效利用,导致性能下滑。这也进一步说明充分利用梯度的位置信息比利用其方向信息更有效。

图 6 展示了一些人体头肩检测结果,里面包括了较好的结果以及虚警、漏检等情况。

4 结 论

针对 HOG 计算速度慢这一瓶颈问题,在其基础上提出了一种新的特征计算方法:多方向线积分的



图 6 头肩检测结果
Fig. 6 Results of head-shoulder detection

梯度特征。该特征计算过程简单,便于在 DSP 等硬件上实现,同时其计算时间仅为 HOG 的三分之一。在人体头肩样本分类问题测试中,MDIG 特征达到了与 HOG 相当的效果。

尽管当前的特征已经充分利用了图像的梯度信息,但并不是所有的分类问题其梯度特征都很明显。仅仅利用梯度信息并不能解决所有的分类问题。文献[11]中将梯度信息(HOG)和区域信息(Haar)结合起来检测行人得到了比 HOG 和 Haar 都要好的结果,这也说明通过合适的方法融合多种信息能比单独利用一种信息更好,对于人体检测,多信息融合是未来的发展趋势,也是今后的研究重点。

参考文献 (References)

- [1] Freund Yoav, Schapire Robert E. A decision-theoretic generalization of on-line learning and an application to boosting [J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [2] Friedman Jerome, Hastie Trevor, Tibshirani Robert. Additive logistic regression: A statistical view of boosting [J]. The Annals of Statistics, 2000, 28(2): 337-407.
- [3] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA: IEEE, 2005: 886-893.
- [4] Li M, Zhang Z, Huang K, et al. Estimating the number of people in crowded scenes by mid based foreground segmentation and head-shoulder detection [C]//Proceedings of International Conference on Pattern Recognition. Tampa, FL: IEEE, 2008: 1-4.
- [5] Li M, Zhang Z, Huang K, et al. Rapid and robust human detection and tracking based on omega-shape features [C]//Proceedings of IEEE International Conference on Image Processing. Cairo: IEEE, 2009: 2545-2548.
- [6] Lin Z, Davis L. Shape-based human detection and segmentation via hierarchical part-template matching [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(4): 604-618.
- [7] Sabzmeydani P, Mori G. Detecting pedestrians by learning shapelet features [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN: IEEE, 2007: 1-8.
- [8] Wu B, Nevatia R. Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors [C]//Proceedings of IEEE International Conference on Computer Vision. Beijing: IEEE, 2005: 90-97.
- [9] Wu B, Nevatia R. Tracking of multiple humans in meetings [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops. New York, NY: IEEE, 2006: 143-143.
- [10] Papageorgiou C, Poggio T. A trainable system for object detection [J]. International Journal of Computer Vision, 2000, 38(1): 15-33.
- [11] Dollar Piotr, Wojek Christian, Schiele Bernt, et al. Pedestrian detection: A benchmark [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL: IEEE, 2009: 304-311.
- [12] Zhu Qiang, Avidan Shai, Yeh Mei Chen, et al. Fast human detection using a cascade of histograms of oriented gradients [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York, NY: IEEE, 2006: 1491-1498.