

融合颜色与纹理的复杂场景下的服装图像分割算法

郭鑫鹏¹ 黄元元¹ 胡作进²

(南京航空航天大学计算机科学与技术学院 南京 210016)¹

(南京特殊教育师范学院数学与信息科学学院 南京 210038)²

摘要 针对复杂场景下拍摄到的服装图像的分割问题,提出一种基于先验知识的融合颜色和纹理特征的无监督分割算法。首先利用块截断编码思想将传统的三维颜色空间截断成为六维空间,得到更为精细的颜色特征,并结合改进的局部二值模式纹理特征实现对图像的特征描述;然后根据目标区域和背景区域在图像中出现的统计规律,提出了一种基于先验知识的两分法来对图像进行分割。由于对图像做了分块处理,因此在子图像块的基础上进行的图像分割将更加高效。实验表明,设计的算法能快速有效地将目标区域从各类不同的复杂场景中分割出来,且整个过程无须人工设定任何参数,对后续的图像理解和图像检索具有重要意义。

关键词 图像分割,块截断编码,纹理特征,先验知识

中图分类号 TP391 **文献标识码** A

Unsupervised Complex-scene Clothing Image Segmentation Algorithm Based on Color and Texture Features

GUO Xin-peng¹ HUANG Yuan-yuan¹ HU Zuo-jin²

(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)¹

(School of Mathematics and Information Science, Nanjing Normal University of Special Education, Nanjing 210038, China)²

Abstract In this paper, an unsupervised segmentation algorithm especially used for clothing images in complex background was proposed. Based on priori knowledge, it combines color and texture features together. First, block truncation encoding thought is used to cut the traditional 3-dimensional color space into 6-dimensional space, so that fine color features can be obtained. Then, the texture feature based on the improved local binary pattern (LBP) algorithm was designed and used to describe the image together with the color feature. After that, according to the statistical appearance-law of the target region and background in the image, called priori knowledge, a kind of bisect method was proposed to do segmentation. Since the image is divided into several sub image blocks, such bisect segmentation will be accomplished more efficient. Experiments show that the algorithm can quickly and effectively extract clothing region from the complex scene without any artificial parameters. This segmentation will play an important role for image understanding and retrieval.

Keywords Image segmentation, Block truncation code, Textural features, Heuristic knowledge

1 引言

图像分割可以将图像中的感兴趣区域从背景中分离出来,本文专门针对复杂场景下的服装图像进行分割处理。服装图像有其自身固有的特点,比如在视觉上,服装本身的颜色和纹理会与背景产生反差,从而给人带来视觉冲击。根据这一特点,国内外的研究人员做了大量的研究。文献[1]在基于图像颜色特征的基础上,提出了一种混合高斯模型的服装图像分割方法,但是此方法只考虑到服装图像的颜色特征而没有考虑图像中客观存在的纹理特征;文献[2]中的 MeanShift 算法将原始待分割图像过分割为具有一致性图像特征的区域集合,再通过人机交互手工大致标记的方法进行分割,取得了不错的效果。但该算法所用特征的维度高达 4096 维,复杂度高,且需要人机交互才能取得最终的分割结果;文献[3]设计

的算法是将 JSeg 算法应用于图像显著性区域的分割算法,利用人脸检测来辅助判别背景区域与人体着装区域,对服装进行分割。文献[4]在服装图像分割时,运用二次聚类的思想取得了不错的结果,其主要思想是首先利用自适应的 Mean Shift 算法,得到区域数目 n 以及 n 个区域的类心向量,再用 K 均值算法进行二次聚类,使得分割结果更加稳定。文献[5]设计的算法是采用直方图阈值法并结合改进的 Fuzzy C-means 聚类算法对图像进行分割,取得了不错的结果。文献[6]设计的算法是在分割过程中融入待分割目标的形状先验知识,把先验信息转化为能量函数,提出了自适应形状先验方法,其主要思想是目标的边缘信息越少,越需要更强的形状先验信息,该方法对灰度图像的分割效果比较好。文献[7]提出了一种改进的基于区域合并的纹理图像分割方法,将 Lab 颜色特征和 Gabor 能量纹理特征进行融合来对图像特征进行描

本文受扬州市“绿杨金凤”优秀博士项目,江苏省“双创”项目资助。

郭鑫鹏(1992—),男,硕士生,主要研究方向为模式识别、图像处理, E-mail: 805861040@qq.com; 黄元元(1975—),女,博士,副教授,主要研究方向为多媒体技术、图像处理、模式识别; 胡作进(1965—),男,博士,教授,主要研究方向为数据处理、机器学习。

述,然后通过人机交互的方式进行图像分割,此算法仍然需要人工参与才能完成分割。文献[8]设计的算法是对文献[2]算法的改进和优化,文献[8]用 LBP 纹理特征代替高达 4096 维的特征,并融合颜色特征对目标图像进行特征描述,然后进行分割,提高了算法的执行效率。文献[9]针对传统活动轮廓分割算法对纹理丰富的图像存在过分割和欠分割现象,分析了各向同性和异性扩散的机理,设计了保边纹理模糊函数,以图像梯度为变量建立了全变分图像纹理模糊模型,取得了不错的实验效果。文献[10]在改进图像分割算法的基础上将颜色特征和纹理特征结合起来,构造局部自适应的正则化参数,较好地避免了前景和背景颜色相似时图像分割算法容易产生分割错误并导致 shrinking bias 的现象,虽然这类改进算法取得了不错的效果同时也降低了算法的复杂度,但需人工参与才能最终完成图像分割。

本文提出一种基于先验知识的融合颜色和纹理特征的无监督服装图像模糊分割算法。该算法可以实现对复杂场景下的服装图像的自适应分割,无须人工设定任何参数。

2 算法设计

服装本身的颜色与纹理特征使其与背景产生反差,因此,为了准确描述图像的内容,应该将颜色与纹理特征相融合。本文首先利用 BTC 块截断思想提取 6 维的颜色特征,并优化 LBP 算法,得到适合与 6 维的颜色特征相融合的纹理特征。对图像做分割实际上就是对图像的像素点做分类,即哪些像素点属于目标区域,哪些像素点属于背景区域。但是若直接在像素级上做分割,首先由于图像包含的像素点数量极大,会导致分割算法的时间复杂度极高;其次,单一的像素点很难体现服装自身的颜色以及纹理特点。因此考虑对图像进行分块处理。对每一个子图像块分别提取上述颜色与纹理特征,并对子图像块做聚类,将其划分为目标区域和背景区域两个类别,从而实现图像的分割。在聚类之前,可以根据一些先验知识确定两类的聚类中心,避免聚类陷入局部最优。

2.1 基于块截断思想的颜色特征

由于常用的 RGB 颜色模型是一种面向硬件的颜色模型,并不符合人类视觉对颜色差异的感知程度^[3],因此需要将 RGB 颜色模型转换成一种均匀的、适合聚类的颜色模型,故本文选用 Lab 颜色模型^[11]。

块截断编码(Block Truncation Coding, BTC)作为一种经典的压缩方法,最初应用于图像的有损压缩。此算法先将图像分割成适当大小的分块,然后在保持图像块中像素的平均值和标准方差的同时,减少图像的灰度值的数量,以此来达到图像压缩的目的^[12-13],本文正是利用这一思想来获取图像块的颜色特征的。传统的 Lab 颜色模型只能得到 3 维的颜色特征,本文利用 BTC 算法可以获取 6 维的颜色特征,对颜色的描述将更加准确。假定有一个 $M \times N$ 的图像块,其第 i 行第 j 列处的像素点 X_{ij} 可以看作是一个包含 L, a, b 3 个分量的特征向量,这 3 个分量可以分别表示为 $X_{ij}(L), X_{ij}(a)$ 和 $X_{ij}(b)$ 。计算该图像块颜色特征的算法步骤如下。

1) 计算图像块的 L, a, b 3 个分量的均值以及标准差,记为 $\bar{L}, \bar{a}, \bar{b}$, 以及 $\sigma_L, \sigma_a, \sigma_b$, 其中:

$$\bar{L} = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N X_{ij}(L) \quad (1)$$

$$\sigma_L = \frac{1}{M \times N} \sqrt{\sum_{i=1}^M \sum_{j=1}^N (X_{ij}(L) - \bar{L})^2} \quad (2)$$

同理计算 \bar{a} 和 \bar{b} 以及 σ_a 和 σ_b 。

2) 对图像块的 L, a, b 3 个分量做截断处理,现以 L 分量为例。将 \bar{L} 看作阈值,与图像块中所有像素点的 L 值进行比较,小于阈值的设定为 L_{low} ,反之设置为 L_{high} :

$$L_{low} = \bar{L} - \sigma_L \sqrt{\frac{p}{M \times N - p}} \quad (3)$$

$$L_{high} = \bar{L} + \sigma_L \sqrt{\frac{M \times N - p}{p}} \quad (4)$$

其中, p 为图像块中 L 分量大于或等于阈值 \bar{L} 的像素点个数。

3) 对图像块中所有像素的 a 和 b 分量也做上述的截断处理,从而得到图像块的 6 维颜色特征 $L_{low}, L_{high}, a_{low}, a_{high}, b_{low}$ 和 b_{high} 。

2.2 改进的 LBP 纹理特征

局部二值模式(Local Binary Pattern, LBP)是一种用来描述图像局部纹理特征的算子^[14]。在像素点的 3×3 邻域内,从左上角的像素开始,按顺时针方向进行,若邻域点的像素值大于或等于中心像素,则记为 1,否则记为 0,从而得到一个 8 位的二进制数^[15-16],如图 1(a)所示。在以 80 为中心的 3×3 邻域内,中心像素点的 LBP 特征即为 11001001,转换为十进制则是 201,这个数值可以反映出像素与其邻域像素的关系。这样,图像块中的所有像素点都会对应一个十进制的 LBP 纹理特征值,统计这些特征值,以该特征值的直方图作为图像块的纹理特征。但是将该纹理特征与前述的颜色特征直接进行融合并不合适,因为由此得到的直方图是 256 维,而之前得到的颜色特征只有 6 维。由于二者的特征维数相差甚远,会导致在图像块的特征描述中,纹理特征会占据主导地位,颜色特征的影响将会非常有限。针对这一问题,本文对原始 LBP 纹理特征提取算法进行了改进,在不影响图像纹理轮廓的前提下实现纹理直方图的降维,如图 1(b)所示。在以 80 为中心的 4 邻域范围内统计中心像素的 LBP 特征,为 1010,转换为十进制为 10。由于考察的邻域像素点数量减少,其二进制模式的位数也随之减少,对应的十进制数仅仅在 0~15 之间,因此将纹理特征从 256 维降低至 16 维。

99	215	33
102	80	79
77	21	101

(a)原始的 LBP 特征图

99	215	33
102	80	79
77	21	101

(b)改进后的 LBP 特征

图 1 纹理特征提取

2.3 图像分割

将原始图像做分块处理,并对子图像块提取颜色和纹理特征后,对图像进行分割,这实际上是对图像块做分类处理,即将子图像块划分为目标区域和背景区域两类。由于类别数已知,因此理论上可以采用 K-means 算法进行聚类。但是, K-means 算法对初始的聚类中心比较敏感,若初始聚类中心选择不当,可能会导致算法收敛到局部最优解。通过浏览分析大量的服装图片,可以发现一个统计规律,如图 2 所示:假定对图像做 10×10 的分割,由于服装是图像的目标区域,因此在绝大多数的图像中,服装区域都会出现在图像比较靠中间的位置,而图中标记出来的几个图像块基本都属于背景区

域。因此,本文提出一种基于先验知识的两分法。



图2 图像分块

如图2所示,若将每个图像分成 10×10 的网格,根据先验知识,图中粗线标记出背景区域的8个图像块属于背景区域的概率非常大。假定一幅图像 S 包含 N 个子图像块,即 $S = \{s_1, s_2, \dots, s_N\}$,每一个子图像块 s_i 经过上述特征提取后可以看作一个22维的特征向量。这些子图像块分属于目标区域 G_1 和背景区域 G_2 这两个类别。初始情况下,图2中所标记的8个子图像块属于 G_2 ,即 $G_2 = \{s_1, s_2, \dots, s_7, s_8\}$,剩余所有子图像块皆属于 G_1 ,即 $G_1 = \{s_9, s_{10}, \dots, s_N\}$ 。接下来需要将 G_1 中原本属于 G_2 的子图像块逐一挑选出来放入 G_2 ,最终得到的 G_1 中所有子图像块所组成的最大连通域即是目标区域。算法的具体步骤如下。

1)在 G_1 中挑选图像块放入 G_2 的过程应该遵守分类最优的原则,即定义目标函数:

$$E = \frac{1}{n_1 n_2} \sum_{i \in G_1, j \in G_2} \|s_i - s_j\| \quad (5)$$

其中, n_1 和 n_2 表示 G_1 和 G_2 中的图像块个数。目标函数表示 G_1 与 G_2 两类之间的类平均距离,其值越大,表明分类效果越好。

2)对 G_1 中的所有图像块做穷举,即分别计算把 s_9, s_{10}, \dots, s_N 并入 G_2 类时的目标函数,若当 s_i 并入 G_2 类时的目标函数值最大,则将 s_i 从 G_1 中取出放入 G_2 ,记录下此时目标函数为 $E(1)$,表示这是第一次做选取得到的目标函数。此时 $G_1 = \{s_j | j \neq i, \text{且 } 9 \leq j \leq N\}, G_2 = \{s_1, s_2, \dots, s_8, s_i\}$ 。

3)重复步骤2),得到最大的目标函数为 $E(2)$,对应图像块记为 s_p 。比较目标函数 $E(1)$ 与 $E(2)$ 。若 $E(2) \geq E(1)$,则将 s_p 并入到 G_2 中,并继续重复上述步骤,直到当前的目标函数 $E(k+1)$ 小于前一次得到的目标函数 $E(k)$,则分类结束,以第 k 次得到的分类结果作为最终的分类结果。

4)此时, G_1 中包含 $N-8-k$ 个子图像块, G_2 中包含 $8+k$ 个子图像块。将 G_1 中包含的子图像块所构成的最大连通域作为最终的目标区域,图像分割过程结束。

上述算法依据统计规律,利用了一定的先验知识,从而有效避免了K-means算法选择初始聚类中心的难题,同时,目标函数的设计使得每次的分类效果可以达到最优,而先验知识的使用可以保证算法不会陷入局部最优。此外,虽然使用穷举法进行分类,但是由于图像块的数量有限,因此算法的执行效率仍然很高。

3 实验与分析

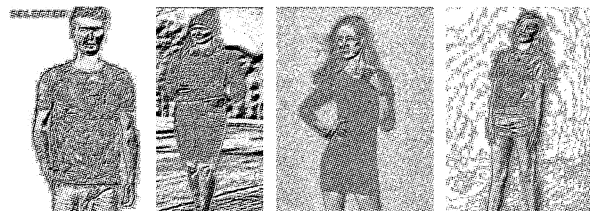
为了验证本文算法的有效性,对2000幅不同场景下的服装图像进行测试。

首先对改进的LBP纹理特征进行测试,如图3所示。图3(a)是原始图像,图3(b)是传统的LBP算法进行特征提取后的LBP值纹理图像,图3(c)是改进后的纹理图像。从图中可以看出,虽然纹理特征从256维降到了16维,但是其仍然能

够很好地保留原始图像的纹理特征。



(a)原始图像



(b)传统的LBP纹理特征图



(c)改进的LBP纹理特征图

图3 LBP纹理特征示意图

接下来,为了验证特征融合以及本文提出的两分法的有效性,设计了5组实验方案。

实验1 仅利用6维的颜色特征,并采用本文的两分法进行图像分割。

实验2 仅利用16维的纹理特征,并采用本文的两分法进行图像分割。

实验3 将传统的Lab三维颜色特征与16维的纹理特征相结合,并采用本文的两分法进行图像分割。

实验4 将本文提出的6维颜色特征与16维的纹理特征相结合,并采用本文的两分法进行图像分割。

实验5 将本文提出的6维颜色特征与16维的纹理特征相结合,并采用K-means算法进行图像分割。

实验结果如图4所示。由图4的实验结果可以发现:

1)实验1与实验2由于仅采用单一的颜色特征或者纹理特征,因此对于背景复杂的图像,或者服装本身具有较复杂的颜色变化或者纹理图案时,分割的效果会受到较大影响,如图4(b)和图4(c)所示。

2)如图4(d)和图4(e)所示,在将颜色与纹理特征结合后,在颜色和纹理特征的相互抑制下,相比于实验1和实验2,实验3和实验4的实验分割结果整体上有了明显的提升。由实验结果可以看出将本文提出的6维截断Lab颜色特征与16维纹理特征相结合并进行图像分割时,有更好的分割结果。

3)从实验5与实验4的对比可以看出,由于K-means算法使用随机的方式来确定聚类中心,导致分割结果不稳定,把背景区域误判为目标区域的情况时有发生。针对K-means算法分割结果不稳定的问题,一种解决方法是进行多次实验,然后用统计的方式确定初始聚类中心,但这无疑增加了分割算法的时间复杂度。

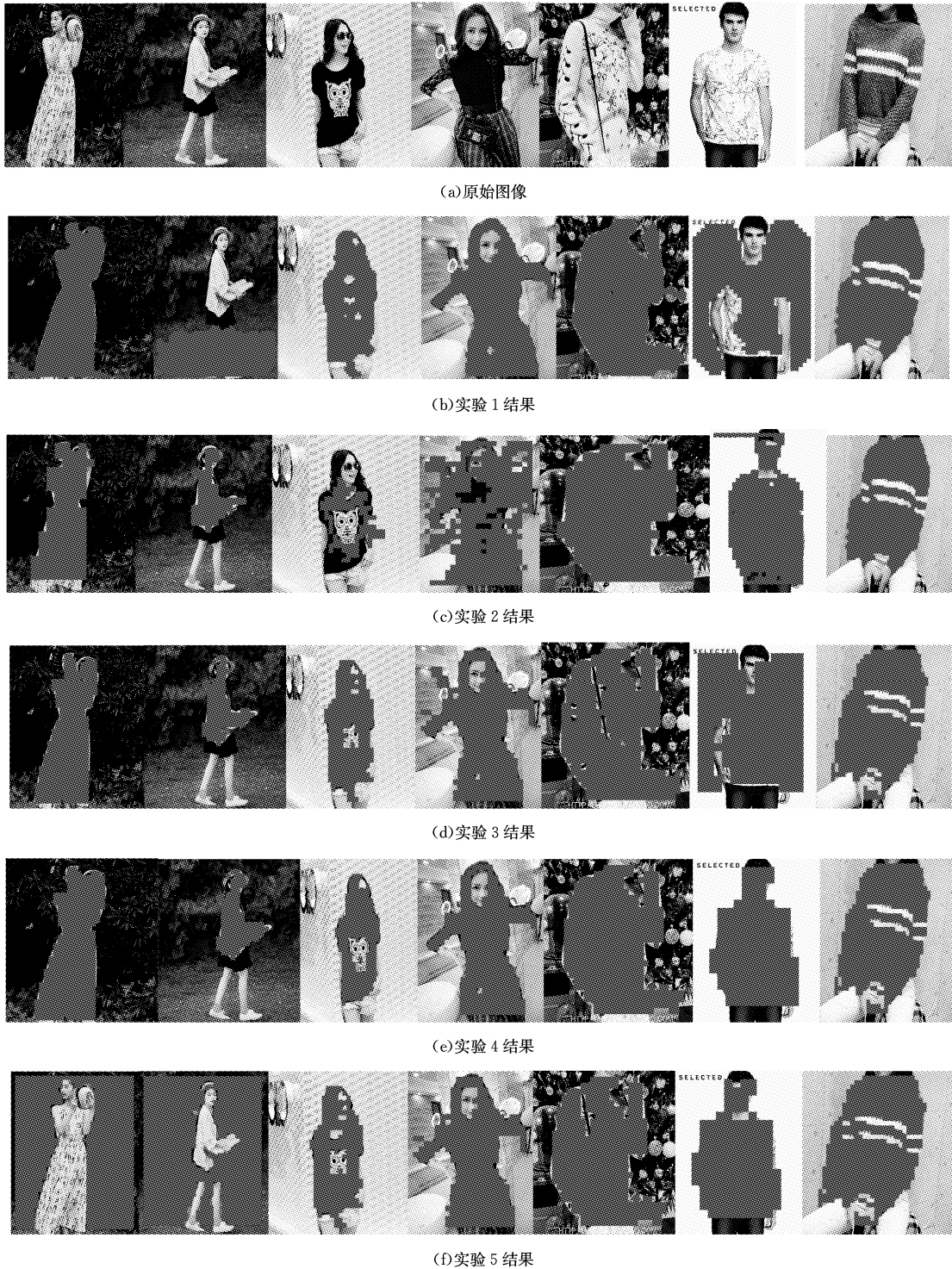


图 4 图像分割测试结果

结束语 本文提出一种针对服装图像的融合颜色和纹理特征的无监督分割算法,可以在复杂场景中将服装的主要部分提取出来。虽然该方法融合了颜色与纹理特征,但是特征维数并不高,而且对图像做了分块处理,使得分类算法可以快速实现。同时,由于利用了基于统计规律的先验知识进行分类,因而在一定程度上保证了分类算法的有效性。实验结果证明,在各类复杂场景下,利用本文提出的算法均可得到较为满意的分割效果,可以为后续的图像理解、图像检索等做好准备。

参 考 文 献

[1] XIANG M,XIE Y,XIANG L. Research on Clothing Image

Segmentation Based on Color Features[J]. Progress in Textile Science & Technology,2013(1):35-37.
 [2] COMANICIU D,MEER P. Mean Shift:A Robust Approach Toward Feature Space Analysis[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2002,24(5):603-619.
 [3] LIU Z. A Garment Image Segmentation Method Based on Saliient Region and JSEG[J]. Journal of Software,2015,10(11):1274-1282.
 [4] XIE Y. Segmentation method combined with mean shift and K-mean clustering algorithm for clothing image [J]. Electronic Measurement Technology,2013,36(8):53-55,60.
 [5] TAN K S,ISA N A M. Color image segmentation using histo-

- gram thresholding-Fuzzy C-means hybrid approach[J]. Pattern Recognition, 2011, 44(1): 1-15.
- [6] WANG H, ZHANG H, RAY N. Adaptive shape prior in graph cut image segmentation[J]. Pattern Recognition, 2013, 46(5): 1409-1414.
- [7] 郑庆庆, 桑农, 高常鑫, 等. 改进的基于区域合并的纹理图像分割方法[J]. 华中科技大学学报(自然科学版), 2011(5): 109-112.
- [8] 张竹伟, 周海英. 基于区域色彩空间和 LBP 算法的图像分割[J]. 微电子学与计算机, 2015(2): 84-87.
- [9] 何坤, 郑秀清, 张永来. 纹理模糊的图像分割[J]. 四川大学学报(工程科学版), 2015, 47(4): 111-117.
- [10] 刘毅, 黄兵, 孙怀江, 等. 结合 CS-LBP 纹理特征的快速图像分割算法[J]. 计算机科学, 2013, 40(5): 300-302.
- [11] 汤顺青. 色度学[M]. 北京: 北京理工大学出版社, 1990.
- [12] SUN Y, XU R, CHEN L, et al. Image compression and encryption scheme using fractal dictionary and Julia set[J]. Iet Image Processing, 2015, 9(3): 173-183.
- [13] DELP E J, MITCHELL O R. Image Compression Using Block Truncation Coding[J]. IEEE Transactions on Communications, 1979, 27(9): 1335-1342.
- [14] OJALA T, PIETIK, INEN M, et al. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(7): 971-987.
- [15] LIU L, LAO S, FIEGUTH P W, et al. Median Robust Extended Local Binary Pattern for Texture Classification[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2016, 25(3): 1368-1381.
- [16] TANG Z, SU Y, MENG J E, et al. A local binary pattern based texture descriptors for classification of tealeaves[J]. Neurocomputing, 2015, 68(C): 1011-1023.

(上接第 201 页)

5.3 实验结果及分析

本文的实验结果是对 DC-ped-dataset 与 CVC 标准数据集测试得到的, 本文在实验过程中将正、负测试集样本进行标号排序, 并设定测试集中正样本标签为 1, 负样本标签为 0。在实验过程中算法若将正样本识别为 0 或者将负样本识别为 1 皆为误识, 则识别率 = 正确识别数/测试集总数。本文在 CVC 标准测试集中选取 3000 幅行人图片作为训练正样本, 选取 3400 幅非行人图片作为训练负样本。选取正、负测试集样本各 600 幅, 分别采用经典 CNN 行人检测算法、经典 HOG+SVM 算法与本文算法进行实验。实验结果如表 3 所列, 本文算法的识别率为 97.825%, 优于经典 CNN 算法(96.937%)与传统 HOG+SVM 算法(96.0125%)。在 DC-ped-dataset 标准测试集中选取 3000 幅行人图片作为训练正样本, 4000 幅非行人图片作为训练负样本。选取正、负测试集样本各 1000 幅。实验结果如表 4 所列, 经典的 CNN 算法识别率为 97.223%, 高于经典 HOG+SVM 行人检测算法, 而本文算法的识别率为 98.325%, 远高于其他两种经典算法。

表 3 CVC 数据集实验效果对比/%

检测方法	CNN+HOG+SVM	HOG+SVM	经典 CNN 算法
识别率	97.825	96.0125	96.937

表 4 DC-ped-dataset 实验效果对比/%

检测方法	CNN+HOG+SVM	HOG+SVM	经典 CNN 算法
识别率	98.325	96.768	97.223

结束语 随着人工智能概念进入公众的视野, 越来越多的研究人员关注其相关的领域, 不仅许多类似于深度学习的新兴算法被提出, 而且一些传统算法也在不停地发展。本文提出的利用 CNN+HOG 的特征提取方式就是一种“新, 老”算法的结合, 使用了多特征融合方法, 首先使用 CNN 在下采样层对数据进行多角度的特征描述, 然后用 HOG 对得到的特征图进行特征提取, 将得到的特征矩阵进行相关性分析, 取

其中相关性低的 3 组特征映射矩阵, 将它们组成一个新的特征矩阵, 最后交由 SVM 进行分类。实验证明, CNN+HOG 的多特征融合的特征描述比传统的单一的 HOG 特征更优秀, 识别率更高。在行人检测领域中, CNN+HOG 的特征表征能力更强, 比经典 CNN 算法更适合用来进行行人检测。

参考文献

- [1] HYVARINEN A. Fast and robust fixed-point algorithms for independent component analysis [J]. IEEE Transactions on Neural Networks, 1999, 10(3): 626-634.
- [2] LOWE D. Distinctive image features from scale Invariant keypoints[J]. IJCV, 2004, 60(2): 91-110.
- [3] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//CVPR. 2005.
- [4] CHENG H S, HU X F. The human body detection based on HOG and SVM technology research in static image[J]. Chinese Journal of Scientific Instrument, 2012, 19(5): 20-23.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image net classification with deep convolutional neural networks [C]//Advances in Neural Information Processing Systems. 2012: 1097-1105.
- [6] HINTON G E, SALAKHUTDINOV R. Reducing dimensionality of data with neural networks[J]. Science, 2006(504): 313.
- [7] BENGIO Y, COURVILLE A C, VINCENT P. Unsupervised feature learning and deep learning: A review and new perspectives [J]. CoRR abs/1206.2012. 5538.
- [8] SERMANET P, KAVUKCUOGLU K, CHINTALA S, et al. Pedestrian detection with unsupervised and multi-stage feature learning[C]//CVPR. 2013.
- [9] ENZWEILER M, EIGENSTETTER A, Schiele B. Multi-cue pedestrian classification with partial occlusion handling[C]//CVPR. 2010.
- [10] VIOLA P, JONES M J, SNOW D. Detecting pedestrians using patterns of motion and appearance[C]//IJCV. 2005: 153-161.
- [11] 田仙仙, 鲍泓, 徐成. 一种改进 HOG 特征的行人检测算法[J]. 计算机科学, 2014, 41(9): 237-259.