

Research on a New HOG Feature Face Recognition Algorithm^{*}

YILIHAMU·YAErmaiti^{*}

(College of Electncian Engineering, Xinjiang University, Urumqi 830047, China)

Abstract: For the reduction of human face recognition rate under non-uniform illumination, Laplace filter and discrete cosine transform (DCT) gradient direction histogram (HOG) face recognition algorithm is proposed. First by Laplace filtering of face image processing, its texture feature is highlighted; Secondly in discrete cosine transform (DCT), high frequency components are effectively filtered out; Then using discrete cosine inverse transformation (IDCT) reconstruction of face image reduces its dimension; Finally through the gradient direction histogram (HOG) operator the inherent characteristics of facial image recognition are extracted and the nearest neighbor method is used. The experimental results show that the proposed algorithm under different characteristic dimension of human face recognition rate as high as 95% in the database and Yale B group of uighurs face database construction in its recognition rate is as high as 98.5%, is better than other traditional algorithm. This algorithm has strong robustness and real time.

Key words: face recognition; non-uniform illumination; gradient direction histogram; discrete cosine transform; uighur face

EEACC: 6140C

doi: 10.3969/j.issn.1005-9490.2019.01.030

一种新的 HOG 特征人脸图像识别算法研究^{*}

伊力哈木·亚尔买买提^{*}

(新疆大学电气工程学院, 乌鲁木齐 830047)

摘 要: 针对人脸在非均匀光照下识别率的降低, 提出了拉普拉斯滤波和离散余弦变换 (DCT) 融合梯度方向直方图 (HOG) 人脸识别算法。首先通过拉普拉斯滤波对人脸图像进行处理, 突出其纹理特征; 其次进行离散余弦变换 (DCT), 有效滤除高频分量; 然后利用离散余弦逆变换 (IDCT) 重建人脸图像, 降低其维数; 最后通过梯度方向直方图 (HOG) 算子提取人脸图像固有特征, 并利用最近邻方法进行识别。实验结果表明, 该算法在不同特征维数下的 Yale B 人脸数据库中识别率高达 95% 以及课题组自建的维吾尔族人脸数据库中其识别率高达 98.5%, 优于其他传统算法, 具有很强的鲁棒性和实时性。

关键词: 人脸识别; 非均匀光照; 梯度方向直方图; 离散余弦变换; 维吾尔族人脸

中图分类号: TP391.4

文献标识码: A

文章编号: 1005-9490(2019)01-0157-06

人脸识别是最重要的生物特征识别技术之一, 可在社会安全、门禁安全系统、教学等众多领域中广泛应用。随着人脸识别技术的成熟发展, 许多相关产品已经很好的应用到人类生活中。同时, 人脸识别技术的研究和应用存在许多问题, 主要为其识别性能受诸多方面因素的影响, 如人脸图像分辨率模糊、光照多样性、多姿态以及丰富的遮挡等问题。

近年来, 针对非均匀光照对人脸识别效果的干扰而提出很多算法, 传统的人脸特征提取识别算法主要有 Gabor 人脸特征提取^[1], 局部二值模式 LBP 人脸特征提取^[2], 局部三值模式 LTP 人脸特征提取^[3], 梯度方向直方图 HOG 人脸特征提取^[4]等。

梯度方向直方图 (HOG) 在 2005 年由 Dalal^[5] 首

先提出, HOG 特征由于对方向、光照等因素具有不敏感性, 在计算机视觉研究中, 如在行人识别领域得到了广泛的应用。到 2011 年被 Deniz^[6] 等人首次应用于人脸识别中, 然而 HOG 特征在人脸识别中的研究成果相对较少, HOG 特征在人脸识别应用中的潜力还有待挖掘。

鉴于以上因素, 本文提出了基于拉普拉斯滤波和离散余弦变换 (DCT) 融合的梯度方向直方图 (HOG) 维吾尔族人脸识别算法。该算法通过拉普拉斯滤波变换突出人脸图像的固有纹理特征, 并通过离散余弦变换 (DCT) 有效滤除高频分量, 然后经过离散余弦逆变换 (IDCT) 处理后构造与原人脸图像相近的图像, 最后利用梯度方向直方图 (HOG) 算子提取其人脸图

项目来源: 国家自然科学基金项目 (61462082, 61866037)

收稿日期: 2018-01-02 修改日期: 2018-03-24

像的特征并利用最近邻方法对人脸图像进行分类识别。实验测试表明,该算法的识别率有了很大的提高,优于其他传统算法,并具有很强的鲁棒性。

1 拉普拉斯滤波

人脸图像经常在不同尺度上表现出不同的纹理、角点以及其他识别信息,其中拉普拉斯滤波就是其中一种多尺度分析工具,经过滤波之后它能够使更多的纹理特征表现出来。其具体表达式为:

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \quad (1)$$

为了对人脸图像处理,将其表示为二维离散形式:

$$\nabla^2 f(x, y) = [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)] - 4f(x, y) \quad (2)$$

$$g(x, y) = f(x, y) + c[\nabla^2 f(x, y)] \quad (3)$$

式中: $g(x, y)$ 为目标图像, $f(x, y)$ 为原始图像。当掩膜中心系数为正时, c 为+1;当掩膜中心系数为负时, c 为-1。人脸图像经过拉普拉斯滤波之后,可以突出人脸图像的边缘固有纹理特征。

2 离散余弦变换(DCT)

离散余弦变换(DCT)被广泛用于图像数据的压缩处理,并将数据信息转换为频域信息,并将重要的人图像数据信息聚集于低频系数上。离散余弦变换(DCT)算法中,把一幅人脸图像矩阵当作二维向量矩阵,在处理单个二维向量矩阵时,可以先将二维特征向量抽取后,再将其前面所抽取的二维特征向量放入其预先准备好的特征库以便进行其后面的识别;当需要将新的人脸图像添加到库中时,首先对其预处理工作,然后对该人脸图像进行 DCT 变换,这样就不用考虑其他人脸图像,即使库中的人脸图像数据很大,那么计算量也不会很大^[7]。

对于一幅 $M \times N$ 的人脸图像 $f(x, y)$, 其 DCT 表达式^[8]可以描述为:

$$C(u, v) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \alpha(u) \alpha(v) \times \cos \frac{(2m+1)u}{2M} \cos \frac{(2n+1)v}{2N} \quad (4)$$

$$\alpha(u) = \begin{cases} (1/M)^{1/2} & u=0 \\ (2/M)^{1/2} & u=1, 2, \dots, M-1 \end{cases} \quad (5)$$

$$\alpha(v) = \begin{cases} (1/N)^{1/2} & v=0 \\ (2/N)^{1/2} & v=1, 2, \dots, N-1 \end{cases} \quad (6)$$

式中 $C(u, v)$ 称为 DCT 内在的固定系数, DCT 变换是正交变换,人脸图像经过 DCT 变换后,其中数值大的系数主要聚集于左上角的非高频部分,也就是

说通过 DCT 变换后的主要信息数据量聚集在少数非高频系数值上。

二维离散余弦变换(DCT)是可逆的,其逆变换称之为离散余弦逆变换(IDCT),其定义为:

$$f(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \alpha(u) \alpha(v) C(u, v) \times \cos \frac{\pi(2m+1)u}{2M} \cos \frac{\pi(2n+1)v}{2N} \quad (7)$$

利用 IDCT,可获得重建的人脸图像,重建的人脸图像与原始人脸图像虽然不是完全相似,但是其主要的信息数据都保留了下来,这样一来,其人脸图像的光照等干扰因素都被一定的削弱,且降低了人脸图像的维数,对后续更好地提取人脸图像特征起到很大的作用。

3 HOG 特征提取描述

HOG 特征描述算子源于 SIFT 算法,是有用的图像局部差分信息的描述子,对表情多姿态变换、光照强弱、复杂场景变化等线性变换具有优越的鲁棒性。梯度方向直方图(HOG)特征提取在模式识别范围内取得了成功的应用,如人脸图像识别^[9]、行人人体检测^[10-11]、目标识别跟踪等。在计算机视觉和图像处理中,HOG 特征描述子是通过局部梯度方向直方图来计算和统计产生其图像固有特征,具体的计算过程如下:

(1) 计算每个像素点的梯度,如下式所示:

$$G_x(x, y) = I(x+1, y) - I(x-1, y) \quad (8)$$

$$G_y(x, y) = I(x, y+1) - I(x, y-1) \quad (9)$$

接着,估算梯度的幅值 $G(x, y)$ 和梯度的方向 $\alpha(x, y)$

$$G(x, y) = \sqrt{G_x^2 + G_y^2} \quad (10)$$

$$\alpha(x, y) = \arctan(G_y/G_x) \quad (11)$$

(2) 将人脸图像隔离成几个大小相似的单元模式 Cell, 设单元模式大小为 $s \times s$ 个二维像素,并以单元模式为其主要的统计单位来估计其梯度方向的质量分布图 Cell-HOG。将梯度方向划分为 K 个均匀的间隔(Bin),用 $V_k(x, y)$ 表示像素点 (x, y) 对第 k 个 Bin 的贡献权重值,权重值的取值算法:

$$V_k(x, y) = \begin{cases} G(x, y) & \alpha(x, y) \in \text{Bin}_k \\ 0 & \alpha(x, y) \notin \text{Bin}_k \end{cases} \quad 1 \leq k \leq K \quad (12)$$

对于每个单元模式 Cell_i , 计算其梯度方向的质量分布图向量 $[H_1^i, H_2^i, \dots, H_K^i]$, 其中

$$H_k^i = \sum_{(x, y) \in \text{Cell}_i} V_k(x, y), \quad 1 \leq k \leq K \quad (13)$$

(3) 相邻的 $t \times t$ 个 Cell 织成一个块 Block。将该块在图像中向右侧或向下方移动一个 Cell 的大小,以获

得下一个块。因此,块与块之间存在着重叠。对于每个块,拼接块内所有单元的梯度方向直方图向量,得到块的梯度方向直方图矢量 Block-HOG,长度定义为 Kt^2 ,接着,将每个块的梯度方向直方图向量归一化处理。使用的方法是按 L_2 模归一化,即对于向量 v ,设

$$v \rightarrow \frac{v}{\sqrt{\|v\|_k^2 + \varepsilon^2}} \quad (14)$$

式中: v 为标准化前的一个 block 的 HOG 二维特征向量, $\|v\|_k$ 表示 k -norm 运算,在此 $k=2$, ε 是极小值,用于阻止式(14)中分母为零的现象,进而导致其计算得到无穷大值。

(4)最后,将所有 Block 的梯度方向直方图融合起来,获得其最终输入图像的 HOG 特征。

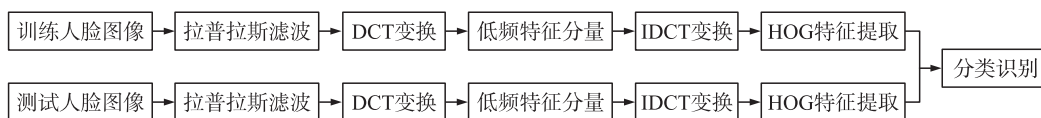


图 1 人脸图像识别流程

具体算法步骤如下:

(1)对训练样本人脸图像进行拉普拉斯滤波变换预处理;

(2)对完成拉普拉斯滤波变换预处理后的人脸图像分成 8×8 人脸图像块,然后将图像块进行 DCT 变换,提取其人脸图像的左上角 10 个低频系数特征分量;

(3)将提取的人脸图像左上角 10 个低频系数特征分量通过离散余弦逆变换 (IDCT) 获得其与原始图像相近的重构的人脸图像;

(4)通过式 (10) 和式 (11) 获得重构人脸图像的梯度;

(5)运用式 (12) 以单元 Cell 为单位模式,统计梯度方向直方图;

(6)以块 block 为单位,对块内的梯度方向直方图进行归一化处理;

(7)将各个块内的梯度方向直方图拼接起来,得到最终人脸图像的 HOG 特征;

(8)最后用最近邻方法对人脸图像进行分类识别。

5 仿真实验分析

本文选用 YaleB 人脸数据库进行本文算法验证性实验。YaleB 人脸数据库是一个可变光照的标准人脸库,该人脸数据库是具有多方向光照变化和多姿态变化的人脸数据库,共涵盖了 10 个人,构建了 576 种多姿态和多光照条件 (9 种姿态 $\times$ 64 种光照) 下的共 5 760 幅人脸图像。所有选用的人脸图像都

4 本文算法实现描述

由于人脸在非均匀光照下的识别率困难以及鲁棒性差的问题,本文提出了基于拉普拉斯滤波和离散余弦变换 (DCT) 融合的梯度方向直方图 (HOG) 人脸识别算法。该算法首先通过拉普拉斯滤波把非均匀光照的影响减弱,突出人脸图像的固有纹理边缘特征,然后利用离散余弦变换 (DCT) 来抽取人脸图像的非高频固有特征分量,同时减少了人脸图像的特征维数,并利用离散余弦逆变换 (IDCT) 构建复原与原始人脸图像相接近的图像信息,这样即可保留其人脸重要的特征信息,最后使用梯度方向直方图 (HOG) 算子抽取其人脸图像的特征。本算法的流程图 1 所示。

做了类似的预处理工作,涵盖了裁剪、矫正、缩放等,如图 2 所示。



图 2 YaleB 人脸数据库

同时,本文也采用了课题组成员自建的新疆维吾尔族人脸信息数据库进行了验证实验。维吾尔族人脸图像均来自新疆各个地区,不同年龄段,不同性别,具有典型的代表性。如图 3 所示。



图 3 非均匀光照下的维吾尔族人脸数据库

5.1 基于不同算法的人脸识别率

本文首先通过不同的算法对 YaleB 人脸数据库和本文课题组自建的新疆维吾尔族人脸数据库进行实验对比。首先分别将两种不同的人脸数据库进行归一化工作处理,即将人脸图像归一化为 64×64、

32×32、16×16 大小的人脸图像,然后运用传统 DCT 算法、传统 LBP 算法、传统 HOG 特征提取算法、DCT+LBP 算法、HOG+LBP 融合算法以及本文算法进行实验比较,最后利用最近邻方法实现人脸的分类。结果如表 1 所示。

表 1 基于不同算法的人脸识别率

图像尺寸	数据库	识别率/%					
		DCT	LBP	HOG	DCT+LBP	HOG+LBP	本文算法
16×16	Yale B	63.39	68.39	67.59	71.49	81.99	91.49
	维吾尔族人脸	84.49	90.75	89.99	93.45	94.89	97.95
32×32	Yale B	68.99	72.79	71.86	76.19	83.59	92.77
	维吾尔族人脸	85.79	91.12	90.34	94.77	95.27	98.17
64×64	Yale B	71.53	74.53	73.43	79.43	85.53	93.23
	维吾尔族人脸	87.52	91.89	90.71	94.80	95.56	98.99

由表 1 结果我们可以看到(64×64 为例),由本文课题组自建的维吾尔族人脸的识别率中,本文提出的基于拉普拉斯滤波和离散余弦变换(DCT)融合的梯度方向直方图(HOG)维吾尔族人脸识别算法最高,达到了 98.99%,其次 HOG+LBP 融合算法的识别率也很高,为 95.56%。其他算法的识别率也是比较高。而对于 YaleB 人脸数据库的识别率中,人脸辨识率达到了 93.23%,说明本文所提出的算法性能依然比较高,其次为 HOG+LBP 融合算法为 85.53%。

从表 1 中我们比照课题组自建的维吾尔族人脸数据库和 YaleB 人脸数据库,我们可以看出,这两个人脸图像数据库的共同点是均处于非均匀光照下,而

且均在本文所提出的算法和 HOG+LBP 融合算法下的人脸辨识率最高,这主要是因为方向梯度方向直方图(HOG)算法具有很重要的几何不变性和光度变换不变性特性,尤其是本文所提出的算法事先进行了拉普拉斯滤波和离散余弦变换(DCT)处理,因此本文算法的人脸辨识率性能好于其他 5 种算法,尤其是在非均匀光照下本文算法更具有更强的鲁棒性。

5.2 基于不同特征维数下的人脸识别率

为了更加体现出本文所算法的优越性,我们以 64×64 人脸图像为例,进行了维吾尔族人脸和 YaleB 人脸两种数据库在不同的特征维数下的人脸识别率的实验,如图 4 和图 5 所示。

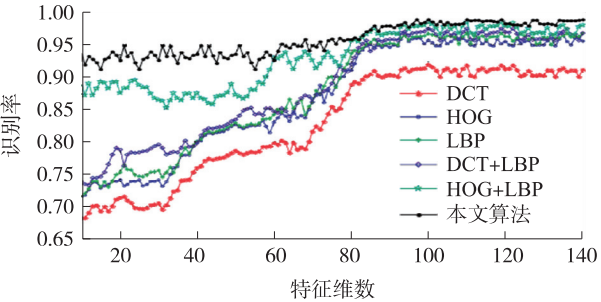


图 4 维吾尔族人脸数据库识别率

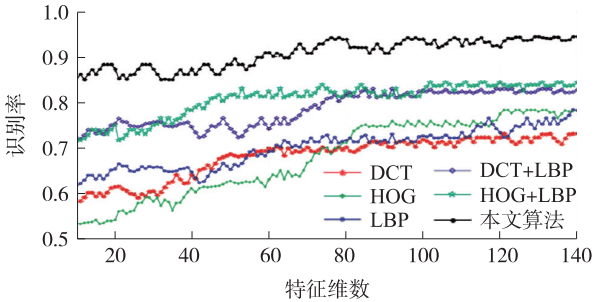


图 5 YaleB 人脸数据库识别率

表 2 不同特征维数下的维吾尔族人脸识别率

单位: %

特征维数	识别算法					
	DCT	HOG	LBP	DCT+LBP	HOG+LBP	本文算法
20	72	74	75	78	87.5	92.5
40	76	80	80	80	87.5	93.5
60	80	83.5	84.5	84.7	92.5	92.5
80	88	91.5	93.5	92.5	96	96
100	92.5	95.5	96.5	97	97.5	98
120	92	95.2	96	96.5	97	97.5
140	92.2	96.2	96.1	96	97	98.5

表 3 不同特征维数下的 YaleB 人脸识别率							单位:%
特征维数	识别算法						
	DCT	HOG	LBP	DCT+LBP	HOG+LBP	本文算法	
20	61.5	53.5	65	75	75	87	
40	64.5	61	64.5	74	79	87	
60	69.5	62.5	69	73.5	82.2	90.5	
80	70	71.5	72	81.5	81.5	94	
100	71.5	74.5	72.5	82	82	94	
120	72	78.2	72.5	82.5	83.5	93	
140	72.2	79	79	82	83.5	95	

从表 2 和表 3 可以看出,本文所提出的算法在 YaleB 人脸数据库和自建的维吾尔族人脸数据库中的识别率最高,尤其是在维吾尔族人脸识别中,这主要是因为维吾尔族人脸数据库在采集过程中体现的是正面效果光照图像,而 YaleB 人脸数据库是具有不同姿态下的光照图像,其次就是本文算法中利用拉普拉斯滤波来突出人脸图像的边缘和纹理特征,减少光照影响,然后通过离散余弦变换(DCT)融合梯度方向直方图(HOG)就可以很好地提取其人脸特,所以维吾尔族人脸数据库辨析率要比 YaleB 人脸数据库高;当然 HOG+LBP 融合算法的识别流程也比较高,这主要是因为这两种算法体现了方向梯度方向直方图(HOG)算法具有几何不变性和光度变换不变性的优秀特性。

综上所述,本文所提出的算法不仅能够增强其人脸图像辨识能力,而且对多变化的不同光照等具有很好的鲁棒性。可以看出本算法提取脸部特征得到的识别效果优越于其他人脸识别算法,它能提取到比传统 LBP、传统 HOG 和传统 DCT 等算法更加丰富有效的人脸纹理信息,对非均匀光照等具有很强的抗干扰能力,体现了本文算法在提取人脸图像特征方面更胜一筹。

5.3 时间测试分析

时间测试实验所用硬件环境为 Intel Corei 7,4G 内存;仿真环境为 MATLAB2014b。在 YaleB 人脸库和自建的维吾尔族人脸库中,选取每个人的 5 幅图像作为训练样本,分别将其他的人脸图像作为测试样本。

由表 4 可见,本文算法无论在训练平均时间还是在识别平均时间上都体现最佳的状态,表中加粗数据为最优状态。虽然本文算法在训练平均时间上排名第二,主要是因为训练过程是离线的,但是对整个识别人脸过程中的计算机系统的运行时间,识别系统没有很大的影响,所以本文算法仍是优越于其他算法,具有很好的实时性。

表 4 不同算法的时间测试			
算法	最高识别率	训练平均时间/s	识别平均时间/s
DCT	72.2%	15.358	0.756
HOG	79%	12.887	0.741
LBP	79%	10.235	0.687
DCT+LBP	82%	9.987	0.274
HOG+LBP	83.5%	2.897	0.307
本文算法	95%	3.148	0.187

最后,为了验证本文算法的优越性,我们计算了相应特征描述子的受试者工作特征(ROC)曲线的 AUC(Area Under Curve)值。AUC 值指处于 ROC 曲线下方的那部分面积大小,通过计算 YaleB 人脸数据库的 ROC 曲线下面积,评估其算法的辨识能力,也可以筛选出最佳算法及其分类器,其值越大就代表着其分类器的性能越好,反之曲线下方面积越小 AUC 值越低,则相应的算法及分类器的性能就越差,我们通过几种特征的 AUC 值来对几种特征进行性能比对,如表所示,通过表中 AUC 值可知,本文提出的特征识别算法的 ROC 曲线更为优越、人脸识别效果更加良好。

表 5 不同算法的 ROC 曲线的 AUC 值	
算法	ROC 曲线下面积
DCT	0.835 4
HOG	0.941 5
LBP	0.899 4
DCT+LBP	0.971 4
HOG+LBP	0.972 1
本文算法	0.989 7

6 结论

本文提出拉普拉斯滤波和离散余弦变换(DCT)融合的梯度方向直方图(HOG)人脸识别算法。该算法通过拉普拉斯滤波突出其人脸图像的边缘和纹理特征,减弱非均匀光照干扰,利用离散余弦变换

(DCT)消除人脸图像高频信息数据,保留其低频信息数据,同时降低了人脸图像维数,并运用离散余弦逆变换(IDCT)重新构建人脸图像,虽然重构的人脸图像与原始人脸图像有一定的差异,但是基本上保留了原始人脸图像的信息数据,进而利用梯度方向直方图(HOG)算子提取出人脸固有特征,进行识别分类。本文所提出的算法在一定程度上能够在光照干扰下很好地提取人脸图像特征,并取得了优越的分类识别效果,在YaleB人脸数据库以及课题组自建的维吾尔族人脸数据库中的实验仿真显示,该算法在不同特征维数下的YaleB人脸数据库中识别率高达95%以及课题组自建的维吾尔族人脸数据库中其识别率高达98.5%,与其他人脸图像识别算法比较其识别率有了很大的提高,识别辨析具有很好的鲁棒性。

参考文献:

- [1] Ahmad B F, Arif W M. Gabor Wavelet Based Face Recognition under Varying Lighting, Pose and Expression Conditions[C]//2015 2nd International Conference on Computing for Sustainable Global Development, 2015:1314-1318.
- [2] Mohamed A A, Yampolskiy R V. An Improved LBP Algorithm for Avatar Face Recognition[C]//XXIII International Symposium on Information, Communication and Automation Automation Technologies, 2011:1-5.
- [3] 李伟生,王立逗,周丽芳. 一种基于LTP自适应阈值的人脸识别方法[J]. 小型微型计算机系统, 2014, 9(9):2099-2103.
- [4] 郭金鑫,陈玮. 基于HOG多特征融合与随机森林的人脸识别[J]. 计算机科学, 2013, 40(10):279-283.
- [5] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection[C]//Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005:886-893.
- [6] Deniz, O, Bueno, G, Salito, J, et al. Face Recognition Using Histograms of Oriented Gradients [J]. Pattern Recognition Letters, 2011, 32(12):1598-1603.
- [7] 王燕,王芸芸. 一种DCT和ELBP融合的人脸特征提取方法[J]. 计算机工程与应, 2017(4):170-175.
- [8] Jadhav D V, Holambe R S. Radon and Discrete Cosine Transforms Based Feature Extraction and Dimensionality Reduction Approach for Face Recognition[J]. Signal Proc, 2008, 88(10):2604-2609.
- [9] Tan Hengliang, Yang Bing, Ma Zhengming. Face Recognition Based on the Fusion of Global and Local HOG Features of Face Images [J]. IET Computer Vision, 2013:224-234.
- [10] 孙锐,陈军. 基于显著性检测与HOG-NMF特征的快速行人检测方法[J]. 电子与信息学报, 2013, 35(8):1921-1925.
- [11] Park W J, Kim D H. Fast Human Detection Using Selective Block-Based HOG-LBP [C]//IEEE International Conference on Image Processing, 2012:601-604.



伊力哈木·亚尔买买提(1978-),男,维吾尔族,新疆乌鲁木齐人,新疆大学,副教授,主要研究领域为模式识别,人脸识别,65891080@qq.com。