

DOI: 10.19322/j.cnki.issn.1006-4710.2017.02.001

一种单幅前视图的人脸识别算法

毛雯, 朱虹, 张岩

(西安理工大学 自动化与信息工程学院, 陕西 西安 710048)

摘要: 本文针对依据单幅正视人脸照片在监控视频中进行目标人物识别的问题,提出了一种通过单应变换获得空间多姿态的虚拟人脸的方法,对得到的人脸提取 Gabor 特征,并采用主成分分析法(PCA)降维之后,通过支持向量机(SVM)对其训练分类器,最后,采用半监督学习的 SVM 进行人脸识别。该方法在全国研究生智慧城市视频挑战赛的数据集上得到测试并验证了其有效性。

关键词: 单幅正视图像; 人脸识别; 单应变换; Gabor 特征; SVM

中图分类号: TP391.41

文献标志码: A

文章编号: 1006-4710(2017)02-0127-05

A face recognition algorithm based on the single front view

MAO Wen, ZHU Hong, ZHANG Yan

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: Based on the single front view of human face images, this paper aims at the solution of the problem by target face recognition in the surveillance video with a method proposed to obtain virtual and multi-pose face samples by homography; it extracts the Gabor feature; using principal component analysis (PCA) to reduce the data dimension; after its use for training SVM classifier, employing the semi-supervised SVM for face recognition. The method is used to verify the validity of the data set of National Graduate contest on smart-city technology and creative design.

Key words: single front view; face recognition; homography; Gabor transform; SVM

因人脸识别的应用领域越来越广泛,为提高其识别准确率,人脸识别方法的研究被大批学者关注。机器学习方法的提出,使得人脸识别准确率有了保障,然而,针对机器学习类算法,其需要大量训练样本。但是,在现实中的一些特殊状况,每个人脸类别只能提供一张人脸的训练样本,因此,基于小样本的人脸识别^[1-3]方法的研究成为一个研究热点。

文献[1]和文献[4]提出对人脸实现三维重建扩充虚拟样本的方法,但获取人的三维人脸图像比较困难,因此在识别应用中并没有能够得到广泛的应用;文献[1]、文献[5]、文献[6]提出,使用二元高次多项式函数对人脸各种姿态的特征集进行拟合;文献[1]、文献[7]、文献[8]提出构造虚拟样本人脸图像并对其进行特征提取,再采用主成分分析法进行人脸识别;文献[9]利用镜像变换、亮度变换和尺度变换产生虚拟人脸。因此,通过对单幅人脸进行一

系列的变形,生成了多种姿态的人脸虚拟图像。在增加了虚拟训练样本后,得到的平均识别率高于基于单幅人脸的识别结果。

而且,在现实生活中,人脸识别在监控视频中的应用也越来越受到人们的关注,但是,要想拿到所关注目标人物的多幅人脸图像是不切实际的,所以,本文所提出的方法,同样是通过单幅前视图扩充多姿态的虚拟训练样本,来解决样本单一对人脸识别率的影响。

1 多姿态虚拟训练样本的生成

本文根据三维物体在二维图像上的投影原理,提出了基于单应变换^[10]的多姿态虚拟人脸的构成方法。

对空间上相同点在不同平面上的位置关系,可以用单应变换公式来表示:

收稿日期: 2016-10-26

基金项目: 国家自然科学基金资助项目(61673318)

作者简介: 毛雯,女,硕士生,研究方向为模式识别。E-mail:1612148294@qq.com

通讯作者: 朱虹,女,教授,博导,博士,研究方向为数字图像处理与模式识别等。E-mail:zhuhong@xaut.edu.cn

$$X_i = \mathbf{H} \cdot x_i \quad (1)$$

其中, $X_i (i=1, 2, \dots, n)$ 为原始人脸上的某一点, n 为人脸面上点的个数, 而 x_i 为生成的虚拟姿态上的对应点, 单应矩阵^[11] \mathbf{H} 是 3×3 的满秩矩阵^[11], 其中, $h_{33}=1$, 则 \mathbf{H} 表示为:

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{pmatrix} \quad (2)$$

给定两对点, 假设有两个图像上的点 (x_1, y_1) 和 (x_2, y_2) , 其齐次坐标为 $(x_1, y_1, 1)$ 和 $(x_2, y_2, 1)$, 代入式(1)可以得到:

$$x_2 = \frac{x_1 h_{11} + y_1 h_{12} + h_{13}}{x_1 h_{31} + y_1 h_{32} + 1} \quad (3)$$

$$y_2 = \frac{x_1 h_{21} + y_1 h_{22} + h_{23}}{x_1 h_{31} + y_1 h_{32} + 1} \quad (4)$$

由式(3)和式(4)可得等价矩阵形式:

$$\mathbf{S}\mathbf{u} = \mathbf{v} \quad (5)$$

其中,

$$\mathbf{S} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1 x_2 & -x_2 y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1 y_2 & -y_1 y_2 \end{bmatrix}$$

$$\mathbf{u} = [h_{11} \quad h_{12} \quad h_{13} \quad h_{21} \quad h_{22} \quad h_{23} \quad h_{31} \quad h_{32}]^T$$

$$\mathbf{v} = [x_2 \quad y_2]^T$$

如果有四对不共线匹配点对, 该方程组就能够垒到 8 行, 存在唯一解, 即可计算出单应矩阵的 8 个系数。

如图 1 所示, 人脸图像尺寸为 $h \times w$, 设摄像机视角为 φ , 人脸面与投影面之间的夹角为 θ , w_{new1} 和 h_{new1} 是通过单应变换后新图的宽度和高度。



图 1 单幅前视图

Fig. 1 Signal front view

根据物体相对运动原理, 当摄像机位置保持不动, 人脸相对摄像机做旋转时, 可以认为人脸位置不动, 摄像机相对人脸做旋转。假设单样本人脸前视图中成像物都为二维平面信息, 再让垂直于摄像机光轴的人脸所在平面(人脸面)分别相对摄像机向上、下、左、右四个方向旋转特定角度, 通过几何模型推导, 完成单样本人脸前视图到多种姿态的变换。

以人脸面向上旋转为例, 此时, 令 $h_{\text{new1}} = h$, 对摄像机和成像面建立直角坐标系, 求解 w_{new1} 时可分为以下两种情况:

当 $\sqrt{\frac{1 + (\cot(\theta))^2}{(\cot(\theta) + \tan(\varphi/2))^2}} < 1$ 时,

$$w_{\text{new1}} = w \sqrt{\frac{(\cot(\theta) + \tan(\varphi/2))^2}{(\cot(\theta) - \tan(\varphi/2))^2}} \quad (6)$$

当 $\sqrt{\frac{1 + (\cot(\theta))^2}{(\cot(\theta) + \tan(\varphi/2))^2}} > 1$ 时,

$$w_{\text{new1}} = 2w \cdot$$

$$\sqrt{\frac{\left(\cos(\theta) - \frac{1}{2}\right)^2 + \left(\sin(\theta) - \frac{1}{2\tan(\varphi/2)}\right)^2}{\left(\frac{2\cos(\theta) - 1}{2\tan(\varphi/2)\sin(\theta) - 1}\right)^2 + \left(\frac{1}{\tan(\varphi/2)}\right)^2}} \quad (7)$$

此时, w_{new1} 已知, 然后找到人脸单应变换前后对应的 4 个点, 求得 \mathbf{H}_{up} 。同理, 可求得 \mathbf{H}_{down} 、 $\mathbf{H}_{\text{right}}$ 和 \mathbf{H}_{left} , 即:

$$\mathbf{H}_{\text{up}} = \begin{bmatrix} \frac{w}{w_{\text{new1}}} & \frac{w(w - w_{\text{new1}})}{2hw_{\text{new1}}} & -\frac{w(w - w_{\text{new1}})}{2w_{\text{new1}}} \\ 0 & \frac{w}{w_{\text{new1}}} & 0 \\ 0 & \frac{w - w_{\text{new1}}}{hw_{\text{new1}}} & 1 \end{bmatrix} \quad (8)$$

$$\mathbf{H}_{\text{down}} = \begin{bmatrix} 1 & \frac{w(w - w_{\text{new1}})}{2h} & 0 \\ 0 & \frac{w_{\text{new1}}}{w} & 0 \\ 0 & \frac{w_{\text{new1}} - w}{hw} & 1 \end{bmatrix} \quad (9)$$

$$\mathbf{H}_{\text{right}} = \begin{bmatrix} \frac{h_{\text{new1}}}{h} & 0 & 0 \\ -\frac{h - h_{\text{new1}}}{2w} & 1 & 0 \\ \frac{h_{\text{new1}} - h}{hw} & 0 & 1 \end{bmatrix} \quad (10)$$

$$\mathbf{H}_{\text{left}} = \begin{bmatrix} \frac{h}{h_{\text{new1}}} & 0 & 0 \\ \frac{h(h - h_{\text{new1}})}{2h_{\text{new1}}w} & \frac{h}{h_{\text{new1}}} & -\frac{h(h - h_{\text{new1}})}{2h_{\text{new1}}} \\ \frac{h - h_{\text{new1}}}{h_{\text{new1}}w} & 0 & 1 \end{bmatrix} \quad (11)$$

根据式(1)可知, $x_i = \mathbf{H}^{-1} \cdot X_i$, 对图 1 依据上述方法从上、下、左、右四个方向进行三维旋转得到的虚拟姿态如图 2 所示。在该方法中, 与西安交通大学人工智能与机器人研究所的东方人脸数据库(XJTU 人脸库)^[12] 相对比, 四个方向的限制旋转角度在 $\pm 45^\circ$ 以内。

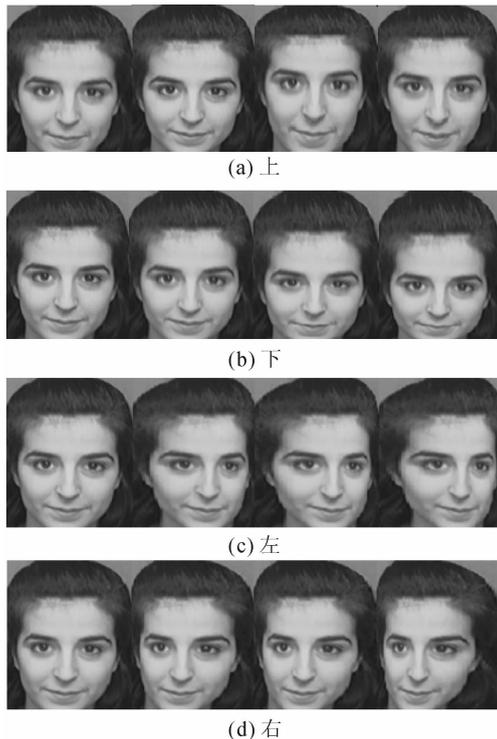


图2 多姿态人脸实验结果

Fig. 2 Multi-pose face experiment results

2 基于 Gabor 小波变换的特征提取

2.1 人脸图像的 2D-Gabor 特征提取

生物学实验发现,二维 Gabor 滤波器与人类视觉系统极度的类似,拥有很好地方向选择性和空间局部性。一个 2D-Gabor 滤波器^[13,14]是一个正弦平面波和高斯核函数的乘积,其具有在空间域和频率域同时取得最优局部化的特性。

Gabor 小波核函数表达式定义为:

$$\Psi_{u,v}(x,y) = \frac{|\mathbf{k}|^2}{\sigma^2} \exp\left(-\frac{|\mathbf{k}|^2 |r|^2}{2\sigma^2}\right) \cdot \left[\exp(i\mathbf{k} \cdot r) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (12)$$

其中, $\mathbf{k} = \begin{bmatrix} k_x \\ k_y \end{bmatrix} = \begin{bmatrix} k_v \cos \varphi_u \\ k_v \sin \varphi_u \end{bmatrix}$, x, y 代表原图某个

像素点的坐标,而 $k_v = 2^{-\frac{v+2}{2}\pi}$, $\varphi_u = \frac{u\pi}{D}$ (D 表示总的

方向数), $r = \sqrt{x^2 + y^2}$, v 和 u 分别代表 Gabor 小波的尺度因子和方向,按照经验值选取 5 个尺度和 8 个方向,即 $v=1,2,\dots,5$ 和 $u=1,2,\dots,8$; 式(12)中

$\frac{|\mathbf{k}|^2}{\sigma^2}$ 的作用是补偿由频率来决定的能量谱的衰弱,

通常 $\sigma = 2\pi$; $\exp\left(-\frac{|\mathbf{k}|^2 |r|^2}{2\sigma^2}\right)$ 是一个高斯函数,它的意义是通过加窗来限制振荡函数 $\exp(i\mathbf{k} \cdot r)$ 的

范围,令其只能在局部有效; $\exp\left(-\frac{\sigma^2}{2}\right)$ 是图像中的直流分量,减去它可以使二维 Gabor 滤波器不受直流分量大小的影响,其目的在于减少光照对 Gabor 特征提取的影响。设输入人脸图像 $\mathbf{I}(x,y)$, 将其与式(12)中的 Gabor 小波核函数 Ψ 进行卷积:

$$\mathbf{H}(u,v,x,y) = \Psi_{u,v}(x,y) \otimes \mathbf{I}(x,y) = \iint \Psi_{u,v}(x,y) \cdot \mathbf{I}(x,y) dx dy \quad (13)$$

由此,每幅图像可以得到 40 个复系数的 Gabor 纹理特征。对图 1 进行 5 个尺度和 8 个方向上的纹理特征提取,其结果如图 3 所示。

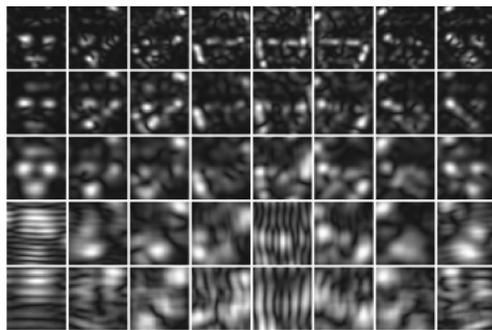


图3 Gabor 特征幅值响应结果

Fig. 3 Gabor feature amplitude response results

2.2 Gabor 特征的降采样

按照上节中所述,最终得到的特征维数是原始图像大小的 40 倍,假设输入的原始图像大小为 80×80 像素,则通过 Gabor 滤波器后的特征维数为 256 000 维 ($80 \times 80 \times 40$), 本文采用了 PCA 算法^[15] 对其进行降维。

PCA 算法的具体步骤如下:

假设训练样本数为 N , 其每个样本 \mathbf{x}_j ($j=1,2,\dots,N$) 的维数为 M ($M \gg N$)。

1) 计算所有训练图片的平均脸 $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \frac{1}{N} \sum_{j=1}^N \mathbf{x}_j \quad (14)$$

2) 计算每个样本与平均脸 $\boldsymbol{\mu}$ 的差值 \mathbf{d}_j :

$$\mathbf{d}_j = \mathbf{x}_j - \boldsymbol{\mu}, j = 1, 2, \dots, N \quad (15)$$

3) 利用差值脸构造协方差矩阵 \mathbf{C} :

$$\mathbf{C} = \frac{1}{N} \sum_{j=1}^N \mathbf{d}_j \mathbf{d}_j^T = \frac{1}{N} \mathbf{A} \mathbf{A}^T \quad (16)$$

$$\mathbf{A} = [\mathbf{d}_1 \quad \mathbf{d}_2 \quad \dots \quad \mathbf{d}_N]$$

4) 求解协方差矩阵 \mathbf{C} 的特征值与特征向量,并构造特征脸空间。

协方差矩阵的维数为 $M \times M$, 由于其计算量会很大,所以采用奇异值分解(SVD)定理,求解 $\mathbf{A}^T \mathbf{A}$ 的特征值和特征向量来表征原始的特征值和特征向量。

5) $\mathbf{A}\mathbf{A}^T$ 的正交归一化特征向量 γ'_o 为:

$$\gamma'_o = \frac{1}{\sqrt{\lambda_o}} \mathbf{A}\mathbf{y}_o, o = 1, 2, \dots, l \quad (17)$$

其中, λ_o 为矩阵 $\mathbf{A}^T\mathbf{A}$ 的 l 个非零特征值, \mathbf{y}_o 为 $\mathbf{A}^T\mathbf{A}$ 对应于 λ_o 的特征向量。一般情况下, 取特征值能量占有所有特征值能量的 90% 以上所对应的特征向量个数 k 个, 组成 PCA 的投影子空间 \mathbf{W} :

$$\mathbf{W} = [\gamma'_1 \gamma'_2 \dots \gamma'_k] \quad (18)$$

6) 任何一幅人脸图像可通过下式投影到 PCA 子空间:

$$\mathbf{T} = \mathbf{W}^T(\mathbf{x}_j - \boldsymbol{\mu}) \quad (19)$$

将训练样本和测试样本投影到 PCA 子空间后, 就可以选择合适的分类器对降维后的数据进行分类。

2.3 人脸识别算法

1) 扩充训练样本: 将上述方法所生成的虚拟样本, 与给定的正视图样本一起, 作为分类器的训练样本。

2) 特征提取: 对每幅人脸图像提取 Gabor 特征, 然后用 PCA 算法对其进行降维。

3) 识别分类: 将训练样本的特征对 SVM 分类器^[16-18]进行训练; 并采用训练好的 SVM 分类器, 对测试样本进行识别。

3 实验结果分析

3.1 实验数据

为了验证本文算法的有效性, 这里采用了 2015 年全国研究生智慧城市技术与创意设计大赛的人脸识别数据集^[19], 如图 4 所示。图中只给出了 10 个人的正视图, 之后, 需要在两个场景下的监控视频中识别出他们。



图 4 智慧城市大赛 10 幅前视图人脸

Fig. 4 10 front view of human from smart-city technology and creative design

3.2 实验结果

通过单应变换的方法对图 4 中的每个人脸生成 33 幅虚拟多姿态人脸。

图 5 是本文在视频帧图像中识别出的目标人物的实验结果。

表 1 给出了对图 4 所示 10 类人的比较实验结

果, 可以看出, 本文方法的识别率明显高于其它三种方法。

同样将从视频帧序列中检测到的人脸进行手动归类, 将除图 4 所示 10 幅人脸外的其余人脸都归为第 11 类, 采用半监督学习的 SVM 算法, 得到 11 类的人脸识别结果, 如表 2 所示。



(a) 场景一



(b) 场景二

图 5 两个场景下的人脸识别结果

Fig. 5 Face recognition results in two scenes

表 1 10 类人脸识别结果

Tab. 1 10 types of face recognition results

方法	训练样本	测试样本	准确率/%
文献[5]	330	1 397	70.57
文献[8]	330	1 397	76.77
文献[19]	330	1 397	78.88
本文方法	330	1 397	87.75

表 2 11 类人脸识别结果

Tab. 2 11 types of face recognition results

方法	训练样本	测试样本	准确率/%
文献[5]	330	7 351	80.56
文献[8]	330	7 351	95.90
文献[19]	330	7 351	94.60
本文方法	330	7 351	97.03

仅对于表 2 中包含的第 11 类人脸的识别结果如表 3 所示。识别率越高表明得到的模型的排它性越好。

由以上实验结果可知, 通过本文算法扩充的多姿态虚拟训练样本, 相较于文献[5]、文献[8]和文献[19], 进一步丰富了训练样本, 对提高识别准确率效果显著。

表3 第11类人脸识别结果

Tab. 3 Face recognition results of type 11

方法	测试样本	准确率/%
文献[5]	5 954	83.47
文献[8]	5 954	96.60
文献[19]	5 954	95.11
本文方法	5 954	98.06

4 结 语

本文提出了一种通过单应变换实现对单幅正视图图像构建多姿态虚拟人脸图像的方法,然后对其提取 Gabor 特征,训练 SVM 分类器,较好的解决了单一训练样本对人脸识别率的影响。实验结果验证了该方法的有效性。

参考文献:

- [1] 邹国锋,傅桂霞,李海涛,等. 多姿态人脸识别综述[J]. 模式识别与人工智能, 2015, 28(7): 613-625.
ZOU Guofeng, FU Guixia, LI Haitao, et al. A survey of multi-pose face recognition [J]. Pattern Recognition and Artificial Intelligence, 2015, 28(7): 613-625.
- [2] ZHANG L, YANG M, FENG X. Sparse representation or collaborative representation: Which helps face recognition [C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 471-478.
- [3] MARASAMY P, SUMATHI S. Automatic recognition and analysis of human faces and facial expression by LDA using wavelet transform [C]//Computer Communication and Informatics (ICCCI), 2012 International Conference on. IEEE, 2012: 1-4.
- [4] GAO Q X, ZHANG L, ZHANG D. Face recognition using FLDA with single training image per person [J]. Applied Mathematics & Computation, 2008, 205(2): 726-734.
- [5] 朱长仁,王润生. 基于单视图的多姿态人脸识别算法[J]. 计算机学报, 2003, 26(1): 104-109.
ZHU Changren, WANG Runsheng. Multi-pose face recognition based on a single view [J]. Chinese Journal of Computers, 2003, 26(1): 104-109.
- [6] 张尤赛,杨姝. 局部加权平均虚拟样本的多姿态人脸识别算法[J]. 江苏科技大学学报(自然科学版), 2013, 27(1): 74-79.
ZHANG Yousai, YANG Shu. Multi-pose face recognition based on virtual samples of local weighted mean [J]. Journal of Jiangsu University of Science and Technology(Natural Science Edition), 2013, 27(1): 74-79.
- [7] 温津伟,罗四维,赵嘉莉,等. 通过创建虚拟样本的小样本人脸识别统计学习方法[J]. 计算机研究与发展, 2002, 39(7): 814-818.
WEN Jinwei, LUO Siwei, ZHAO Jiali, et al. A small sample face recognition statistical learning method based on virtual samples [J]. Journal of Computer Research and Development, 2002, 39(7): 814-818.
- [8] 龚卫国,黄潇莹,李伟红. 基于正弦变换的人脸姿态矫正及识别研究[J]. 计算机工程与应用, 2010, 46(22): 213-216.

- GONG Weiguo, HUANG Xiaoying, LI Weihong, et al. Face pose correction for facial recognition based on sine transform [J]. Computer Engineering and Applications, 2010, 46(22): 213-216.
- [9] 单桂军. 基于虚拟样本扩张法的单样本人脸识别算法研究[J]. 科学技术与工程, 2013, 13(14): 3908-3911.
SHAN Guijun. Virtual sample generating for face recognition from a simple training sample per person [J]. Science Technology and Engineering, 2013, 13(14): 3908-3911.
- [10] GONG J, ZHENG C, TIAN J, et al. An image-sequence compressing algorithm based on homography transformation for unmanned aerial vehicle [C]//International Symposium on Intelligence Information Processing and Trusted Computing. IEEE, 2010: 37-40.
- [11] 张雪波,方勇纯,马博军,等. 基于单应矩阵的摄像机标定方法及应用[J]. 控制工程, 2010, 17(2): 248-251.
ZHANG Xuebo, FANG Yongchun, MA Bojun, et al. Homography-based camera calibration method and its application [J]. Control Engineering of China, 2010, 17(2): 248-251.
- [12] The Institute of Artificial Intelligence and Robotics of Xi'an Jiaotong University. The oriental face database [EB/OL]. [2013-11-27]. <http://www.aiar.xjtu.edu.cn/dfrlsjk3.htm>.
- [13] 王科俊,邹国锋. 基于子模式的 Gabor 特征融合的单样本人脸识别[J]. 模式识别与人工智能, 2013, 26(1): 50-56.
WANG Kejun, ZOU Guofeng. A sub-pattern Gabor features fusion method for single sample face recognition [J]. Pattern Recognition and Artificial Intelligence, 2013, 26(1): 50-56.
- [14] GU W, XIANG C, VENKATESH Y V, et al. Facial expression recognition using radial encoding of local Gabor features and classifier synthesis [J]. Pattern Recognition, 2012, 45(1): 80-91.
- [15] KREMIC E, SUBASI A, HAJDAREVI K. Face recognition implementation for client server mobile application using PCA [C]//Information Technology Interfaces (ITI), Proceedings of the ITI 2012 34th International Conference on. IEEE, 2012: 435-440.
- [16] LEE J, SEO Y H. An efficient head pose determination and its application to face recognition using multi-pose face DB and SVM [C]//Ninth International Conference on Broadband and Wireless Computing, Communication and Applications. IEEE, 2014: 527-531.
- [17] LIU Y L, YU L. The applications of wavelet transform and fast PCA and SVM in human face identification [C]//International Conference on Computing and Networking Technology. 2012: 175-179.
- [18] ZHANG E, LI Y, ZHANG F. A single training sample face recognition algorithm based on sample extension [C]//International Conference on Advanced Computational Intelligence. 2013: 324-327.
- [19] 中国研究生智慧城市技术与创意设计大赛. 人脸检测与识别 [EB/OL]. <http://www.smartcity-competition.com.cn>. (责任编辑 周 蓓)