

# 基于多尺度中心化二值模式的人脸表情识别

付晓峰, 韦巍

(浙江大学 电气工程学院, 浙江 杭州 310027)

**摘要:** 现有局部二值模式(LBP)算子存在不足: 产生的直方图维数过长、鉴别力不高、对噪声反应敏感. 针对此类问题, 提出中心化二值模式(CBP)算子, 其优点: 1) 通过比较邻域中近邻点对, 大大降低了直方图维数; 2) 考虑中心像素点的作用并赋予其最高权重, 实现鉴别力的提高; 3) 改变LBP算子的符号函数, 明显减弱白噪声对图像的影响. 此外, 为提高识别率, 将多尺度CBP(MCBP)直方图作为人脸表征. 为增强算法对表情图像中细小变形的鲁棒性, 引入图像欧式距离(IMED)并将其嵌入MCBP方法. 在JAFPE和Cohn-Kanade表情库的实验结果表明: 所提方法优于其它表情识别方法, IMED可增强MCBP的表情识别能力.

**关键词:** 中心化二值模式; 多尺度; 图像欧式距离; 表情识别

**中图分类号:** TP391 **文献标识码:** A

## Facial expression recognition based on multi-scale centralized binary pattern

FU Xiao-feng, WEI Wei

(Institute of Electrical Engineering, Zhejiang University, Hangzhou Zhejiang 310027, China)

**Abstract:** The existing local binary pattern(LBP) operators have disadvantages of long histograms produced by them, low discrimination and high sensitivities to noise. To deal with these problems, we propose the centralized binary pattern(CBP) operator. The CBP operator has several advantages: 1) It significantly reduces the histogram dimensionality by comparing pairs of neighbors in the neighborhood; 2) It enhances the discrimination by emphasizing the effect of the center pixel point through giving it the largest weight; 3) It decreases the white noise influence on face images by modifying the sign function of the existing LBP operator. Moreover, the multi-scale CBP(MCBP) histogram is used as face representation to increase the recognition accuracy. Furthermore, in order to improve the robustness to small deformation of expressional images, the image Euclidean distance(IMED) is introduced and embedded in MCBP. Experiments on JAFFE and Cohn-Kanade facial expression databases demonstrate that the proposed method outperforms other modern approaches and show that IMED can enhance the performance of MCBP in facial expression recognition.

**Key words:** centralized binary pattern; multi-scale; image Euclidean distance; facial expression recognition

### 1 引言(Introduction)

随着人机交互(HCI)系统<sup>[1,2]</sup>的发展,近些年来人脸表情识别越来越受到关注. 目前,用于人脸表情识别的几大流行方法有: 基于Gabor小波<sup>[3]</sup>、基于活动外观模型(AAM)<sup>[4]</sup>、基于局部二值模式(LBP)<sup>[5,6]</sup>等. 其中, 基于Gabor小波的方法非常耗时间和耗内存, 不适合于建立快速高效的人机交互系统. 基于AAM的方法虽然可获得更为可靠的表情特征, 却具有计算复杂、初始参数难以获取等缺点. 基于LBP的方法以其计算简便、抗光线干扰、判别力强等优点正引起人们的高度关注. 相比于Gabor小波特征, LBP特征能够更快地从人脸图像中提取出来,

而且维数小很多, 同时保留了人脸的有效信息.

然而, 现有LBP算子具有以下几点不尽人意的地方: 1) 由它们产生的直方图维数过长, 导致识别速度低下(尤其是在大型人脸库上); 2) 由于不考虑LBP算子中心像素点的作用, 在特定情况下它们失去某些局部结构信息; 3) 它们产生的二值数据对噪声反应敏感. 因此, 本文提出中心化二值模式(CBP)算子, 它克服了现有LBP算子的缺点. 具体地, 1) 通过比较环形邻域中近邻点对, CBP算子大大降低直方图维数; 2) 相比于现有LBP算子, CBP算子的鉴别力大大提高, 因为它把中心像素点考虑进来并赋予其最大权重; 3) CBP算子对噪声不敏感原因在于它改变

了现有LBP算子的符号函数. 此外, 为提高识别率本文使用多尺度中心化二值模式(MCBP)直方图作为人脸表征. 另外, 为增强算法对表情图像中存在的细小变形的鲁棒性, 引入图像欧式距离(IMED)<sup>[7]</sup>并将之嵌入MCBP方法, 最终提取到的表情特征具有优点: 特征维数大大降低、更强的鉴别力、抗噪声能力强、不易受细小变形的干扰. 最后, 采用Chi square statistic<sup>[8]</sup>和最近邻分类器完成人脸表情分类. 在两个知名人脸表情库上的实验结果表明: 嵌有IMED的MCBP方法(简称MCBP-IMED)优于其它用于人脸表情识别的方法, 而且IMED有助于MCBP方法在人脸表情识别中的应用.

## 2 多尺度中心化二值模式(Multi-scale centralized binary pattern)

### 2.1 中心化二值模式(Centralized binary pattern)

传统LBP算子<sup>[9]</sup>标记图像中每个像素点得到一个二进制数, 将此二进制数转换为十进制数即此像素点的LBP码(见式(1)), 把每幅图像中所有像素点的LBP码所构成的直方图看作该图像的LBP特征.

$$\text{LBP}(M, R) = \sum_{m=0}^{M-1} s(g_m - g_c) 2^m. \quad (1)$$

其中:  $M, R$ 分别指代近邻点个数、半径,  $g_c$ 是中心像素点,  $g_m (m = 0, \dots, M-1)$ 是 $g_c$ 的近邻点, 符号函数 $s(x)$ 的定义如下:

$$s(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (2)$$

均匀LBP<sup>[9]</sup>与传统LBP不同之处是: 均匀LBP将至多存在2次“1到0”或者“0到1”转变的二进制数视为均匀模式, 而其它不属于此模式的二进制数都被看作一个标记. 可见, 均匀LBP是 $2^M$ 个传统LBP的子集. 此外, 研究人员发现, 为有效表征人脸, 提取的特征应该包含空间信息, 因此把人脸分为 $k$ 个子区域 $B_1, B_2, \dots, B_k$ (见图1), 在各子区域上提取LBP特征. 图1(a)的划分是根据人脸某些部位的信息对表情识别贡献较大如眼睛、嘴部, 它们被赋予较大的权值, 而其它部位的贡献较小则被赋予较小的权值, 权值的赋予可见图1(b). 黑色区域对应的权值为1.0, 灰色为2.0, 白色为4.0. 其中各个区域的尺寸是根据不同人在不同表情时面部器官的大致范围确定.

通过观察可知, 子区域尺寸越大识别率越低, 原因在于更多的空间信息被丢失掉, 因此子区域的数目 $k$ 不可能太小. 假设二值模式直方图bin的数目为 $n$ , 则最终特征向量长度为 $n \times k$ , 采用如此长度的特征向量会大大降低识别速度(尤其是在大型人脸库上). 因此, 本文提出对二值模式进行降维的措

施. 如图2所示, 如果“近邻点对”的连线穿过中心像素点则比较该对近邻点的灰度值. 对于没有落在像素点中心的近邻点, 采用双线性插值法估计其灰度值. 经分析可知, 图2的中心化二值模式(CBP)优于传统LBP表现在两个方面: 1) 维数的降低; 2) 相比LBP, CBP捕捉到更好的梯度信息, 有利于提高识别率.

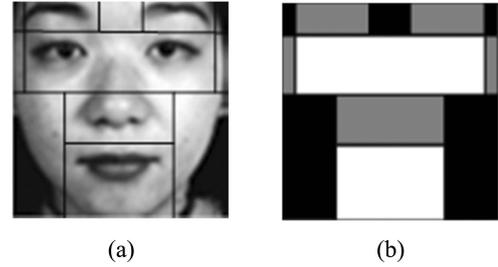


图1 一张人脸图像被分割为12个子区域及其权值集合  
Fig. 1 A face image divided into 12 sub-regions and the weights set for weighted dissimilarity measure

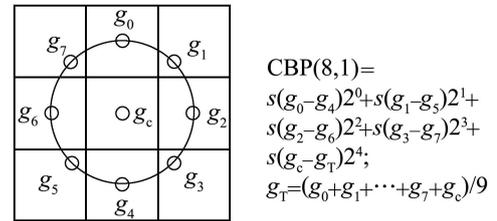


图2 CBP(8,1)算子

Fig. 2 CBP(8,1) operator

此外, 在某些特定情况下, 传统LBP丢失了局部结构信息. 以LBP(8,1)为例, 采用式(1)仅仅可得到512种二值模式中的256种, 原因在于中心像素点 $g_c$ 一直被置为0. 然而, 在大多数情况下中心像素点要比其近邻点提供更重要的信息, 因此要把 $g_c$ 考虑进来并赋予其最高权值.

另外, 传统LBP存在一个严重的问题即由式(1)得到的二值数据对噪声反应敏感. 因此, 在实际应用中, 图像中存在的噪声严重限制了LBP算子提取特征的效果. 原因在于LBP算子考虑的是像素点间的关系, 而此关系是相互的, 即如果 $s(g_a - g_b) = 0$ 那么 $s(g_b - g_a) = 1$ . 因此为降低白噪声对图像的影响, 本文将符号函数更改如下:

$$s(x) = \begin{cases} 1, & |x| \geq C, \\ 0, & |x| < C. \end{cases} \quad (3)$$

其中 $C$ 是阈值常数.

综上, 本文提出的中心化二值模式算子, 它可以由式(4)表示. 对“中心化二值模式”的命名有两个缘由: 1) 要比较的“近邻点对”必须满足其连线穿

过中心像素点 $g_c$ ; 2) 将中心像素点考虑进来后, 其灰度值与所有近邻点及 $g_c$ 之和的平均值进行比较, 而且 $g_c$ 拥有最高权重. 可见, 所提出的模式是中心化了的二值模式, 因此命名为“中心化二值模式”简称CBP.

$$\text{CBP}(M, R) = \sum_{m=0}^{(M/2)-1} s(g_m - g_{m+(M/2)}) 2^m + s\left(g_c - \frac{1}{M+1} \left(\sum_{m=0}^{M-1} g_m + g_c\right)\right) 2^{\frac{M}{2}}. \quad (4)$$

其中符号函数 $s(x)$ 如式(3)所示. 由式(4)可见, CBP算子将中心像素点考虑进来且赋予其最高权重的操作增强了中心像素点的作用而且有益于提高鉴别能力. 除此之外, CBP算子产生的直方图维数远小于传统LBP算子产生的直方图维数. 以 $(M=8, R=2)$ 为例, 传统LBP直方图维数为256, 均匀LBP为59, CBP为32.  $M$ 值越大, CBP在维数方面的优势越明显. 此外, 不同于传统LBP要比较近邻点与中心像素点之间的差异, CBP通过比较“近邻点对”之间的差异更好地捕捉到梯度信息, 大大提高CBP的鉴别能力. 另外, 由于式(3)的符号函数, 从图像中提取到的CBP特征对白噪声反应不敏感. 总之, 相比于传统LBP, CBP具有优势: 维数大大降低、更强的鉴别能力、对白噪声反应不敏感.

## 2.2 多尺度中心化二值模式(Multi-scale centralized binary pattern)

多分辨率表征已经应用于纹理分类, 结果显示: 相比于单尺度LBP, 多分辨率表征可取得更高的准确率<sup>[9]</sup>. 为提高人脸表情识别率, 本文将CBP拓展为多尺度中心化二值模式(MCBP), 通过改变参量 $R$ 以及合并多个CBP算子的信息可获得MCBP. 具体地, 根据图1将人脸图像划分为 $k$ 个子区域, 在每个子区域上提取多尺度CBP直方图并求其和值 $h_j$  ( $j=1, \dots, k$ ), 将 $h_j$ 作为子区域特征, 如式(5)所示.

$$h_j = \sum_{i=1}^t \text{CBPH}(M, R_i), \quad j = 1, \dots, k. \quad (5)$$

其中:  $h_j$ 是第 $j$ 个子区域特征,  $R_i$ 是 $t$ 个变化的半径参量中的第 $i$ 个,  $\text{CBPH}(M, R_i)$ 指代由算子 $\text{CBP}(M, R_i)$ 产生的直方图. 由于各种尺度的CBP算子具有相同的参量 $M$ , 所以各个CBP直方图维数相同则可以相加求和. 如此操作的好处: 所得子区域特征维数小于把各种尺度CBP直方图串接起来构成的子区域特征维数, 而且保留了各种尺度直方图信息. 最后,  $k$ 个子区域特征 $h_j$  ( $j=1, \dots, k$ ) 串接起来构成的特征向量 $\mathbf{kf}$ 作为人脸表征.  $\mathbf{kf}$ 本质上就是人脸图像的多尺度CBP直方图.

## 3 图像欧式距离(Image Euclidean distance)

Cohn-Kanade表情库<sup>[10]</sup>部分图像(如图3所示, ©Jeffrey Cohn) 显示出一个人的同种表情可以有很多种面貌. 例如, 一个人的高兴表情可以是微笑、大笑等任意程度, 如此会在属于同类表情的人脸图像中出现细小的变形(deformation). 因此, 在表情识别中有必要采取对细小的变形不敏感的措施, 而图像欧式距离(IMED)<sup>[7]</sup>正是具有此种特性的方法. 除此特性之外, IMED还具有两种优点: 1) 计算的简便性; 2) 它可以很有效地被嵌入到图像识别方法中. 本文把IMED嵌入到MCBP方法中.



图3 Cohn-Kanade表情库部分图像  
Fig. 3 Some examples from Cohn-Kanade facial expression database

尺寸为 $P \times Q$ 的图像可以被向量化为位于空间 $R^{PQ}$ 中的向量 $\mathbf{x}$ , 向量化图像 $\mathbf{x}$ 和 $\mathbf{y}$ 间的IMED可以按式(6)计算, 有关IMED的更多细节详见文献[7].

$$d_{IM}^2(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{PQ} g_{ij} (x^i - y^i)(x^j - y^j) = (\mathbf{x} - \mathbf{y})^T G (\mathbf{x} - \mathbf{y}). \quad (6)$$

将“ $G^{1/2}$ 转换”应用于图像 $\mathbf{x}$ 和 $\mathbf{y}$ 可得转换后图像 $\mathbf{u}, \mathbf{v}$ , 如此操作可避免不必要的重复性计算IMED, 如式(7)所示.

$$\mathbf{u} = G^{1/2} \mathbf{x}, \quad \mathbf{v} = G^{1/2} \mathbf{y}. \quad (7)$$

则 $\mathbf{x}$ 和 $\mathbf{y}$ 间的IMED可以简化为 $\mathbf{u}$ 和 $\mathbf{v}$ 间的传统欧式距离:

$$(\mathbf{x} - \mathbf{y})^T G (\mathbf{x} - \mathbf{y}) = (\mathbf{u} - \mathbf{v})^T (\mathbf{u} - \mathbf{v}). \quad (8)$$

因此, 本文将转换后图像作为MCBP的输入就自动把IMED嵌入到MCBP方法中. 此外, 人脸表情库上的大量实验表明, 相比于传统欧式距离(TED), IMED可使类间距离更大而使类内距离更小. 此即, 在这类人脸库上, 由TED计算得的类间距离小于由IMED计算得的类间距离, 而由TED计算得的类内距离却大于由IMED计算得的类内距离. 显然, IMED比TED给出了更好的图像距离关系, 此特性有益于人脸表情分类.

总之, 嵌有IMED的MCBP(简称MCBP-IMED)方法提取出的特征具有优点: 维数大大降低、很强的鉴别能力、对噪声不敏感、不易受细小变形的

干扰.

#### 4 实验结果及分析(Experimental result and analysis)

本文在JAFFE库<sup>[11]</sup>和Cohn-Kanade表情库<sup>[10]</sup>上对算法MCBP-IMED与均匀LBP, MCBP, Gabor小波进行比较验证. 算法MCBP-IMED与MCBP唯一的区别是: 前者需要将IMED嵌入MCBP方法中而后者不需要此步骤.

为检测算法的广泛性(generality)效果, 采用10-fold交互检验(cross validation)方案<sup>[5]</sup>. 首先, 每张图像要经过剪切、调整尺寸为 $64 \times 64$ 大小以及消除光线干扰的预处理. 在均匀LBP方法中, 选取均匀LBP(8,2)算子从人脸图像各子区域(如图1所示)提取均匀LBP直方图, 并串接起来构成人脸表征. 最后采用Chi square statistic和最近邻分类器完成人脸表情分类. 在MCBP-IMED与MCBP方法中选取3种CBP算子: CBP(8,1), CBP(8,2), CBP(8,3)在人脸每个子区域上提取多尺度CBP直方图并求其和值作为子区域特征. 将各个子区域特征(用向量表示)串接起来构成人脸表征. 最后采用Chi square statistic和最近邻分类器完成表情分类. 对于IMED, 本文选择Gaussian函数构建量度系数 $g_{ij}$ <sup>[7]</sup>. 在Gabor小波方法中, 选取与文献[3]中相同的二维Gabor小波, 并采取最近邻分类器完成表情识别. 本文实验是采用MATLAB7.1软件在PentiumIV3.73GHz, 2G内存PC机上进行的.

##### 4.1 JAFFE库(JAFFE database)

从JAFFE库中选取196张图像, 算法MCBP-IMED与均匀LBP, MCBP, Gabor小波的比较结果如表1所示.

表1 JAFFE库上 $64 \times 64$ 分辨率时不同方法的识别结果比较

Table 1 Comparison of different methods with resolution level  $64 \times 64$  for the images of JAFFE database

方法	识别率/%	识别一个样本所用时间/s
Gabor小波	83.51	23.5671
均匀LBP	85.78	0.2594
MCBP	92.84	0.1453
MCBP-IMED	94.83	0.2069

由表1可见, MCBP方法的识别率超过均匀LBP, Gabor小波方法, 而它识别一个样本所用时间却小于后两者. 此现象揭示出在表情特征提取和识别速度两个方面, MCBP都优于后两者. 原因是: 1) 不像LBP比较近邻点与中心像素点间的差异,

MCBP通过比较“近邻点对”间的差异更好地捕捉到梯度信息, 此举有利于MCBP鉴别力的提高; 2) 不像LBP没有考虑中心像素点的作用, MCBP却充分考虑它的作用并给它分配最高权重, 此项操作大大提高了MCBP的鉴别能力; 3) MCBP具有多尺度特性; 4) Gabor特征维数是163840( $=40 \times 64 \times 64$ ), LBP特征维数是708( $=59 \times 12$ ), 而MCBP特征维数只有384( $=32 \times 12$ ),  $M$ 值越大, MCBP特征在维数方面的优势越明显. 另外, 方法MCBP-IMED优于MCBP, 可见IMED可增强MCBP识别表情的能力, 原因在于IMED对表情图像中的细小变形不敏感, 而且IMED给出的较好的图像距离关系有益于表情分类. 总之, 以识别率和识别一个样本所用时间两个指标综合评判以上几种方法可见MCBP-IMED达到最佳效果.

由于在许多应用场合, 图像的分辨率会受多种因素的影响而不是很高, 因此在低分辨率图像上比较MCBP-IMED与其它方法尤显重要. 获取低分辨率图像的方式与以上实验获取 $64 \times 64$ 图像的方式类似, 唯一不同的是把尺寸分别调整为 $48 \times 48$ ,  $32 \times 32$ ,  $16 \times 16$ . 比较结果如图4所示.

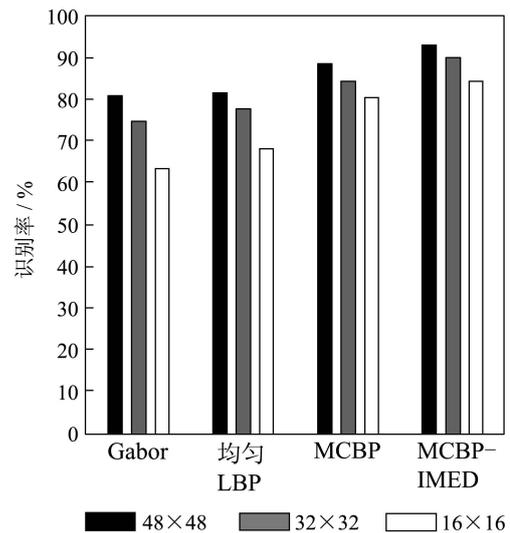


图4 不同方法在低分辨率情况下识别结果比较

Fig. 4 Comparison of different methods with low resolution levels for the images of JAFFE database

由图4可见, 在各种分辨率下方法MCBP-IMED都超过其它方法, 即使在 $16 \times 16$ 分辨率时其识别率仍可达84.13%, 显示出更好的稳定性和更高的鲁棒性.

##### 4.2 Cohn-Kanade表情库(Cohn-Kanade expressive database)

从Cohn-Kanade表情库中选取1820张图像, 为比较各种算法对噪声的抗扰能力, 在每张图像中加入

较弱的Gaussian噪声. 最终结果如表2所示.

由表2可见, MCBP方法的抗噪声能力比均匀LBP方法强, 原因是: 由于式(3)的符号函数, 从图像中提取到的MCBP特征对白噪声反应不敏感, 并阐明了一个事实—图像中存在的噪声严重限制了LBP算子提取特征的效果. 另外, 实验中发现: 由于均匀LBP产生的直方图维数过长, 导致在此大型人脸库上识别速度低下. 由表2仍可见, 方法MCBP-IMED优于MCBP, 再次揭示出IMED增强了MCBP在表情识别方面的能力.

表2 有噪声时不同方法在Cohn-Kanade表情库上识别结果比较

Table 2 Performance comparison of different methods on Cohn-Kanade database under noisy circumstance

方法	识别率/%	识别一个样本所用时间/s
Gabor小波	83.20	35.8890
均匀LBP	84.52	3.2284
MCBP	93.89	1.8516
MCBP-IMED	94.91	2.7092

## 5 结论(Conclusion)

本文提出一种识别人脸表情的方法. 针对现有LBP存在的不足, 提出多尺度中心化二值模式(MCBP). 此外, 为增强算法对表情图像中存在的细小变形的鲁棒性, 本文引入图像欧式距离并将它嵌入到MCBP方法中. 实验结果不仅表明本文方法MCBP-IMED优于其它表情识别方法而且揭示出IMED可增强MCBP识别表情的能力.

## 参考文献(References):

- [1] 刘海波, 顾国昌, 沈晶, 等. 智能机器人神经心理模型[J]. 控制理论与应用, 2006, 23(2): 175-180.  
(LIU Haibo, GU Guochang, SHEN Jing, et al. Neuropsychological model for intelligent robot[J]. *Control Theory & Applications*, 2006, 23(2): 175-180.)
- [2] 宋亦旭, 贾培发. 人工情感及其应用[J]. 控制理论与应用, 2004, 21(2): 315-320.

- (SONG Yixu, JIA Peifa. Artificial emotion and its applications[J]. *Control Theory & Applications*, 2004, 21(2): 315-320.)
- [3] 姚伟, 孙正兴, 张岩. 面向脸部表情识别的Gabor特征选择方法[J]. 计算机辅助设计与图形学学报, 2008, 20(1): 79-84.  
(YAO Wei, SUN Zhengxing, ZHANG Yan. Optimal Gabor feature for facial expression recognition[J]. *Journal of Computer-aided Design & Computer Graphics*, 2008, 20(1): 79-84.)
- [4] 左坤隆, 刘文耀. 基于活动外观模型的人脸表情分析与识别[J]. 光子·激光, 2004, 15(7): 853-857.  
(ZUO Kunlong, LIU Wenyao. Facial expression recognition using active appearance models[J]. *Journal of Optoelectronics · Laser*, 2004, 15(7): 853-857.)
- [5] SHAN C, GONG S, MCOWAN P W. Robust facial expression recognition using local binary patterns[C] // *IEEE International Conference on Image Processing*. Genova, Italy: IEEE, 2005, 2.
- [6] 孙宁, 冀贞海, 邹采荣, 等. 基于2维偏最小二乘法的图像局部特征提取及其在面部表情识别中的应用[J]. 中国图象图形学报, 2007, 12(5): 847-853.  
(SUN Ning, JI Zhenhai, ZOU Cairong, et al. Image local feature extraction method based on two-dimensional partial least square and its application in facial expression recognition[J]. *Journal of Image and Graphics*, 2007, 12(5): 847-853.)
- [7] WANG L, ZHANG Y, FENG J. On the Euclidean distance of image[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(8): 1334-1339.
- [8] AHONEN T, HADID A, PIETIKÄINEN M. Face description with local binary patterns: application to face recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(12): 2037-2041.
- [9] OJALA T, PIETIKÄINEN M, MÄENPÄÄ T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971-987.
- [10] KANADE T, COHN J F, TIAN Y. Comprehensive database for facial expression analysis[C] // *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition*. Grenoble, France: IEEE, 2000: 46-53.
- [11] MICHAEL J L, SHIGERU A, MIYUKI K, et al. Coding facial expressions with Gabor wavelets[C] // *Third IEEE International Conference on Automatic Face and Gesture Recognition*. Nara, Japan: IEEE, 1998: 200-205.

## 作者简介:

付晓峰 (1981—), 女, 博士研究生, 研究方向为智能控制、智能机器人、模式识别等, E-mail: saqieer98@yahoo.com.cn;

韦巍 (1964—), 男, 教授, 博士生导师, 浙江大学电气学院副院长, 研究方向为智能控制、智能机器人等.