

文章编号:1007-2780(2013)03-0440-06

B2DPCA 和 ELM 人脸识别算法研究

李定珍,郭建昌

(南阳理工学院 电子与电气工程学院,河南 南阳 473004, E-mail:lidingzhen136@163.com)

摘 要: 提出一种新型、高效的基于 B2DPCA(双向二维主成分分析)和 ELM(极端学习机)的人脸识别算法,该算法是根据曲波变换分解人脸图像和一种改进的降维技术,通过 B2DPCA 生成识别特征集来训练和测试 ELM 分类器,提高识别精度。通过大量实验,并把实验结果与现存技术进行比较,结果表明 B2DPCA+ELM 算法有效地提高了识别准确率,并降低了对原型数量的依赖。将来有望能把局部特征和基于曲波分解的全局信息结合起来应用到识别精度和分类速度上。

关 键 词: 人脸识别;双向二维主成分分析;极端学习机;降维技术;识别准确率

中图分类号: TP391.4 **文献标识码:** A **DOI:** 10.3788/YJYXS20132803.0440

Face Recognition Algorithm Based on B2DPCA and ELM

LI Ding-zhen, GUO Jian-chang

(Institute of Electronics and Electrical Engineering, Nanyang Institute of Technology,
Nanyang 473004, China, E-mail:lidingzhen136@163.com)

Abstract: A new human face recognition algorithm was proposed based on B2DPCA (bidirectional two dimensional principal component analysis) and ELM (extreme learning machine). This method was based on curvelet image decomposition of human faces and an improved dimensionality reduction technique. Discriminative feature sets were generated by using B2DPCA to train and test ELM classifier. The recognition accuracy can be improved by using this method. Extensive experiments were performed by using databases and results were compared with state of the existing techniques. The results showed recognition accuracy and minimal dependence on the number of prototypes were significantly improved by using B2DPCA and ELM algorithm. The local characteristics and global information based on curvelet decomposition are expected to apply to the recognition accuracy and speed of classification in the future.

Key words: human face recognition; B2DPCA; extreme learning machine; dimensionality reduction technique; recognition accuracy

1 引 言

人脸识别系统必须能可靠地完成人脸检测、特征提取和人脸识别等功能,分为基于部件识别和基于面部识别两部分^[1]。当人脸大小和形状发

生微小变化时将会引起提取特征发生变化,因此可靠的面部特征提取是一个极其复杂的任务。所以,研究人员提出把人脸视为二维强度模式的基于面部识别系统,并通过检测和匹配统计特性来实现。

收稿日期:2012-09-13; 修订日期:2012-11-27

基金项目:河南省教育厅科技攻关项目(No.12B510024)

作者简介:李定珍(1972—),女,河南南阳人,硕士研究生,副教授,主要从事图像处理、智能仪器测试及自动控制等研究。

Kirby 等利用 PCA(主成分分析法)把人脸描述成加权特征向量的线性组合,但 PCA 的人脸识别系统存在辨别能力差和计算量过大的缺点。J. Lu 等为改进人脸识别,提出用 LDA(线性判别分析法)使类间散布矩阵和类内散布矩阵比例达到最大值^[2]。J. M. Fellus 提出 EP(evolutionary pursuit)法和弹性束图匹配法(EBGH-elastic bunch graph matching)以期产生最优轴投影。因此为了改善高维图像的人脸识别系统性能,近年来研究的基于曲波变换的人脸识别算法有基于曲波变换的 PCA 法、LDA 法和 PCA+LDA 法^[3-5]。但是他们都有其自身局限性:比如对视点的变化、原型数量、分类速度等有较大敏感性。

在上述研究基础上,作者提出把曲波变换与 B2DPCA(双向二维主成分分析)和 ELM(极端学习机)相结合来识别人脸,用以消除原有算法中的固有缺陷。

2 图像识别算法的分析与选择

2.1 曲波变换

小波变换在传达准确的时间和空间信息上是一个显著的多分辨率分析工具,但小波变换存在很多问题,针对这些问题 Candes 和 Donoho 在 1999 年提出曲波变换理论^[6],并构造了曲波的紧框架,对于具有光滑奇异性曲线的目标函数,曲波提供了稳定、高效和近乎最优的表示。曲波变换直接以边缘为基本表示元素,曲波变换是各向异性的,具有很强的方向性,非常有利于图像边缘的高效表示^[7]。并符合生理学研究所指出的“最优”图像表示方法应具有的特征。

第一代曲波变换,该理论由脊波理论衍生而来。曲波变换和小波变换相同之处:都是多尺度变换,以刻度和空间位置参数为构成基础。但曲波变换有方向参数,曲波金字塔包含一种方向特征的基本原理。此外曲波变换基于各向异性尺度法则。第二代曲波与第一代在构造上已完全不同,第一代的构造思想是通过足够小的分块将曲线近似地当作每个分块中的直线来看待,然后利用局部的脊波分析其特性。而第二代曲波和脊波理论并没有关系,实现过程也无需用到脊波。本文主要利用曲波变换分解图像。

2.2 B2DPCA-双向二维主成分分析法

Karhunen-Loeve 展开式,又被称作主成分分

析法(PCA),是一种被广泛应用在模式识别和压缩的数据表示技术。直接用于人脸识别时,需将图像矩阵转换成向量,导致求解高阶矩阵计算量大。Yang 等人提出了 2DPCA(二维主成分分析)图像表示法^[8],与 PCA 不同,2DPCA 是基于二维矩阵而不是一维向量。因此,图像的矩阵在进行特征提取前不需要矢量化,而是使用原始图像矩阵直接计算图像协方差矩阵。

设 $\mathbf{X}$ 表示一个 q 维的单位化列向量,2DPCA 的设计思想是将 $p \times q$ 图像的矩阵 $\mathbf{A}$ 经线性变换 $\mathbf{Y} = \mathbf{A}\mathbf{X}$ 直接投影到 X 上,得到一个 A 的投影特征向量 $\mathbf{Y}$ 。可用投影样本的总体散布情况来衡量投影向量 $\mathbf{X}$ 对样本的识别能力。投影样本的总体散布可用投影特征向量的协方差矩阵的迹进行描述。

这样,可采用判定准则 $J(\mathbf{X}) = \text{tr}(\mathbf{S}_x)$ 。其中 $\mathbf{S}_x$ 表示训练样本的投影特征向量的协方差矩阵, $\text{tr}(\mathbf{S}_x)$ 表示 $\mathbf{S}_x$ 的迹。找出判定准则的最大值,其意义是找到一个使所有样本都能被涉及到的投影方向 X ,以使这些投影样本的总体散布被最大化。

协方差矩阵 $\mathbf{S}_x$ 可表示为式(1):

$$\mathbf{S}_x = E[\mathbf{Y} - E(\mathbf{Y})][\mathbf{Y} - E(\mathbf{Y})]^T = E[(\mathbf{A} - E\mathbf{A})\mathbf{X}][(\mathbf{A} - E\mathbf{A})\mathbf{X}]^T \quad (1)$$

所以,

$$\text{tr}(\mathbf{S}_x) = \mathbf{X}^T [E(\mathbf{A} - E\mathbf{A})^T (\mathbf{A} - E\mathbf{A})] \mathbf{X} \quad (2)$$

定义一个图像的协方差矩阵 G_i 。

$$G_i = E[(\mathbf{A} - E\mathbf{A})^T (\mathbf{A} - E\mathbf{A})] \quad (3)$$

从定义可证明 G_i 是一个非负的 $q \times q$ 矩阵。使用训练图像样本可直接计算 G_i 。假设共有 M 个训练图像样本,第 α 个训练图像样本用一个 $P \times q$ 矩阵 $A_\alpha (\alpha = 1, 2, \dots, M)$ 表示,所有训练图像样本的平均图像用 A 表示。则 G_i 可由下式计算得到:

$$G_i = \frac{1}{M} \sum_{\alpha=1}^M (A_\alpha - A)^T (A_\alpha - A) \quad (4)$$

则,

$$J(\mathbf{X}) = \mathbf{X}^T G_i \mathbf{X} \quad (5)$$

使判定准则 $J(\mathbf{X})$ 取最大值的线性向量 $\mathbf{X}_{\text{opt}}$ 称作最佳投影轴。一般情况下,只有一个最佳投影轴是不够的,通常需要选择满足正交条件和使 $J(\mathbf{X})$ 取最大值的一簇投影轴 $X_1, X_2, \dots, X_M$ 。实际上,最佳投影轴是 G_i 对应于前 M 个最大本征值的正交本征向量。

2DPCA 的实质是对图像矩阵按行进行图像压缩抽取特征,消除了图像列的相关性,因此基于 2DPCA 识别的局限性,即是它的可操作性只能沿着行方向,其特征数量仍然较大,影响分类速度。针对这一问题,提出了采用双向压缩的二维主成分分析消除图像行间和列间的相关性,Zhang 和 Zhou 就是在基于假设训练图像零均值的基础上提出了 B2DPCA 法^[9]。因此,可用行/列图像矢量的外积计算图像的协方差矩阵。利用式(4)计算出 G_{tRow} 和 G_{tCol} ,先按照 A_a 和 A 的行向量计算,接着用同样方法对其列向量计算。对 G_{tRow} 和 G_{tCol} 的最佳投影轴进行取值,分别记作 X_{1opt} 和 Z_{1opt} , G_t 和 G_{tRow} 都是基于行计算的,因此最佳投影轴 X_{opt} 和 X_{1opt} 是相似的,通过变换把 A_a 表示成 $Z_{1opt}^T A_a X_{1opt}$ 。

为保证行列关系和生成区别特征集,降维算法沿行列是独立完成的,B2DPCA+ELM 是利用最佳投影轴生成图像协方差矩阵,并对其进一步优化。一旦最佳投影轴把 X_{opta} 计算出来,图像的矩阵 A_a 沿其列向量降维,利用式(6)得到图像集 A_β ,把新生成的图像集看作一个新数据库,计算得出新的图像协方差矩阵 G_β 和新的最佳投影轴 $X_{opt\beta}$ 。将 $X_{opt\beta}^T$ 左乘 A_β 得到输出矩阵 A_θ ,如式(7)所示。

$$A_\beta = A_a X_{opta} \quad (6)$$

$$A_\theta = X_{opt\beta}^T A_\beta \quad (7)$$

Kong 等人将 2DPCA 进行改进^[10],提出了 B2DPCA 算法来降维。B2DPCA 算法的主要步骤:

INPUT:输入图像的矩阵 $A_{N \times N}$, $N \in Z^+$,
OUTPUT:输出 $M \times M$ 矩阵 A_θ , $N \in Z^+$ 。利用式(4)计算 A_a 图像的协方差矩阵 G_a ;计算 $J(X) = X^T G_a X$;最佳投影轴 $X_{opta} = \{X_1, X_2, \dots, X_M\}$, X_i 表示正交向量;沿列降维得到 $A_\beta = A_a X_{opta}$;计算 A_β 的图像协方差矩阵 G_β ;同式(2)计算 $X_{opt\beta}$;通过式(7)进行行降维。

2.3 ELM-极端学习机

传统前馈神经网络多采用梯度下降算法调整权值,缺点是学习速度较慢、泛化性能差,为了解决这些缺陷,Huang 等人提出了极端学习机算法(ELM-Extreme learning machine)^[11],该算法对单隐层神经网络的输入权值和隐层节点偏移量进行随机赋值,并且只通过一步计算即可解析求出

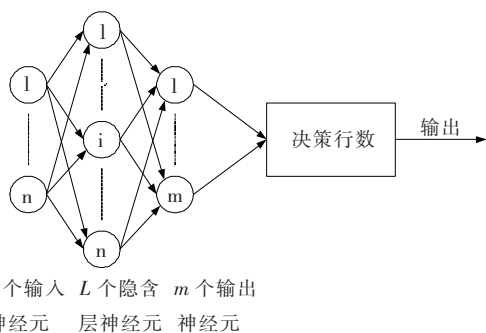


图 1 极端学习机分类器结构

Fig. 1 Extreme learning machine classifier structure

网络输出权值,如图 1 所示。

考虑到 N 个不同样本集 (x_i, t_i) , 其中 $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$, $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m$, 可把含有 L 个隐节点和一个激活函数 $\xi(x)$ 的 ELM 建模为式(8):

$$\sum_{i=1}^L \gamma_i \xi(x_n) = \sum_{i=1}^L \gamma_i \xi_i(w_i x_n + b_i) = o_n, \quad n = 1, 2, \dots, N \quad (8)$$

其中 $w_i = [w_{i1}, w_{i2}, \dots, w_{im}]$ 表示输入, $\gamma_i = [\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{im}]$ 表示隐含层权值向量, b_i 是隐节点偏移量。极端学习机以最小的误差,近似于 N 个样本,如式(9)。式(9)可表示成 $\delta \gamma = \tau$, $\delta = (w_1, \dots, w_L, b_1, \dots, b_L, x_1, \dots, x_N)$ 。

$$\sum_{i=1}^L r_i \xi_i(w_i x_n + b_i) = t_n, n = 1, 2, \dots, N \quad (9)$$

δ 的第 i 列就表示输入为 $x_1, x_2, \dots, x_N$ 的第 i 层隐节点的输出。如果激活函数 $\xi(x)$ 是无穷可微的,隐节点数 $L \ll N$,极端学习机的训练需要把误差函数 E (式 10)最小化。

$$E = \sum_{n=1}^N \left(\sum_{i=1}^L \gamma_i \xi_i(w_i x_n + b_i) - \tau_n \right)^2 \Rightarrow E = \|\delta \gamma - \tau\| \quad (10)$$

经典神经网络 δ 是利用梯度下降算法调整的^[12],在学习效率为 ρ 情况下,把输入权值 w_i , 隐含层权值 γ_i , 隐节点偏移量 b_i 进行迭代调整。 ρ 值过小将会使学习算法收敛速度过慢,而 ρ 值过大又会导致不稳定,为解决这一问题,ELM 采用了极小范数最小二乘法。因此,网络的输入权值和偏移量可直接随机赋值而不必采用梯度下降算法迭代调整,而 $\delta \gamma = \tau$ 的最小二乘法可以简化这个问题。隐含层输出矩阵 δ 是一个非方阵,范数最小二乘法可以归纳为 $\gamma = \delta^* \tau$, 其中 δ^* 表示 δ 的广义逆矩阵。ELM 表示一个线性系统的最小二

乘法^[13],所以可得到一个无限小的训练误差,用式(11)表示。

$$\|\delta\gamma - \tau\| = \|\delta\delta^* \tau - \tau\| \equiv \min_{\gamma} \|\delta\gamma - \tau\|$$

(11)

3 B2DPCA+ELM 人脸识别算法

由以上分析,作者提出 B2DPCA+ELM 的人脸识别算法。该算法是用曲波变换的图像分解和使用降维后的系数来识别,利用 B2DPCA 得到的特征集来训练和测试 ELM 分类器,算法示意图如图 2 所示。

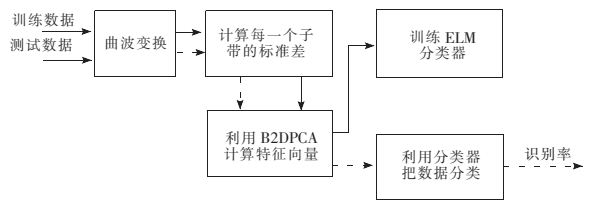


图 2 B2DPCA+ELM 的人脸识别算法示意图

Fig. 2 Schematic diagram of human face recognition algorithm

数据库中每个图像大小减小一倍转换成灰度图像,每个图像数据库随机分为训练集和测试集,这样每个目标图像有 40%~45%当作原型,剩下的图像用在测试阶段。曲波变换用来生成初始特征向量,因为它不仅在存有奇异点的高维矩阵中显示优越的性能,而且能以最小的噪声影响增强高频成分的定位。首先把输入图像尺寸重置,因为当全局信息在同一层次的时候,类似图像尺寸支持曲波特征向量的生成。此外,所有数据库中图像的曲波分解是在三尺度 8 个角方向下计算的,因此,长生 25 个不同的子带。

计算出每个子带的标准差,选取标准差最大的子带作为初始特征向量,尺寸为 $U \times V$,其中 $U \times V \ll R \times C$,其中 $R \times C$ 是输入图像尺寸。和 T. Mandal^[14] 等人选择两个子带相比,仅仅选择了一个子带,那是因为两个标准差不同子带的区别是十分明显的,所有被测试数据库标准差都有显著差别。该方法的基础是选择最大标准差的子带,这在降维时可以最大限度地节约计算成本。

B2DPCA 是用来生成独特特征集的,并且使整体框架计算复杂度最简化。2DPCA 法以一个单一的协方差矩阵分别沿行和列对图像降维,然而 B2DPCA+ELM 方法是沿与坐标轴无关的正

交方向降维。通过初始特征矩阵的初次降维提取中间特征,即沿其列选取曲波子带。然后沿中间特征的行对其进行降维以得到最终特征集,每个特征集尺寸为 $U' \times C'$,其中 $U' \times C' \ll U \times V$ 。

B2DPCA+ELM 方法能保存临近像素点之间的关键邻域信息^[14],并且能生成区别性的特征。对于每一个数据集,随机选择其中降维后的曲波特征集用于 ELM 的训练,而同一个数据集中剩余特征集用来判断框架的可分离性。

4 人脸识别实验与结果分析

4.1 B2DPCA+ELM 算法的实验

用 B2DPCA+ELM 算法,利用 ORL 和 Gtech 人脸数据库进行了大量实验。所有图像维数降低一倍,并从 RGB 转成灰度图像。在所有数据库中,每个目标图像的 40%~45%作为原型,剩下的用作测试目的。利用曲波变换在三尺度 8 个角方向下训练和测试图像,先利用 B2DPCA 得到降维后的近似曲波系数,接着用 ELM 矢量化、训练和测试。对每个数据库,每个实验做 100 次,保留平均结果。下面仅把 B2DPCA+ELM 方法的实验结果和 PCA+LAD 方法^[15]得到的结果相比较。

4.2 实验结果分析

通过 ORL 和 GTech 数据库,用不同数目主成分,采用 B2DPCA+ELM 法和基于曲波的 PCA+LAD 法实验得到的平均识别率进行比较,结果如表 1 所示。

表 1 ORL 和 GTech 人脸数据库平均识别率

Table 1 Average recognition rate of ORL and GTech face database

主成分 数量	ORL 数据库		GTech 数据库	
	PCA+ LAD 法	B2DPCA+ ELM 法	PCA+ LAD 法	B2DPCA+ ELM 法
5	79.11	93.99	71.12	89.15
10	89.17	99.55	88.31	93.52
15	94.19	98.18	90.32	97.45
20	98.32	98.76	95.63	97.12
25	97.51	98.57	96.33	96.82
30	97.52	99.92	94.66	97.01
35	98.41	99.07	96.02	97.42
40	96.68	99.27	96.12	97.71
45	97.43	99.74	94.25	97.87
50	97.51	99.84	93.24	97.98

由表 1 看出 B2DPCA+ELM 法识别准确率高于 PCA+LAD 法。B2DPCA+ELM 法对 ORL 数据库和 GTech 数据库识别率的提高是明显的,并适用于具有挑战性(视点在各个方向)的数据库。

值得一提的是,增加主成分数量不是提高识

别率的必要条件,利用人脸识别的局部信息也有可能产生更好的结果,如图 3 所示。图 3 是 ORL 数据库中原型数量(a)~(d)分别占 30%、40%、60%、70%时的识别率,其中 x 轴表示主成分的数量, y 轴表示平均识别率。

由图 3 可看出,B2DPCA+ELM 法和其他人

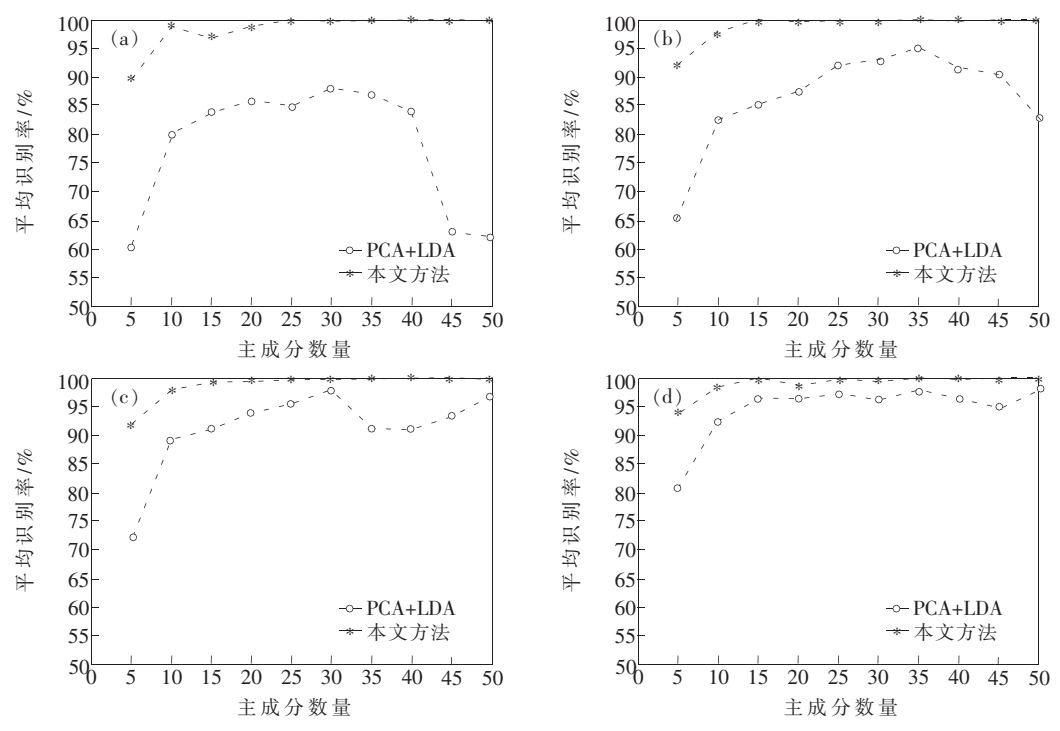


图 3 原型数量所占不同比例时的平均识别率
Fig. 3 Average recognition rate of different proportion prototype number

脸识别算法相比除了提高了识别准确率以外,另一个优点是原型的数量对于识别率影响非常小,降低了对原型数量的依赖。

5 结 论

对人脸识别算法技术进行研究,基于曲波特征空间提出了一种新型高效的 B2DPCA+ELM 的人脸识别算法。根据曲波变换来分解人脸图像,而曲波变换的主要作用是实现高维稀疏,作者利用 B2DPCA 对稀疏特征降维得到不同特征集,

最后把这些特征集作为 ELM 的输入来分析学习最优模型,从而提高识别精度。通过 ORL 和 GTech 数据库,用不同的数目主成分,进行了大量实验,并把实验数据与现存技术进行的实验结果做了比较。结果显示:B2DPCA+ELM 的人脸识别算法不仅比现存方法提高了对人脸识别准确率,而且在用以训练的原型数量上具有独立性,降低了对原型数量的依赖。将来,希望能把局部特征和基于曲波分解的全局信息结合起来应用到识别精度和分类速度上。

参 考 文 献:

[1] 张俊峰,王宇,葛哋哋. 基于 Gabor 小波变换的人脸表情识别技术 [J]. 计算机测量与控制, 2010, 18(4): 906-911.
[2] Lu J, Plataniotis K N, Venetsanopoulos A N. Face recognition using LDA-based algorithms [J]. *IEEE Transactions on Neural Networks*, 2003, 14(1): 195-200.
[3] Feng G C, Yuen P C, Dai D Q. Human face recognition using PCA on wavelet subband [J]. *J. Electronic Imaging*,

- 2000,9 (2):226-233.
- [4] Chien J T, Wu C C. Discriminant wavelet faces and nearest feature classifiers for face recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(2):1644-1649.
- [5] Mandal T, Wu Q M J, Yuan Y. Curvelet based face recognition via dimension [J]. *Reduction Elsevier Signal Processing*, 2009, 89(3):2345-2353.
- [6] Candes E J, Demanet L, Donoho D L, et al. Fast discrete curvelet transforms [J]. *Multiscale Modeling and Simulation*, 2006, 5 (3):861-899.
- [7] 吴春颖. 浅谈曲波变换的应用 [J]. 福建电脑, 2011, (10):79-80.
- [8] Yang J, Zhang D, Frangi A F, et al. Two-dimensional PCA: a new approach to appearance based face representation and recognition. [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(1):131-137.
- [9] Zhang D, Zhou Z H. (2D)2PCA: two-directional two-dimensional PCA for efficient face representation and recognition [J]. *Elsevier Neurocomputing*, 2005, 69(1):224-231.
- [10] Kong H, Wang L, Teoh E K, et al. Generalized 2D principal component analysis for face image representation and recognition [J]. *Neural Networks*, 2005, 6(1):53-56.
- [11] Huang G, Zhu A, Siew C. Extreme learning machine theory and applications [J]. *Elsevier Neurocomputing*, 2006, 70(1-3):489-501.
- [12] 沈凌云, 郎百和, 朱明. 一种基于人工神经网络的人脸识别方法 [J]. 液晶与显示, 2011, 26(6):836-840.
- [13] 黎新伍. 基于 BP 神经网络的 CRT 显示器色彩管理新模型 [J]. 液晶与显示, 2008, 23(2):235-240.
- [14] 吉倩倩, 苏光大, 向守兵. 嵌入式邻域图像并行处理机的液晶显示系统设计 [J]. 液晶与显示, 2011, 26(6):768-773.
- [15] Mandal T, Wu Q M J, Yuan J. Curvelet based face recognition via dimension reduction [J]. *Elsevier Signal Processing*, 2009, 89 (3):2345-2353.