

## 基于基空间转换的姿态人耳识别

王 瑜, 穆志纯, 封 筠

(北京科技大学 信息工程学院, 北京 100083)

**摘 要:**人耳识别是近年来新兴起的一种生物特征识别技术,然而姿态问题一直是其难点问题之一,针对此问题提出了一种基于基空间转换的新方法.首先,利用主元分析和核主元分析方法得到姿态人耳图像和正侧面人耳图像的基空间,通过计算两种基空间之间的线性转换关系求出姿态转换矩阵,然后将待测的姿态人耳图像特征集利用基空间姿态转换矩阵转变成正侧面人耳图像特征集,最后用支持向量机进行分类识别.实验结果表明,该方法与没有经过姿态转换的方法相比,识别率显著提高.

**关键词:**基空间转换;人耳识别;主元分析;核主元分析;姿态转换矩阵;支持向量机

**中图分类号:** TP 391.41 **文献标识码:** A

### 0 引言

人耳自动识别是近年来新兴起的一种生物特征识别技术,虽然起步较晚,但由于不需要与被识别者直接发生身体上的接触而倍受青睐,目前用于人耳识别的方法很多,按照所提取的特征来划分可以归纳为两大类<sup>[1]</sup>:一类是基于几何特征的方法,这类方法通过寻找人脸或人耳轮廓和内部结构的关键点,构建几何特征;一类是基于代数特征的方法,例如主元分析法.虽然这些方法都取得了满意的效果,但都是基于无姿态变化或是小姿态变化下进行研究的,一旦在大角度姿态背景下便会表现得非常不稳定,有时甚至会完全失效,因此研究不同姿态下的人耳识别成为热点和难点.

针对人耳姿态问题,笔者在研究2D灰度图像的基础上,提出了一种基于姿态转换的新方法.首先利用PCA或KPCA方法得到姿态人耳图像和非姿态人耳图像的基空间,通过计算两种基空间之间的线性转换关系获得姿态转换矩阵,然后将待测的姿态人耳图像特征集利用姿态转换矩阵转换成正侧面人耳图像特征集,最后利用SVM分类器进行分类识别,实验结果表明了该方法处理姿态问题的有效性以及利用人耳图像进行身份识别的可行性.

### 1 姿态转换

姿态转换的思想依据主要有3点<sup>[2]</sup>:首先,人类视觉系统能够识别不同姿态的特征,并且当训练和测试的姿态相同时识别率会提高.其次,人类具有充分利用先验信息的特点,如果能够准确的知道人类获得先验知识的规则,就完全可以利用这些规则来进行不同姿态下的特征识别.第三,虽然人类还没有发现在姿态特征与非姿态特征之间究竟存在着什么样的关联性规则,以及人类自身是如何利用这些规则来鉴别姿态特征的,但是可以肯定的是,这种关联是确实存在的,姿态转换也是有意义的.

因为人耳图像是一种高维数据表达方式,需要将人耳图像由高维空间转换到低维空间,计算基空间并提取特征,进行姿态转换和分类识别.

#### 1.1 子空间表示

笔者为了更全面地说明姿态转换方法的有效性,利用主元分析(PCA)<sup>[3]</sup>和核主元分析(KPCA)两种子空间表示方法计算基空间并提取特征,PCA方法较为简单,这里不再介绍.

核主元分析方法(Kernel PCA)<sup>[4]</sup>的目的是寻找输入图像经映射后的高维空间 $W$ 的主元.

首先利用一个非线性映射 $\Phi$ 将训练样本图像 $x_i$ 映射到高维空间 $F$ 中, $i=1, \dots, N$ 为训练样

收稿日期:2007-10-20;修订日期:2007-12-10

基金项目:国家自然科学基金资助项目(60573058);北京市教委重点学科共建项目(XK100080537).

作者简介:王 瑜(1977-),女,吉林长春人,北京科技大学博士研究生,主要从事模式识别及图像处理方面的研究, E-mail: danahuiguniang@126.com.

本个数,且假设已被中心化,即满足  $\sum_{i=1}^N \Phi(x_i) = 0$ , 然后计算高维空间  $F$  中经过映射后样本图像的协方差矩阵  $C$ , 如下式所示:

$$C = \frac{1}{N} \sum_{i=1}^N \Phi(x_i) \Phi(x_i)^T \quad (1)$$

根据

$$\lambda v = Cv \quad (2)$$

求协方差矩阵  $C$  的本征值 ( $\lambda \geq 0$ ) 以及相对应的本征向量  $v_k (k = 1, \dots, m)$ .

由于所有本征向量  $v_k$  均可表示为空间  $F$  中样本  $\Phi(x_i)$  的线性张量

$$v_k = \sum_{i=1}^N w_k^i \Phi(x_i) \quad (3)$$

考虑等式

$$\lambda(\Phi(x_i)v) = \Phi(x_i)Cv \quad (4)$$

定义一个  $N \times N$  的有序核函数矩阵  $K_{ij} = K(x_i, x_j) = (\Phi(x_i), \Phi(x_j))$ , 并将式(1)、式(3)代入到式(4)得到

$$N\lambda w = Kw \quad (5)$$

最初的核变换不能保证  $\sum_{i=1}^N \Phi(x_i) = 0$ , 因此需要对核矩阵  $K$  进行归一化处理, 表达式为

$$\tilde{K}_{ij} = (K - A_c K - K A_c + A_c K A_c)_{ij} \quad (6)$$

式中:  $A_c$  是一个  $N \times N$  矩阵,  $(A_c)_{ij} = 1/N$ . 则特征值等式可以写为

$$N\lambda w = \tilde{K}w \quad (7)$$

根据式(7), 得到一组非零本征值 ( $\lambda \geq 0$ ) 以及相对应的满足归一化条件式(8)的本征向量  $w_k (k = 1, \dots, m)$ :

$$\lambda_k(w_k, w_k) = 1 \quad (8)$$

式中:  $w_k$  为第  $k$  大本征值所对应的本征列向量, 称为第  $k$  个基向量, 且这  $m$  个基向量是彼此正交的,  $W = (w_1, \dots, w_m)$ ,  $m \leq n$ ,  $W$  称为基空间.

因此对于任意一幅  $n$  维图像  $x_i, i = 1, \dots, N$  为训练样本个数,  $n$  为图像像素数, 为了提取其非线性主成分, 只需在高维空间  $F$  中, 计算该样本映射后  $\Phi(x_i)$  在本征向量  $v_k (k = 1, \dots, m)$  上的投影, 同时将式(5)代入, 可以得到

$$s_i^k = (v_k, \Phi(x_i)) = \sum_{j=1}^N w_k^j \tilde{K}(x_i, x_j) \quad (9)$$

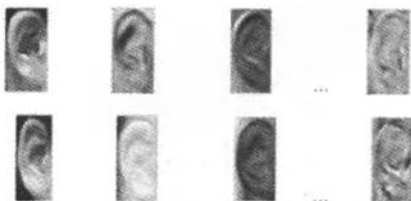
将所有  $\Phi(x_i)$  在  $v_k$  上的投影值形成一个列向量  $s_i = (s_i^1, \dots, s_i^m)^T$ , 作为任意样本图像  $x_i$  的特征列向量, 由这些特征列向量组成的集合构成

特征集, 即特征空间  $S = \{s_1, \dots, s_N\}$ .

## 1.2 姿态转换矩阵的获得

图1显示了正侧面人耳图像和带姿态人耳图像及其它们各自的基空间, 且它们的基空间分别用  $W^F$  和  $W^P$  表示. 为了获取它们之间存在着一定的关联性, 设一个  $m \times m$  维的方阵  $U$  满足:

$$W^F = W^P U \quad (10)$$



人耳图像 第一个基向量 第二个基向量 ... 第  $m$  个基向量

图1 正侧面人耳图像与带有姿态人耳图像的基空间表示

Fig.1 Basis space representing of the frontal and posed ear images

因为通常情况下, 基空间  $W$  的行数大于列数, 所以在  $(W^P)^T W^P$  为非奇异的情况下, 转换矩阵  $U$  利用下式来获得

$$U = \left( (W^P)^T (W^P) \right)^{-1} (W^P)^T (W^F) \quad (11)$$

又因为<sup>[2]</sup>

$$(S^F)^T = (S^P)^T U \quad (12)$$

式中:  $S^P$  为姿态人耳图像特征集,  $S^P = (s_1^P, \dots, s_N^P)$ ;  $N$  为样本图像个数;  $s_i^P$  为第  $i$  幅姿态人耳图像的特征列向量;  $S^F$  为利用姿态转换矩阵  $U$  新生成的正侧面人耳图像的特征集,  $S^F = (s_1^F, \dots, s_N^F)$ ;  $s_i^F$  为新生成的第  $i$  幅正侧面人耳图像的特征列向量, 且与  $S^P$  中的特征列向量  $s_i^P$  一一对应. 将式(11)代入到式(12), 即可求出  $S^F$ .

## 2 分类器设计

经过特征提取和姿态转换以后, 需要选择合适的分类器对人耳特征集进行分类识别. 由于SVM以结构风险最小化为准则, 在解决小样本、非线性及高维识别问题中表现出了许多特有的优势<sup>[3]</sup>, 因此采用支持向量机(SVM)做分类器.

对于给定的训练集  $(s_i, y_i) (i = 1, \dots, N)$ ,  $s_i \in R^n$  且为特征列向量,  $y \in \{+1, -1\}$  为类标记,  $N$  为样本个数. 则最优分类面函数设为

$$g(s_i) = \sum_{j=1}^N \alpha_j y_j K(s_i, s_j) + b^0 \quad (13)$$

式中:  $b^0$  为分类阈值;  $K(s_i, s_j)$  为核函数. 利用Lagrange优化方法可以将上述最优分类面问题转

化为在约束条件  $\sum_{i=1}^N y_i \alpha_i = 0$  和  $0 \leq \alpha_i \leq c$  ( $i = 1, \dots, N$ ;  $c$  为 Lagrange 乘子设定的上限) 下对  $\alpha_i$  求解下列函数的最大值问题<sup>[6]</sup>:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j K(s_i, s_j) \quad (14)$$

其中  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$ , 这个问题的最优化解中, 只有一部分 (通常是少部分)  $\alpha_i$  不为零, 其对应的样本就是支持向量。得到的最优分类函数是:

$$f(s) = \text{sgn} \left\{ \sum_{i=1}^N \alpha_i^* y_i K(s, s_i) + b_i^* \right\} \quad (15)$$

式中:  $\alpha_i^*$  为  $\alpha_i$  的最优解;  $b_i^*$  是分类阈值, 可以用任意一个支持向量满足公式  $y_i \left( \sum_{j=1}^N \alpha_j y_j K(s_i, s_j) + b^{op} \right) - 1 = 0$  求得。

SVM 基本理论是针对两分类问题, 在此基础上, 又发展了解决多分类问题的方法, 笔者选用“一对一”的方法以减少识别死区。

### 3 基于基空间转换的姿态人耳识别的实现

实验中所应用的人耳图像是自建的 USTB 彩色图像库三中的图像, 共 79 人, 其中选用了正侧面图像 ( $0^\circ$ ) 和向右旋转  $5^\circ$ 、 $20^\circ$ 、 $35^\circ$  和  $45^\circ$  4 种姿态图像, 每种情况 2 幅图像, 如图 2 所示。实验过程中统一将图像归一化为  $70 \times 138$  像素, 训练时使用 59 人的 5 种人耳图像 (共  $59 \times 5 \times 2 = 590$  幅) 求出 4 种姿态转换矩阵  $U$ , 测试时使用剩下的 20 人的 5 种图像 (共  $20 \times 5 \times 2$  幅) 来评估识别率。具体实验步骤如下 (以  $5^\circ$  为例进行说明)。

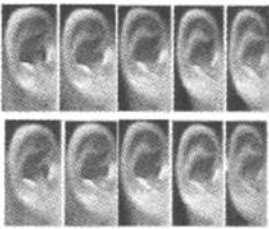


图 2 人耳图像实例 (从左至右:  $0^\circ$ ,  $5^\circ$ ,  $20^\circ$ ,  $35^\circ$ ,  $45^\circ$ )

Fig. 2 Example of ear images (From left to right:  $0^\circ$ ,  $5^\circ$ ,  $20^\circ$ ,  $35^\circ$ ,  $45^\circ$ )

第一步: 将 59 人的正侧面人耳图像组成数据集,  $A_{59}^F = (x_{11}, x_{12}, \dots, x_{a1}, x_{a2})$ ,  $x_{i1}, x_{i2}$  ( $i = 1, \dots, 59$ ) 分别为第  $i$  个人的两幅图像, 且均为列向量, 由每幅图像中第一列像素至最后一列像素首尾相接组成。同理, 这 59 人的  $5^\circ$  人耳图像可以组成数据集  $A_{59}^P$ 。

第二步: 将这两个数据集利用 PCA 和 KPCA 进行降维, 分别得到基空间  $W_{59}^F$  和  $W_{59}^P$ , 然后利用式 (13) 求出  $5^\circ$  的姿态转换矩阵  $U$ 。

第三步: 将剩下的 20 人的正侧面人耳图像和  $5^\circ$  人耳图像按照第一步介绍的方法组成数据集  $A_{20}^F$  和  $A_{20}^P$ , 然后利用 PCA 和 KPCA 方法进行特征提取, 分别得到特征集  $S_{20}^F$  和  $S_{20}^P$ , 再将  $S_{20}^P$  利用第二步求得的姿态转换矩阵  $U$  和式 (14) 生成正侧面人耳图像特征集  $(S_{20}^F)_{\text{new}}$ 。

第四步: 将第三步得到的特征集  $S_{20}^F$  进行训练得到 SVM 多类分类器, 然后对新生成的人耳图像特征集  $(S_{20}^F)_{\text{new}}$  进行分类识别。

### 4 实验结果及分析

笔者利用 PCA 和 KPCA 两种方法计算基空间和特征提取, PCA 方法选用了 40 个基向量, KPCA 选用了 118 个基向量, 且 KPCA 和 SVM 均选用了 Gaussian 核函数, 也就是径向基核函数, 即

$$K(x_i, x_j) = \exp \left( -\frac{\|x_i - x_j\|^2}{2\sigma^2} \right)$$

使用姿态转换矩阵前后的识别率如表 1 所示。

表 1 识别结果比较

Tab. 1 Recognition results				%
方法	$5^\circ$	$20^\circ$	$35^\circ$	$45^\circ$
NPT + PCA + SVM	97.5	35.0	17.5	10.0
NPT + KPCA + SVM	100.0	87.5	62.5	17.5
PT + PCA + SVM	97.5	82.5	77.5	55.0
PT + KPCA + SVM	100.0	87.5	80.0	60.0

注: NPT 为非转换; PT 为转换

(1) 不管是否使用姿态转换矩阵, KPCA 方法的识别率都要高于 PCA 方法, 这是因为 PCA 方法提取特征时只考虑了图像数据中的二阶统计信息, 而 KPCA 方法能够充分考虑输入数据的高阶非线性统计信息, 因此能够取得更好的识别效果。

(2) 由于人耳图像的信息随着旋转角度的增加而不断减少, 所以识别率逐渐降低。在小角度旋转的情况下, 姿态转换方法的优势并不明显, 但是当角度增大到一定程度时, 姿态转换的优势便突显出来, 尤其在  $45^\circ$  的情况下, 没有经过姿态转换的识别率只有 10%, 如此低的识别率已经表明, 在这种旋转角度下, 几乎不能使用人耳特征作为有效的生物特征识别依据, 但经过姿态转换以后, 可以将识别率明显提高, 达到 60%。

另外, 图 3 将本文方法与 Hyung - Soo Lee 等提出的特征空间转换方法进行了比较, 这里特征

提取仅使用了 PCA 方法。

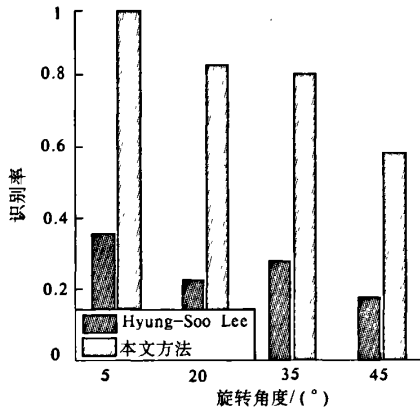


图3 不同姿态转换方法识别率比较

Fig.3 Recognition results of different pose transformation methods

之所以会出现这样的结果,是因为 Hyung - Soo Lee 等提出的方法需要经过图像重构的过程,难免会出现重构误差。另外,人脸具有明显的特征信息,如眼睛、鼻子、嘴等,而人耳的沟壑、脊等特征信息远没有人脸丰富,因此对于带有姿态的人耳图像和正侧面人耳图像来说,基空间之间的相关性比特征空间之间的相关性更大,故而本文提出的方法识别率表现更好。

## 5 结论

笔者针对人耳识别中的姿态问题,提出了一

种基于基空间姿态转换矩阵的新方法,将带有姿态的人耳图像特征集转换成正侧面人耳图像特征集,然后利用 SVM 进行分类识别。实验结果表明,当对带有姿态的人耳图像进行匹配识别时,尤其是在大角度姿态变化下,该方法表现出了显著的优势。

## 参考文献:

- [1] 张海军, 穆志纯. 基于复合结构分类器的人耳识别[J]. 北京科技大学学报, 2006, 28(12): 1186 - 1190.
- [2] LEE H S, KIM D J. Generating frontal view face image for pose invariant face recognition [J]. Pattern Recognition Letters, 2006, 27: 747 - 754.
- [3] MENG J C, Zhang W B. Volume measure in 2DPCA - based face recognition [J]. Pattern Recognition Letters, 2007, 28(10): 1203 - 1208.
- [4] LIU C J. Gabor - based kernel PCA with fractional power polynomial models for face recognition [J]. IEEE trans pattern Anal Mach Intell, 2004, 26(5): 572 - 580.
- [5] HEISELE B, HO P, WU J, et al. Face recognition: component - based versus global approaches [J]. Computer Vision and Image Understanding, 2003, 91: 6 - 21.
- [6] 齐保林, 李凌均, 李志农. 基于支持向量机的故障模式识别研究[J]. 郑州大学学报:工学版, 2007, 28(1): 9 - 11.

## Posed Ear Recognition Based on Basis Space Transformation

WANG Yu, MU Zhi - chun, FENG Jun

(School of Information Engineering, University of Science and Technology of Beijing, Beijing 100083, China)

**Abstract:** Ear recognition is a kind of new biometrics in recent years. Pose change, however, is one of the difficult problems to ear recognition. Therefore' this paper proposes a new method based on basis space transformation to solve ear pose problem. Firstly, basis spaces including the posed ear images and the frontal ear images are obtained using PCA or KPCA. Secondly, the pose transformation matrix is gained in the light of linear transformation relation between two kinds of basis spaces. Thirdly, features of the posed ear images to be test are transformed into those of corresponding frontal ear images on the basis of the pose transformation matrix. Finally, the generated frontal ear features are identified by Support Vector Machine method. Experimental results show that the recognition rate with the pose transformation outperforms that without pose transformation remarkably.

**Key words:** basis space transformation; ear recognition; PCA; KPCA; pose transformation matrix; SVM