

# 一类多频带主分量分析方法

郭志波<sup>1</sup> 严云洋<sup>2</sup> 庞成<sup>1</sup>

(1. 扬州大学信息工程学院, 扬州, 225127; 2. 淮阴工学院计算机工程学院, 淮安, 223003)

**摘要:** 主分量分析是模式识别领域使用较广的一种特征抽取方法, 但是由于经典的主分量分析在处理图像矩阵时需要将图像展开成向量形式, 使得计算量很大。本文提出了一种多频带主分量分析方法, 该方法不仅减少了运算过程中的计算量, 而且在一定程度上提高了整体性能。首先通过二维离散余弦变换将图像转变成频率数据, 再按照频率变化将数据分成多个频带, 然后在此基础上设计了针对多个频带数据的主分量分析方法。通过对 ORL 和 NUST603 图像库进行实验证明, 本文方法不仅具有快速提取图像特征的能力, 而且综合性能优于相应的主分量分析。

**关键词:** 主分量分析; 多频带主分量分析; 特征抽取

**中图分类号:** TP391.4      **文献标志码:** A

## Multi-Band Principal Component Analysis Method

Guo Zhibo<sup>1</sup>, Yan Yunyang<sup>2</sup>, Pang Chen<sup>1</sup>

(1. The College of Information Engineering, Yangzhou University, Yangzhou, 225127, China; 2. Faculty of Computer Engineering, Huaiyin Institute of Technology, Huai'an, 223003, China)

**Abstract:** Principal component analysis (PCA) is the well-known method in pattern recognition. However, expanding original image matrices into the same dimensional vectors in classical PCA increase the computational complexity. Here one presents a kind of multi-band principle component analysis (MBPCA). The process can reduce the computational complexity thus improving the overall performance. Firstly, the image is transformed into frequency data by the two-dimensional discrete cosine transform. Secondly, frequency data is divided into a plurality of frequency bands according to its frequency range. Finally, a principal component analysis method using a plurality of frequency bands is designed. The experiments on ORL and NUST603 face database show that the proposed method has the ability to quickly extract image features and performs better than the corresponding principal component analysis.

**Key words:** principal component analysis (PCA); multi-band PCA (MBPCA); feature extraction

## 引 言

主分量分析(Principal component analysis, PCA)是模式识别中使用较为广泛的一种降维分析方法,

通过该方法可以抽取出少数的鉴别特征来代替整个样本,大大降低了样本的维数,为后面更加快速且有效的分类工作打下基础。当前,人脸识别<sup>[1]</sup>是模式识别和机器视觉领域的一个研究热点。基于经典主分量分析(Classical principal component analysis, CPCA)的 Eigenface 识别方法<sup>[2-3]</sup>是其中使用较广的分析方法,许多在人脸识别方面的 PCA 方法<sup>[4-13]</sup>都是基于 CPCA 发展起来。当在人脸图像上使用 CPCA 时,需要先将图像矩阵展开成向量形式,再用得到的协方差矩阵进行降维分析。由于图像矩阵转化为向量后,造成向量的维数变得非常大,如需要处理分辨率为  $N \times N$  图像,得到协方差矩阵的维数就是  $N^2 \times N^2$ ,所以对于高维的图像完成 CPCA 就比较费时间。本文提出了一种多频带主分量分析方法(Multi-band principle component analysis, MBPCA),不仅减少了过程中的计算量,而且提高了整体性能。首先通过 FFT 正交变换方法将图像转变成频率数据,再按照频率变化将数据分成多个频带,然后设计了一种针对多个频带数据的 PCA 方法。如果频带数目是 16,则通过 MBPCA 方法需要的计算量大约是 CPCA 方法的 1/16。为了检验算法的有效性,使用 ORL, Yale 人脸库进行了实验,实验结果证明本文算法综合性能优于相应的主分量分析。

## 1 相关知识

### 1.1 基本定义

设  $\omega_1, \omega_2, \dots, \omega_c$  为  $c$  个模式类,样本  $\mathbf{X} = \{\mathbf{x}_i\}, i=1, 2, \dots, n, \mathbf{x}_i$  为  $m$  维实向量,类间散布矩阵  $\mathbf{S}_b$ ,类内散布矩阵  $\mathbf{S}_w$  和总体散布矩阵  $\mathbf{S}_t$  分别定义为

$$\mathbf{S}_b = \sum_{i=1}^c P(\omega_i) (\mathbf{m}_i - \mathbf{m}_0) (\mathbf{m}_i - \mathbf{m}_0)^T \quad (1)$$

$$\mathbf{S}_w = \sum_{i=1}^c P(\omega_i) E\{(\mathbf{X} - \mathbf{m}_i) (\mathbf{X} - \mathbf{m}_i)^T / \omega_i\} \quad (2)$$

$$\mathbf{S}_t = \mathbf{S}_b + \mathbf{S}_w = E\{(\mathbf{X} - \mathbf{m}_0) (\mathbf{X} - \mathbf{m}_0)^T\} \quad (3)$$

式中:  $P(\omega_i)$  为第  $i$  类训练样本的先验概率(一般在人脸问题上取等概率);  $\mathbf{m}_i = E\{X/\omega_i\}$  为第  $i$  类训练样本的均值;  $\mathbf{m}_0 = E\{\mathbf{X}\} = \sum_{i=1}^m P(\omega_i) \mathbf{m}_i$  为全体训练样本的均值。

由式(1~3)的定义知,  $\mathbf{S}_w, \mathbf{S}_b, \mathbf{S}_t$  均为非负定矩阵。而且当  $\mathbf{S}_w$  可逆时,  $\mathbf{S}_w$  与  $\mathbf{S}_t$  均为正定矩阵。

### 1.2 经典主分量分析

实际上, CPCA 旨在通过准则函数式(4)找到一个最优的投影矩阵  $\mathbf{A}_{\text{opt}} = \arg \max_{\mathbf{A}} (J_{\rho}(\mathbf{A}))$ 。

$$J_{\rho}(\mathbf{A}) = \text{tr}[(\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{S}_t \mathbf{A})] \quad (4)$$

通过  $\partial J_{\rho}(\mathbf{A}) / \partial \mathbf{A} = 0$ , 可以得到 PCA 的最优投影轴, 即  $\mathbf{A}_{\text{opt}}$  的列向量为特征方程  $\mathbf{S}_t \mathbf{A} = \lambda \mathbf{A}$  的  $d$  个最大的特征值所对应的标准正交特征向量  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d$ , 其满足以下条件

$$\mathbf{S}_t \mathbf{u}_j = \lambda_j \mathbf{u}_j \quad j = 1, \dots, d \quad (5)$$

其中  $\lambda_1 \geq \dots \geq \lambda_d$ 。

## 2 多频带主分量分析

### 2.1 多频带分解

首先需要将图像数据转变成频率数据,可以通过多种正交变换的方法,比如说离散傅里叶变换(Discrete Fourier transform, DFT)、离散余弦变换(Discrete cosine transform, DCT)、离散正弦变换(Discrete sine transform, DST)、离散沃尔什变换(Discrete walsh transform, DWT)以及离散小波变换(也称 Discrete wavelet transform, DWT)等。一般对于图像数据,采用离散余弦变换、离散小波变换及离散沃

尔什变换性能较为突出,在这些变化中离散沃尔什变换要求数据长度必须是2的整数次方,而离散小波变换尤其是Gabor小波,虽然在图像局部特征提取方面应用较广,但是本文由于在识别时只是使用最低频分量,而离散余弦变换得到的最低频子图保存的能量最大,其形成的空间关系更接近原始图像空间,并且离散余弦变换计算速度明显快于离散小波变换,所以综合考虑,采用了二维离散余弦变换(Two-dimensional discrete cosine transform, 2DDCT)方法来获取频率数据。如果图像 $\mathbf{X}$ 是 $m \times n$ 维矩阵,则 $\mathbf{X}$ 的2DDCT变换为 $\mathbf{F} = \mathbf{C}_m \mathbf{X} \mathbf{C}_n^T$ ,其中 $\mathbf{F}$ 为 $m \times n$ 维频率数据, $\mathbf{C}_k = (c_{ij}) (k = m, n)$ 为DCT矩阵。DCT矩阵的构造方法如下

$$\begin{cases} c_{ij} = \frac{1}{\sqrt{k}} \\ c_{ij} = \sqrt{\frac{2}{k}} \cos \frac{(2j-1)(l-1)\pi}{2k} \end{cases} \quad i, j = 1, \dots, k \quad (6)$$

如图1所示,图1(a)是原图像,图1(b)是通过得到的频率数据,按照频率变化,将频率数据划分成 $d$ 个频带,如图1(c)所示,然后将这些频带数据展开成向量形式 $\mathbf{f}^{(d)}$ 。由于2DDCT可以通过二维快速傅里叶变换(2DFFT)进行,这样可以大大加快训练样本的多频带分解速度。

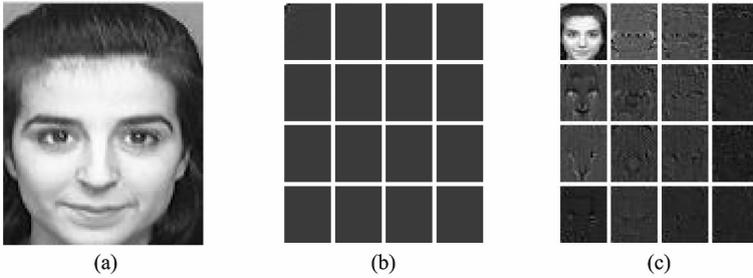


图1 一幅图像的多频带分解

Fig. 1 Multi-band decomposition of an image

## 2.2 MBPCA 算法

如果有 $l$ 个图像样本 $\mathbf{X}_i, i = 1, 2, \dots, n, \mathbf{X}_i$ 是 $m \times n$ 维实矩阵,通过多频带分解可以得到 $l \times d$ 个矢量 $\mathbf{f}_i^{(j)} (i = 1, 2, \dots, l; j = 1, 2, \dots, d)$ 。此时总体散布矩阵 $\hat{\mathbf{S}}_l$ 被定义为

$$\hat{\mathbf{S}}_l = \frac{1}{l \times d} \sum_{j=1}^d \sum_{i=1}^l (\mathbf{f}_i^{(j)} - \mathbf{m}_0^{(j)}) (\mathbf{f}_i^{(j)} - \mathbf{m}_0^{(j)})^T$$

式中: $\mathbf{m}_i^{(j)}$ 为第 $i$ 类样本 $j$ 频带的中心, $\mathbf{m}_0^{(j)}$ 为所有样本 $j$ 频带的中心。

MBPCA 准则为

$$\begin{cases} J_m(\mathbf{A}) = \text{tr}(\mathbf{A}^T \hat{\mathbf{S}}_l \mathbf{A}) \\ \mathbf{A}_{\text{opt}} = \arg \max_{\mathbf{A}} (J_m(\mathbf{A})) \end{cases} \quad (7)$$

由于通过多频带分解,在一定程度上提高了离散度矩阵的秩,只要频带数大于一定数目就可以使得 $\hat{\mathbf{S}}_l$ 满秩。另外,可以证明,得到的多频带数据的总体离散度和直接通过原始数据获得的总体离散度相同,也就是说MBPCA所在的数据空间并没有损失信息。

证明:多频带数据的总体离散度为

$$\hat{\mathbf{S}} = \frac{1}{l} \sum_{j=1}^d \sum_{i=1}^l (\mathbf{f}_i^{(j)} - \mathbf{m}_0^{(j)})^T (\mathbf{f}_i^{(j)} - \mathbf{m}_0^{(j)}) \quad (8)$$

原始数据的总体离散度为

$$S = \frac{1}{l} \sum_{i=1}^l (\hat{\mathbf{X}}_i - \mathbf{m}_0)^T (\hat{\mathbf{X}}_i - \mathbf{m}_0) \quad (9)$$

式中:  $\mathbf{m}_0$  为所有样本  $\mathbf{X}_i$  的向量形式  $\hat{\mathbf{X}}_i$  的中心,  $i=1, 2, \dots, l$ 。

$f_i^{(j)}$  可以表示为  $\mathbf{U}_j^T \hat{\mathbf{X}}_i$ , 其中  $\mathbf{U}_j$  为第  $j$  个频带的投影矩阵, 所以多频带数据的总体离散度可以有下面形式

$$\begin{aligned} \hat{S} &= \frac{1}{l} \sum_{j=1}^d \sum_{i=1}^l (\hat{\mathbf{X}}_i - \mathbf{m}_0)^T \mathbf{U}_j \mathbf{U}_j^T (\hat{\mathbf{X}}_i - \mathbf{m}_0) \\ &= \frac{1}{l} \sum_{i=1}^l [(\hat{\mathbf{X}}_i - \mathbf{m}_0)^T (\sum_{j=1}^d \mathbf{U}_j \mathbf{U}_j^T) (\hat{\mathbf{X}}_i - \mathbf{m}_0)] \end{aligned} \quad (10)$$

而  $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_j]$  是  $(m \times n) \times (m \times n)$  维矩阵, 其中每个列向量都来自于离散余弦正交矩阵  $(\mathbf{C}_m)^T$  中任意列向量和  $\mathbf{C}_n$  中任意行向量相乘之后得到矩阵的展开向量, 所以  $\mathbf{U}$  中每个列向量都是相互正交的, 则有

$$\sum_{j=1}^d \mathbf{U}_j \mathbf{U}_j^T = \mathbf{U} \mathbf{U}^T = \mathbf{I} \quad (11)$$

式中:  $\mathbf{I}$  为单位矩阵, 根据式(9, 10)得到

$$\hat{S} = \frac{1}{l} \sum_{i=1}^l [(\hat{\mathbf{X}}_i - \mathbf{m}_0)^T (\hat{\mathbf{X}}_i - \mathbf{m}_0)] = S \quad (12)$$

综上所述, 得到的多频带数据的总体离散度和直接通过原始数据获得的总体离散度相同。为了降低投影特征的维数, 实验中发现最低频频带的数据在投影空间上进行投影所得到的特征作为鉴别特征就可以得到较好的实验结果, 通过该投影方法在保持较好的识别精度基础上, 使得投影特征维数也大大降低。

在抽取样本  $\mathbf{X}_i$  的特征时, 只需根据下式可以直接抽取  $\mathbf{X}_i$  的鉴别特征

$$y_i = \mathbf{A}_{\text{opt}}^T f_i^{(1)} \quad (13)$$

式中:  $f_i^{(1)}$  为样本  $\mathbf{X}_i$  展开向量的最低频数据。

### 3 实验结果与分析

实验中使用了 ORL 人脸库中所有样本图像, 含有 40 人, 每人 10 幅, 共 400 幅图像。实验中选用 4 次随机样本组合(对每人而言): 随机选择 5 幅用于训练, 剩余 5 幅用于测试; 然后将 4 次样本选择模式下测得的结果等概率相加得到最终的实验结果。实验中使用图像样本的分辨率为  $92 \times 112$ 。MBPCA 采用  $8 \times 8$  个频带。



图 2 ORL 人脸库中某人的 10 幅图像

Fig. 2 Ten images of someone in ORL human face database

实验结果如表 1 所示,表中粗体标出的是每种方法的测试最大精度。从表 1 中可以看出,无论使用最小距离还是最近邻测量方法,MBPCA 的识别精度都比 CPCA 有显著的提高。

表 1 在 ORL 人脸库的实验结果 %

Tab. 1 Comparison of experiment results on ORL face database

特征维数	最小距离		最近邻	
	MBPCA	CPCA	MBPCA	CPCA
19	88.2	83.0	91.1	88.9
29	89.6	84.6	92.3	91.1
39	<b>91.3</b>	86.8	<b>93.4</b>	91.5
59	91.2	87.7	93.2	91.7
79	91.3	<b>88.3</b>	93.3	<b>92.2</b>

MBPCA 训练时间为 8.3 s,而 CPCA 为 475 s,可以知道 MBPCA 计算最优投影矢量集的速度远快于 CPCA。另外,对 Yang 提出的 2DPCA 在本文数据库上进行了实验,最好结果为最近邻 93.1%,最小距离 90.5%,虽然精度和本文方法相近,但其抽取特征的维数高达  $112 \times 10 = 1120$  维,而本文方法中最优结果的特征维数只有 39 维,在后期进行人脸识别时要远快于 2DPCA。为了检验 MBLDA 对较大人脸库的识别能力,最后采用 NUST603 人脸库来进行实验,NUST603 人脸库是一个规模较大的人脸数据库,其中包含 96 人,每人 10 幅图像,每幅图像为  $32 \times 32$  的标准图像,MBPCA 采用  $2 \times 2$  个频带并采用同 ORL 人脸库上进行实验时相同的样本选择模式进行实验。

表 2 中的人脸识别率为最终的平均实验结果,其中粗体标出的是每种方法的最大精度,在测试结果中,可以看出在 NUST603 人脸库上,无论使用最小距离还是最近邻测量方法,MBPCA 的识别率都优于 PCA 方法。

表 2 在 NUST603 人脸库上的实验结果 %

Tab. 2 Comparison of experiment results on NUST603 face database

特征维数	最小距离		最近邻	
	MBPCA	PCA	MBPCA	PCA
15	91.5	81.5	95.1	85.6
35	<b>92.5</b>	88.3	<b>95.8</b>	92.0
55	92.3	89.5	95.3	93.6
75	92.3	89.9	95.5	<b>94.2</b>
85	92.2	<b>90.1</b>	95.5	94.1
95	92.3	90.1	95.4	94.2

## 4 结束语

本文提出一种多频带主分量分析,不仅降低了训练时间,而且提高了识别性能。MBPCA 提取的鉴别特征维数小,同时在识别精度上明显优于 PCA,通过对 ORL,NUST603 人脸库实验证明了该算法的有效性。

## 参考文献:

- [1] Zhao W, Chellappa R, Rosenfeld A, et al. Face recognition: A literature survey[R]. Technical Report CAR-TR-948, UMD CS-TR-4167R, 2002.
- [2] Kirby M, Sirovich L. Application of the KL procedure for the characterization of human faces[J]. IEEE Trans Pattern Anal-

ysis and Machine Intelligence, 1990, 12(1): 103-108.

- [3] Turk M, Pentland A. Eigenfaces for recognition[J]. *Journal of Cognitive Neuroscience*, 1991, 3(1): 71-86.
- [4] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces: Recognition using class special linear projection [J]. *IEEE Trans Pattern Analysis and Machine Intelligence*, 1997, 19 (7): 711-720.
- [5] Swets D L, Weng J. Using discriminant eigenfeatures for image retrieval[J]. *IEEE Trans Pattern Analysis and Machine Intelligence*, 1996, 18(8): 831-836.
- [6] Yang J, David Z, Yang J Y. 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.
- [7] Guo Z B, Yan Y Y. A retrieve space principal component analysis based on the image retrieve[C]//*Proceedings of the 2009 Chinese Conference on Pattern Recognition and the 1st CJK Joint Workshop on Pattern Recognition, CJKPR, Piscataway, USA; IEEE, 2009:63-67.*
- [8] 雷刚, 吴朋. 基于虚拟信息的单样本分块人脸识别方法[J]. *数据采集与处理*, 2009, 24(4):443-448.  
Lei Gang, Wu Peng. Sub-block face recognition method based on virtual information with one training image per person[J]. *Journal of Data Acquisition and Processing*, 2009, 24(4):443-448.
- [9] 郑豪. 基于 L1 范式的分块 2DPCA 人脸识别方法[J]. *数据采集与处理*, 2011, 26(6):666-670.  
Zheng Hao. Method of face recognition based on L1 norm and block two-dimensional principal component analysis[J]. *Journal of Data Acquisition and Processing*, 2011, 26(6):666-670.
- [10] 张国云, 章兢. 基于 Gabor 小波的多尺度 PCA 支持向量机人脸识别方法[J]. *数据采集与处理*, 2009, 24(4):438-442.  
Zhang Guoyun, Zhang Jing. Multi-scale PCA support vector machine algorithm based on Gabor wavelet for face recognition [J]. *Journal of Data Acquisition and Processing*, 2009, 24(4):438-442.
- [11] Yu Z Y, Gao S B. Fuzzy Two-dimensional principal component analysis and its application to face recognition [J]. *AISS: Advances in Information Sciences and Service Sciences*, 2011, 3(11): 335-341.
- [12] Zeng Y, Feng D Z. The face recognition method of the two-direction variation of 2DPCA[J]. *International Journal of Digital Content Technology and Its Applications*, 2011, 5(2): 216-223.

#### 作者简介:



郭志波 (1975-), 男, 博士, 副教授, 研究方向: 模式识别、图像处理, E-mail: zhibo\_guo@163.com。



严云洋 (1968-), 男, 博士, 教授, 研究方向: 模式识别、机器视觉, E-mail: zhibo\_guo@163.com。



庞成 (1989-), 男, 硕士研究生, 研究方向: 模式识别、图像处理, E-mail: pangcheng4432@126.com。

