一种新的基于DCT变换的线性判别分析

黄国宏[1],刘 刚[2]

(1.广东工业大学 信息工程学院,广东 广州 510006;
2.上海电力学院 电力与自动化工程学院,上海 200090)

摘 要:对于高维复杂模式识别问题,传统的线性判别分析通常首先采用PCA变换来降低模式的维数,然后再求取最优判别矢量集。然而PCA变换是以判别信息的损失为代价的,故无法保证所提取的特征是最优的。DCT变换具有“能量聚集特性”和变换的保距特性,文中正是基于此特性,提出一种新的基于DCT变换的线性判别分析方法,同时,也给出了一种在该模型下的最优判别矢量集的直接求解方法。实验表明,文中算法具有计算速度快、识别率高的优点。

关键词:DCT变换;主元分析;特征提取;线性判别分析

中图分类号:TP391 **文献标识码**:A **文章编号**:1673-629X(2008)06-0097-04

A Novel Linear Discriminant Analysis Based on DCT

HUANG Guo-hong[1], LIU Gang[2]

(1.Sch. of Information, Guangdong University of Technology, Guangzhou 510006, China;
2.Sch. of Electric Power and Automation Eng., Shanghai University of Electric Power, Shanghai 200090, China)

Abstract: For the issue of high-dimensional complex pattern recognition, classical linear discriminant analysis usually use PCA transformation to reduce the dimensionality of the patterns, and then solve the optimal discriminant vector set. However, the process of reducing the dimensionality is at a cost of losing the discriminant information. DCT transformation has the properties of “energy compaction” and distance preserving. Based on the properties, takes the DCT as the preprocessing method of data dimensionality reduction. At the same time, a direct algorithm to solve the optimal discriminant vector set is also proposed. The experimental result shows that the present method is superior to the existing methods in terms of correct classification rate.

Key words: DCT; PCA; feature extraction; linear discriminant analysis

0 引 言

线性判别分析(LDA),也称为Fisher判别分析,是一种经典的特征提取算法,其基本思想就是在Fisher准则函数取得最大值的条件下,寻找一组变换使得变换后样本的类内离散度最小且类间离散度达到最大,亦即可分性最好[1,2]。目前,基于线性判别分析LDA的特征提取方法已被广泛地应用于人脸识别[3~5]及多媒体信息检索[6,7]等领域。

众所周知,对于高维小样本问题(如人脸识别),样本的类内离散度矩阵 S_w 通常是奇异的,这就导致无法直接通过使得Fisher准则函数最大化来求解最优判别矢量集。针对这种情况,Swets等最先提出结合主元分析(PCA)的线性判别方法[6],通过PCA变换来降低原始图像的维数,消除类内散度矩阵 S_w 的奇异性,然后再利用线性判别分析求取最优判别矢量集。在此基础上,Belhumeur等提出了著名的Fisherfaces方法[3],Yang等提出了最优组合判别矢量集方法[7]等,并且都取得了不错的效果。

尽管这种降维方法可以消除类内离散度矩阵的奇异性,但都是以判别信息的损失为代价的,从而无法保证所提取的特征是最优的。根据主元分析理论,如果选择所有非零特征值对应的特征向量作为主元,就不会损失样本的任何信息。然而,Liu等通过实验证明这种选取方法反而会导致识别性能的下降[8]。因此,如何对高维模式进行有效的降维仍是一个值得研究的问题。

DCT(Discrete Cosine Transform)变换具有“能量聚集特性”和变换的保距特性[9],被广泛地应用于信号处理和数据压缩[10,11]。它的本质就是在一定损失的情况下,采用较少数目的点数来表示原点数。也就是说一个高维数据经过DCT变换后可以用较少数目

收稿日期:2007-09-19

作者简介:黄国宏(1975-)男,博士,讲师,研究方向为图像处理与分析、机器视觉、模式识别等;刘 刚,博士,讲师,研究方向为图像处理与分析、机器视觉、模式识别等。

的维数来表示，且数据的本质特征没有发生变化。

文中正是基于这种特性，尝试利用 DCT 变换来代替主元分析，提出了一种新的基于 DCT 变换的线性判别分析方法，并取得了不错的效果。该方法通过 DCT 变换来提取人脸的整体信息，然后再利用线性判别分析求解最优判别矢量集。由于 DCT 有成熟的快速算法，有效地解决了计算量大的问题，同时，也给出了一种在该模型下的最优判别矢量集的直接求解方法，实验表明，文中算法与传统线性判别分析相比，具有计算速度快、识别率高的优点。

1 基于 DCT 变换的线性判别分析

1.1 人脸图像的二维 DCT 变换

灰度图像可以看作二维矩阵，一幅 $N\times N$ 的图像矩阵 $f(x,y)$ 的 DCT 变换定义为：

$$F(u,v)=\frac{2c(u)c(v)}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (1)$$

$u=0,1,\cdots,N-1;v=0,1,\cdots,N-1$。其中：

$$c(k)=\begin{cases}\frac{1}{\sqrt{2}}, k=0\\ 1,\text{其它}\end{cases}$$

由 DCT 变换的定义可知，图像 DCT 变换的系数数目与图像中的像素数目相等。因此，如何选择变换系数、选择哪些系数来表述人脸图像成为问题的关键。庆幸的是，图像的 DCT 变换具有能量集中的特点，即变换后原图像的绝大部分信息(能量)都集中到低频部分。如图 1 所示。

图 1 人脸图像(左)及 DCT 变换系数幅值图(右)

根据 DCT 变换所具有的能量集中的特点，选用部分低频系数作为图像的特征，即从变换系数矩阵的左上角开始，从左到右、从上到下提取 DCT 系数。由于过多的系数对应于图像的高频成分，它们更易受到噪声的影响，从而影响后续的识别；此外，选用过多的 DCT 系数也将会增加后续的计算量。

1.2 基于 DCT 变换的线性判别分析

假定 $\omega_1,\omega_2,\cdots,\omega_L$ 为 L 个已知模式类，X 为一经过 DCT 变换处理后的 n 维训练样本。设 $\boldsymbol{m}_0$ 为全体训练样本的均值向量，$\boldsymbol{m}_i,\boldsymbol{S}_i,\boldsymbol{P}_i(i=1,2,\cdots,L)$ 分别为 ω_i 类的均值向量、协方差矩阵和先验概率，则类间散度矩阵 $\boldsymbol{S}_b$，类内散度矩阵 $\boldsymbol{S}_w$ 和总体散度矩阵 $\boldsymbol{S}_t$ 分别为：

$$\boldsymbol{S}_b=\sum_{i=1}^{L}\boldsymbol{P}_i(\boldsymbol{m}_i-\boldsymbol{m}_0)(\boldsymbol{m}_i-\boldsymbol{m}_0)^{\mathrm{T}} \quad (2)$$

$$\boldsymbol{S}_w=\sum_{i=1}^{L}\boldsymbol{P}_iE[(X-\boldsymbol{m}_i)(X-\boldsymbol{m}_i)^{\mathrm{T}}|\omega_i]=\sum_{i=1}^{L}\boldsymbol{P}_i\boldsymbol{S}_i \quad (3)$$

$$\boldsymbol{S}_t=E[(X-\boldsymbol{m}_0)(X-\boldsymbol{m}_0)^{\mathrm{T}}]=\boldsymbol{S}_b+\boldsymbol{S}_w \quad (4)$$

Fisher 准则函数可定义为：

$$J(\boldsymbol{\varphi})=\frac{\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{\varphi}}{\boldsymbol{\varphi}_{S_w}^{\mathrm{T}}\boldsymbol{\varphi}} \quad (5)$$

推广的 Fisher 准则函数定义如下：

$$J(\boldsymbol{\varphi})=\frac{\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{\varphi}}{\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{S}_t\boldsymbol{\varphi}} \quad (6)$$

其中，$\boldsymbol{\varphi}$ 为任一使 $J(\boldsymbol{\varphi})$ 达到最大值的单位向量，即模式样本在 $\boldsymbol{\varphi}$ 方向上的投影具有最大的类间散度和最小的类内散度，亦即可分性最好。文中采用推广的 Fisher 准则函数式(6)。

受文献[7]的启发，从最优化的角度，利用 Fisher 准则即为广义 Rayleigh 商，且广义 Rayleigh 商所对应的广义特征方程存在共轭正交的特征向量这一结论[12]，给出了具有统计不相关的最优判别矢量集[13]的直接求解方法。

定义 1 设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶实对称矩阵，且 $\boldsymbol{B}$ 正定，$x\in R^n$，称

$$R(x)=\frac{x^{\mathrm{T}}\boldsymbol{A}x}{x^{\mathrm{T}}\boldsymbol{B}x},x\neq 0 \quad (7)$$

为矩阵 $\boldsymbol{A}$ 相对于矩阵 $\boldsymbol{B}$ 的广义 Rayleigh 商。

由定义 1 可知，Fisher 准则函数式(6)即为类间散度矩阵 $\boldsymbol{S}_b$ 相对于总体散度矩阵 $\boldsymbol{S}_t$ 的广义 Rayleigh 商。

定理 1 广义特征方程 $\boldsymbol{A}x=\lambda\boldsymbol{B}x$ 存在 n 个特征向量 $x_1,x_2,\cdots,x_n$ 满足如下条件：

$$x_i^{\mathrm{T}}\boldsymbol{B}x_j=\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases} \quad i,j=1,\cdots,n \quad (8)$$

证明：由于 $\boldsymbol{B}$ 是正定对称矩阵，必存在正交阵 $\boldsymbol{U}=(u_1,u_2,\cdots,u_n)$，使得

$$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{U}=\boldsymbol{\Lambda}=\mathrm{diag}(a_1,\cdots,a_n) \quad (9)$$

其中，$u_1,u_2,\cdots,u_n$ 为 $\boldsymbol{B}$ 的标准正交特征向量，$a_1,\cdots,a_n$ 为所对应的特征值，且满足 $a_i>0,i=1,\cdots,n$。

令 $\boldsymbol{W}=\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}$，则 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{W}=\boldsymbol{I}$，有 $\boldsymbol{B}=(\boldsymbol{W}^{-1})^{\mathrm{T}}\boldsymbol{W}^{-1}$，代入特征方程 $\boldsymbol{A}x=\lambda\boldsymbol{B}x$，得

$$\boldsymbol{A}x=\lambda(\boldsymbol{W}^{-1})^{\mathrm{T}}\boldsymbol{W}^{-1}x \quad (10)$$

令 $y=\boldsymbol{W}^{-1}x$，则有 $x=\boldsymbol{W}y$，代入上式，得

$$\boldsymbol{W}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{W}y=\lambda y \quad (11)$$

令 $\boldsymbol{S}=\boldsymbol{W}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{W}$，得 $\boldsymbol{S}y=\lambda y \quad (12)$

于是,广义特征值问题 $Ax=\lambda Bx$ 等价地转化为对称矩阵 S 的普通特征值问题式(12)。显然 S 也是实对称矩阵,所以它的特征值 $\lambda_1,\cdots,\lambda_n$ 均为实数,且存在着完备的标准正交特征向量系 $y_1,\cdots,y_n$,即有

$$y_i^{\mathrm{T}}y_j=\begin{cases}0 & i\neq j\\ 1 & i=j\end{cases} \tag{13}$$

把 $y_i=W^{-1}x_i, i=1,\cdots,n$,代入式(13),得

$$y_i^{\mathrm{T}}y_j=x_i^{\mathrm{T}}(W^{-1})^{\mathrm{T}}W^{-1}x_j=x_i^{\mathrm{T}}Bx_j \tag{14}$$

综合式(13)与式(14),有

$$x_i^{\mathrm{T}}Bx_j=\begin{cases}0 & i\neq j\\ 1 & i=j\end{cases} \tag{15}$$

得证。

定理2 具有统计不相关性(即关于 S_t 共轭正交)的最优判别矢量集 $\Psi=(x_1,x_2,\cdots,x_d),(d\leqslant n)$,可构成如下线性变换:

$$Z=\Psi^{\mathrm{T}}Y \tag{16}$$

进行线性特征提取,其中 Y 为模式向量。

综上所述,求解具有统计不相关性的最优判别矢量集 $x_1,x_2,\cdots,x_d$ 的具体步骤可归纳如下:

Step1:计算总体散度矩阵 S_t 的所有特征值 $a_1,a_2,\cdots,a_n$ 及其对应的单位正交特征向量 $u_1,u_2,\cdots,u_n$;

Step2:令 $W=U\Lambda^{-1/2}$,其中 $U=(u_1,\cdots,u_n)$,$\Lambda=\mathrm{diag}(a_1,\cdots,a_n)$,则总体散度矩阵 $S_t=(W^{-1})^{\mathrm{T}}W^{-1}$,代入广义特征方程 $S_bx=\lambda S_tx$,得 $S_bx=\lambda(W^{-1})^{\mathrm{T}}W^{-1}x$;

Step3:令 $y=W^{-1}x$,则有 $W^{\mathrm{T}}S_bWy=\lambda y$;

Step4:令 $S=WS_bW$,则原广义特征方程转化为普通特征方程 $Sy=\lambda y$,并求其所有特征值 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$ 及其对应的特征向量;

Step5:选取前 $d(d\leqslant n)$ 个最大特征值 $\lambda_1,\lambda_2,\cdots,\lambda_d$ 对应的特征向量 $y_1,y_2,\cdots,y_d$,代入 $y=W^{-1}x$,即可得广义特征方程 $S_bx=\lambda S_tx$ 的前 d 个最大特征值对应的特征向量 $x_1,x_2,\cdots,x_d$,即可得具有统计不相关性的最优判别矢量集 $\Psi=(x_1,x_2,\cdots,x_d)$。

2 实验结果与分析

实验采用ORL人脸图像数据库,该数据库由40人,每人10幅 92×112 图像所组成,其中有些图像是拍摄于不同时期;人脸脸部表情与脸部细节有变化,例如:笑或不笑,眼睛睁着或闭着,带或不带眼镜;人脸姿态有变化,深度旋转与平面旋转可达到20°;人脸的尺度也有多达10%的变化。图2是ORL人脸数据库中的一些图像。

图2 ORL人脸图像数据库中的部分人脸

2.1 DCT系数数目对识别率的影响

实验首先分析需要多少DCT系数来描述人脸的整体信息,采用"留一法",即每次取一幅人脸图像作为测试样本,其余作为训练样本,在不同数目DCT系数的情况下使用最近邻分类器,实验结果如图3所示。

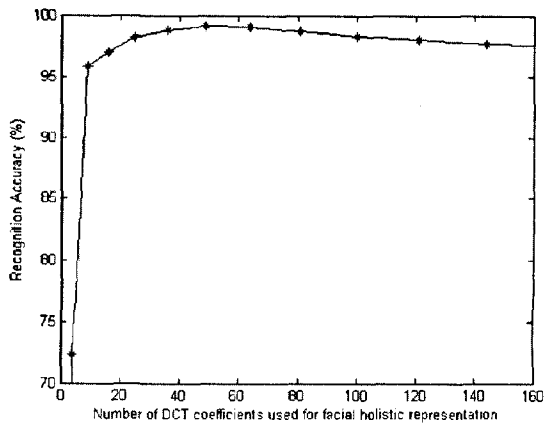

图3 基于DCT系数数目的识别结果

从图3可以看出,随着DCT系数数目的增加,识别率开始逐渐增加,到达峰值后缓慢地降低。在选用48个DCT系数时取得最好的识别率,继续增加系数数目,则识别率缓慢下降。这是由于新增加的系数对应于图像的高频成分,它们更易受到噪声的影响,从而影响识别率。

2.2 与其它方法识别率比较

随机选取每人 $k(k=3,4,5)$ 幅图像作为训练样本,剩余样本作为检测样本,采用最近邻分类器,分别采用三种方法进行识别实验,表2所示为三种方法对应的最高识别率。

表2 三种方法识别率比较(%)

方法\每人训练样本数目	3	4	5
Eigenfaces[14]	87.5	88.5	89.6
Fisherfaces	88.7	93.8	94.0
DCT-LDA	89.5	95.0	96

通过与Eigenfaces和Fisherface方法进行对比发现,就识别率而言,文中提出的方法要优于另外两种方法。从理论上讲,在对人脸图像降维过程中,对以均方误差为准则的人脸重建来说,PCA方法是一种最优方法,但从模式分类和辨识的观点来看,PCA方法则不

一定能够取得最优的结果。

可以看出，较之于以样本最优重建为目标的 PCA，具有"能量聚集特性"的 DCT 变换更适合于人脸图像的压缩，且对噪声干扰不敏感。

3 结束语

DCT 变换是信号处理和数据压缩中常用的一种方法，具有"能量聚集特性"和变换的保距特性。文中正是利用了 DCT 变换的特性，提出了基于 DCT 变换的线性判别分析方法，即在人脸识别的初期利用 DCT 变换来对人脸图像进行降维，由于 DCT 有成熟的快速算法，有效地解决了计算量大的问题。同时，也给出了在该模型下的最优判别矢量集的直接求解方法。实验证明了算法的有效性。

参考文献：

[1] Fukunaga K. Introduction to Statistical Pattern Recognition [M]. 2nd Edition. New York: Academic Press, 1990.

[2] Mika S, RVatsch G, MVuller K. A mathematical programming approach to the Kernel Fish algorithm[C]//in: Leen T K, Dietterich T G, Tresp V. (Eds.). Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2001: 591 – 597.

[3] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces versus fisherfaces: Recognition using class specific linear projection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19: 711 – 720.

[4] Zhao W. Discriminant component analysis for face recognition [C]//Proceedings of the International Conference on Pattern Recognition. Barcelona: [s. n.], 2000.

[5] Yang J, Jin Z, Yang J, et al. The essence of kernel Fisher discriminant: KPCA plus LDA[J]. Pattern Recognition, 2004, 37(10): 2097 – 2100.

[6] Swets D, Weng J. Using discriminant eigenfeatures for image retrieval[J]. IEEE Trans. Patt. Anal. and Mach. Intell., 1996, 18(8): 831 – 836.

[7] 杨 健. Fisher 线性鉴别分析的理论研究及其应用[J]. 自动化学报, 2003, 29(4): 481 – 493.

[8] Liu C, Wechsler H. Robust coding schemes for indexing and retrieval from large face databases[J]. IEEE Trans. Image Processing, 2000, 9(1): 132 – 137.

[9] 胡永刚. 高维数据降维的 DCT 变换[J]. 计算机工程与应用, 2006, 42(32): 21 – 23.

[10] Kailath T. Modern Signal Processing[M]. [s. l.]: Springer Verlag, 1985.

[11] Egecioglu. Dimensionality Reduction and Similarity Computation by Inner – Product Approximations[J]. IEEE Transactions on Knowledge and Data Engineering, 2004, 16 (6): 714 – 726.

[12] 程云鹏. 工程中的矩阵理论[M]. 天津: 天津大学出版社, 1998.

[13] Zhong J, Yang J Y, Hu Z S, et al. Face Recognition based on uncorrelated discriminant transformation[J]. Pattern Recognition, 2001, 33(7): 1405 – 1416.

[14] Turk M, Pentland A. Eigenfaces for recognition[J]. J. Cognitive Neurosci, 1991, 3(1): 71 – 86.

（上接第 96 页）

管理、设备维修管理、设备安全管理、知识库管理、系统服务管理等模块。

在各个模块中，应用轻量级开发思想，按表示层、业务逻辑层、持久层逐层设计，在查询和统计操作中，需要编写复杂的 SQL 语句和处理复杂的视图结构。经过比较选择，应用自己设计的持久层达到系统的性能要求，维护了系统的稳定。

3 结束语

文中详细介绍了这一持久层设计的过程，并通过和其它持久层方案的简单比较说明了这一持久层的特点和应用场合，同时 DAO 模式的应用提高了访问的透明度和降低了各层之间的耦合度，使得系统具备良好的伸缩性和移植性。

当然在应用这一持久层的同时也应该注意一些问题，比如增加实体 Bean 和 DAO 模式的采用不可避免地增加了系统的复杂度，另外在 DAO 层中的操作不能像 Hibernate 的 HQL 那么简便，还是得写出完整的 SQL 操作语句。

参考文献：

[1] Bauer C, King G. Hibernate in Action[M]. American: Manning Publications Co, 2004.

[2] Gorton P, Liu Yan, Trivedi N. An extensible, lightweight architecture for adaptive J2EE applications[M]. [s. l.]: ACM Press, 2006: 47 – 54.

[3] Berry C A, Crupi J, Malks D. J2EE 核心设计模式[M]. 北京: 机械工业出版社, 2002.

[4] 田 珂, 谢世波, 方 马. J2EE 数据持久层的解决方案[J]. 计算机工程, 2003, 29(22): 93 – 95.

[5] 魏 勇, 唐文彬, 郭 梅. 基于 DAO 模式的 J2EE 应用程序的数据库访问设计[J]. 计算机应用, 2003, 23(2): 356 – 357.