

基于 DCT 和线性判别分析的人脸识别

尹洪涛¹, 付 平¹, 沙学军²

(1. 哈尔滨工业大学自动化测试与控制研究所, 哈工大科学园 3033 信箱, 黑龙江哈尔滨 150080;
2. 哈尔滨工业大学通信技术研究, 黑龙江哈尔滨 150080)

摘 要: 提出基于离散余弦变换和线性判别分析的人脸识别方法. DCT 变换本身并不进行数据压缩, 它只是将图像源数据映射到另一个域, 如何在新的数据域中选择最有效的 DCT 系数作为识别特征成为关键问题. 本文从选择有效特征角度出发, 引入特征选择算法, 根据可分性判别确定将哪些 DCT 系数作为特征, 然后对选出的 DCT 系数进行线性判别分析提取识别特征. 在 ORL 人脸库上的实验结果证明了提出方法的有效性.

关键词: 人脸识别; 离散余弦变换; 线性判别分析

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2009) 10-2211-04

Face Recognition Based on DCT and LDA

YIN Hong-tao¹, FU Ping¹, SHA Xue-jun²

(1. Automatic Test and Control Institute Science Park of Harbin Institute of Technology, P. O. Box 3033, Harbin, Helongjiang 150080;
2. Communication Research Center, Harbin Institute of Technology, Harbin, Helongjinag 150080)

Abstract: A face recognition method based on the discrete cosine transform and linear discriminant analysis is presented. Discrete cosine transform cannot compress the data and it only maps the resource data to another data field. How to select the DCT coefficients that are most effective to classify is an important problem. In order to obtain the effective features, a feature selection algorithm according to the separability discriminant is use to select the DCT coefficients. The experiments on ORL and Yale face databases show that the improved method is effective.

Key words: face recognition; DCT(discrete cosine transform); LDA(linear discriminant analysis)

1 引言

人脸识别是生物特征识别中最受人们关注的一个分支. 近年来, 人脸识别已成为模式识别和人工智能领域中的一个非常活跃的研究方向. 过去几十年中该课题的研究已经取得了较大的进展, 众多研究者提出了很多识别方法, 如基于主成分分析的方法^[1,2]、基于线性判别分析的方法^[3,4]、基于独立成分分析的方法等. 离散余弦变换(DCT)是信号处理过程中常见的一种时域频域变换. 离散余弦变换是一种实数域变换, 其变换核为实数的余弦函数, 并具有快速算法, 而且离散余弦变换是次最佳变换, 变换后的系数分布比较集中, 因为图像的主要信息为低频信息, 而 DCT 在频域中将图像的主要信息集中在低频部分, 所以被广泛应用于语音及图像数据压缩领域.

DCT 系数也可以作为特征进行识别分类. Hafed^[5]等人首先提出一种基于 DCT 的人脸识别方法, 他们对整个人脸图像进行离散余弦变换, 再取很少的 DCT 系

数作为特征进行人脸识别. 同时 Ramasubramanian^[6]提出了离散余弦变换与主成分分析相结合的人脸识别方法. 文献[7]提出一种基于离散余弦变换与 LDA 相结合的人脸识别方法, 首先利用 DCT 将图像进行降维, 然后在低维空间中利用 LDA 进行特征提取. Chen^[8]等证明了直接在 DCT 域上和空域上进行主成分分析和线性判别分析可以获得相同的识别结果. 在这些方法中只是简单将低频 DCT 系数作为特征进行人脸识别, DCT 变换本身并不进行数据压缩, 它只是将图像源数据映射到另一个域, 如何在新的数据域中选择最有效的 DCT 系数作为识别特征成为关键问题. 已有的方法都是按矩形或“Z”字型顺序选择低频 DCT 系数作为特征进行人脸识别. 本章从选择有效特征角度出发, 对离散余弦变换与线性判别分析相结合的人脸识别方法进行了改进.

2 离散余弦变换

对于一幅 $M \times N$ 的图像 $f(x, y)$ 矩阵, 其离散余弦变换定义为:

$$C(u, v) = a(u) a(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x-1)u}{2M} \cos \frac{(2y-1)v}{2N} \quad (1)$$

其中, $C(u, v)$ 称为矩阵 $f(x, y)$ 的 DCT 系数. $u=0, 1, \dots, M-1, v=0, 1, \dots, N-1, a(u), a(v)$ 分别定义为:

$$a(u) = \begin{cases} \sqrt{1/M}, & u=0 \\ \sqrt{2/M}, & u=1, 2, \dots, M-1 \end{cases} \quad (2)$$

$$a(v) = \begin{cases} \sqrt{1/N}, & v=0 \\ \sqrt{2/N}, & v=1, 2, \dots, N-1 \end{cases} \quad (3)$$

经过 DCT 变换后, 图像的二维 DCT 系数构成一个与原图像大小相同的矩阵, 其低频系数集中在矩阵的左上角, 是图像中变化较慢的成分, 高频系数集中在矩阵的右下角, 是图像的细节和边缘成分.

3 线性判别分析

基于 Fisher 准则的线性判别分析是特征提取的最有效方法之一, 其目标是从高维特征空间里提取出最具有判别能力的低维特征, 这些特征能使同一类别的样本聚集在一起, 不同类别的样本尽量分开, 即选择使得样本类间离散度和样本类内离散度的比值最大的特征. 假设有 c 个模式类, 样本的类内散布矩阵 S_w 和类间散布矩阵 S_b 可以表示为:

$$S_w = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} (y_j^{(i)} - \bar{y}^{(i)}) (\bar{y}^{(i)} - y_j^{(i)})^T \quad (4)$$

$$S_b = \frac{1}{N} \sum_{i=1}^c N_i (\bar{y}^{(i)} - \bar{y}) (\bar{y}^{(i)} - \bar{y})^T \quad (5)$$

式中, N 表示样本总数; N_i 表示第 i 类样本的数量; $y^{(i)}$ 表示第 i 类样本的均值; $y_j^{(i)}$ 表示第 i 类中第 j 个样本; \bar{y} 表示所有样本的平均值.

Fisher 准则函数定义为:

$$J(W) = \frac{|W^T S_b W|}{|W^T S_w W|} \quad (6)$$

Fisher 准则函数将样本的类间离散度和类内离散度非常巧妙地结合在一起, 取极大化目标函数 $J(W)$ 的矩阵 W_{fld} 作为投影方向, 其物理意义是: 特征样本在 W_{fld} 上投影后, 类间离散度和类内离散度之比达到最大.

4 类别可分性判据

设有 c 类模式, 样本总数为 N , 各类样本的数量分别为 N_1, N_2, \dots, N_c , 令 $x_k^{(i)}, x_l^{(j)}$ 分别表示第 i 类和第 j 类的样本, 各类样本之间的平均距离定义为

$$\bar{d}(x) = \frac{1}{2} \sum_{i=1}^c P_i \sum_{j=1}^c P_j \frac{1}{N_i N_j} \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} d(x_k^{(i)}, x_l^{(j)}) \quad (7)$$

式中, P_i, P_j 表示相应类别的先验概率, $d(x_k^{(i)}, x_l^{(j)})$ 表

示 $x_k^{(i)}$ 和 $x_l^{(j)}$ 之间的距离, 对于欧氏距离有

$$d(x_k^{(i)}, x_l^{(j)}) = \left[(x_k^{(i)} - x_l^{(j)})^T (x_k^{(i)} - x_l^{(j)}) \right]^{\frac{1}{2}} \quad (8)$$

样本总类内散布矩阵定义为

$$S_w = \sum_{i=1}^c P_i \frac{1}{N_i} \sum_{k=1}^{N_i} (x_k^{(i)} - m_i) (x_k^{(i)} - m_i)^T \quad (9)$$

样本总类间散布矩阵定义为

$$S_b = \sum_{i=1}^c P_i (m_i - m) (m_i - m)^T \quad (10)$$

样本总体散布矩阵定义为

$$S_t = \frac{1}{N} \sum_{i=1}^N (x_i - m) (x_i - m)^T \quad (11)$$

式中, m_i 表示第 i 类样本的平均矢量, m 表示所有样本的平均矢量, 分别定义如下

$$m_i = \frac{1}{N_i} \sum_{k=1}^{N_i} x_k^{(i)} \quad (12)$$

$$m = \sum_{i=1}^c P_i m_i \quad (13)$$

用各类样本的样本频率 N_i/N 表示先验概率 P_i , 可得

$$J = \frac{1}{d^2(x)} = \text{Tr}[S_w + S_b] = \text{Tr}[S_t] \quad (14)$$

所以可以将样本总体散布矩阵的迹作为类别的可分性判据, 其值越大, 表示样本越分散.

5 提出的方法

在人脸识别系统中, 提取人脸图像特征的目的有两个, 一是降低图像的维数, 减小分类时计算的复杂程度; 另一个是选出最有代表性的特征提高分类性能.

人脸图像经过离散余弦变换后, 可得到一个与原图像大小相等的系数矩阵, 如何选取 DCT 系数来有效地表示和区分人脸成为关键性问题. 如图 1 所示为一幅人脸图像及其 DCT 系数分布情况, 1(b) 是人脸图像的 DCT 系数矩阵. 从 DCT 系数的分布可以看出, 人脸图像经 DCT 变换后, 数值大的频谱系数主要成分集中在比较小的范围, 且主要位于低频部分. 该矩阵左上角的第一个数据是人脸图像的平均亮度值, 后面各项系数的分布和大小可以反映亮度起伏的剧烈程度. 若系数较大, 说明该频率的分量在人脸图像中变化明显, 若系数小, 则说明该频率的分量在人脸图像中变化不明显. 因此, DCT

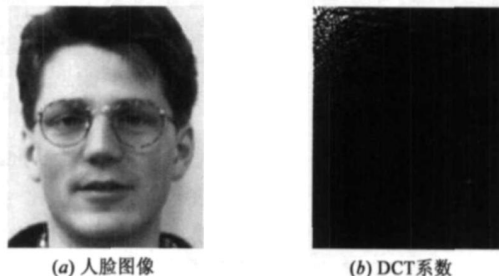


图1 人脸图像及其DCT系数

能将大量的信息集中在相对较少的低频系数上。

从图像压缩和重构的角度出发,可以选择系数矩阵左上角的低频部分,它代表图像信息的低频分量,是图像信息的主体,即图像亮度变化的主体部分。但是从识别的角度,这样做并不一定能取得最佳的识别效果。为了挑选出更有效的特征,我们引入特征选择算法,根据可分性判据确定将哪些 DCT 系数作为识别特征。

本文提出的基于 DCT 和线性判别分析的人脸识别方法框图如图 2 所示。

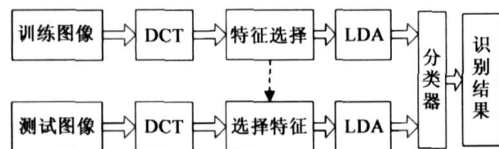


图2 基于DCT和LDA的人脸识别框图

训练过程主要分为以下几步：

(1) 首先对训练集中每一幅人脸图像进行离散余弦变换,求出 DCT 系数。

(2) 然后分别在不同频率的 DCT 系数上计算样本总体散布矩阵的迹,并以该值作为类别的可分性判据,按这些值的大小进行排序。

(3) 将迹值较大频率的 DCT 系数作为人脸的特征进行线性判别分析,求得最佳投影矩阵和训练样本的识别特征。

识别过程主要分为以下几步：

(1) 对于一幅待识别的人脸图像,首先对图像进行离散余弦变换,求出 DCT 系数。

(2) 然后按训练过程中选定的顺序选取 DCT 系数,并将所选定的 DCT 系数向训练过程中确定的最佳投影矩阵投影,获得识别特征。

(3) 最后利用最近邻分类器对人脸图像进行分类。

6 实验

为了验证所提出算法的性能,实验在 ORL 人脸库上进行,在 ORL 人脸库中有 40 个人的 400 幅图像,每人 10 幅,这些人脸图像是在不同时间,不同光照,不同头部角度,不同面部表情和不同人脸细节条件下拍摄得到的。图像尺寸为 92 × 112 像素。

首先比较特征选择方法对识别效果的影响。每人随机选取 5 幅图像作为训练样本,其他图像用于测试。先对样本进行离散余弦变换,然后分别使用三种方式选择 DCT 系数,一种是选取图像的 DCT 系数中低频部分的一个正方形子块作为特征,一种是按“Z”字形顺序从低频部分开始选取,另一种是按本文所采用可分性判据进行选择。在 DCT 系数选定之后,再对选定的系数进行线性判别分析,提取判别特征,最后使用最近邻分类器进行分类。每组实验都重复 20 次,取平均值作为识别结果。

实验结果如图 3 所示。

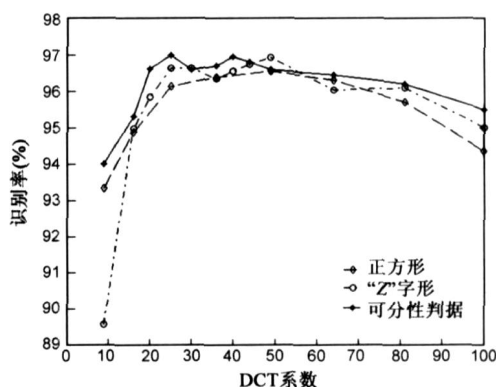


图3 不同DCT系数选择方法的识别率比较

按选取图像的 DCT 系数中低频部分的一个正方形子块作为特征的方法,在使用 49 个 DCT 系数时,该方法达到最高识别率为 96.55 %;按“Z”字形顺序从低频部分开始选取 DCT 系数的方法,也是在使用 49 个 DCT 系数时,达到最高识别率为 96.92 %;而按本节所采用可分性判据进行选择 DCT 系数的方法,是在使用 25 个 DCT 系数时,达到最高识别率为 97.00 %。所以从最高识别率看,按可分性判据选取 DCT 系数的方法取得了最高识别率,而且所用的 DCT 系数最少。从整体来看,效果最差的是按正方形方法选取 DCT 系数,效果最好的是按可分性判据选择 DCT 系数,在大多数情况下,识别率均高于另外两种方法。

传统基于 DCT 的人脸识别方法,基本上选择低频 DCT 系数作为识别特征,而本文提出的方法是在全部 DCT 系数范围内进行选择,所以能够选择出更有利于分类的特征,从而提高识别性能。

还同文献[5,7]中的两种基于离散余弦变换的方法进行了比较。在文献[5]中将 DCT 系数矩阵中左上角正方形内的低频系数直接用于分类,我们称之为 DCT 方法;在文献[7]中,先对图像进行降维处理,再对降维后的人脸图像进行离散余弦变换,然后选取 DCT 系数矩阵中左上角正方形内的低频系数进行线性判别分析,提取识别特征,我们称之为 DCT+LDA 方法。

实验中,每人分别随机选取 5 幅图像作为训练样本,其它图像用于测试,每组实验重复进行 20 次,取平均值作为识别结果。识别结果如表 1 所示。

表 1 三种方法在 ORL 人脸库上的识别率(%)比较

DCT 系数	16	25	36	49	64	81
DCT	94.45	94.77	94.95	94.43	94.97	94.10
DCT+LDA	94.85	96.62	96.45	96.65	96.45	95.85
提出方法	95.30	97.00	96.70	96.60	96.45	96.20

从表 1 可以看出,本文提出的方法只使用 25 个 DCT 系数就取得了最佳识别效果,我们还可以发现三种方法中,所使用 DCT 系数的数量与所取得的识别率并不成

正比,使用太多的 DCT 系数并不能取得最好的识别率.在重构一幅人脸图像时,要准确重构原始图像需要尽可能多的 DCT 系数,但是对于识别并不需要太多的 DCT 系数.

7 结论

本文研究了基于离散余弦变换和线性判别分析的人脸识别方法.针对 DCT 系数选择问题,我们从选择有效特征角度出发,引入特征选择算法,根据可分性判别选择 DCT 系数进行线性判别分析,仿真实验验证了提出方法的有效性.

参考文献:

- [1] L Sirovich, M Kirby. Application of Karhunen-Loeve procedure for the characterization of human faces[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1990, 3(1): 71 - 79.
- [2] M Turk, A Pentland. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 72 - 86.
- [3] D L Swets, J Y Weng. Using discriminant eigenfeatures for image retrieval[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1996, 18(8): 831 - 836.
- [4] P N Belhumeur, J P Hespanha, D J Kriegman. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1997, 19(7): 711 - 720.
- [5] Z M Hafed, M D Levine. Face recognition using the discrete cosine transform[J]. International Journal of Computer Vision, 2001, 43(3): 167 - 188.
- [6] D Ramasubramanian, Y V Venkatesh. Encoding and recognition of faces based on the human visual model and DCT[J]. Pattern Recognition, 2001, 34(12): 2447 - 2458.
- [7] 张燕昆, 刘重庆. 一种新颖的基于 LDA 的人脸识别方法[J]. 红外与毫米波学报, 2003, 22(5): 327 - 330.
- ZHANG Yan-Kun and LIU Chong-Qing. A novel face recognition method based on linear discriminant analysis[J]. Journal of Infrared and Millimeter Waves, 2003, 22(5): 327 - 330. (in Chinese)
- [8] W Chen, J E Meng, S Wu. PCA and LDA in DCT domain[J]. Pattern Recognition Letters, 2005, 26(15): 2474 - 2482.

作者简介:



尹洪涛 男, 1972 年出生于黑龙江海林市, 博士, 哈尔滨工业大学自动化测试与控制系统师. 主要研究方向为数字图象处理和模式识别.
E-mail: yinht@hit.edu.cn



付平 男, 1965 年生于黑龙江哈尔滨, 博士, 现为哈尔滨工业大学自动化测试与控制系统教授, 博士生导师, 中国电子学会及中国计量测试学会高级会员, 主要研究方向为计算机自动测试与控制, 图像处理.
E-mail: fuping@hit.edu.cn

沙学军 男, 1966 年生于黑龙江齐齐哈尔, 博士, 现为哈尔滨工业大学通信技术研究所教授, 博士生导师, 主要从事通信网及专用移动通信系统和宽带无线接入的研究.