

压缩域上人脸识别的研究^{*}

廖红文¹, 冯国灿¹, Jiang Jianmin²

(1. 中山大学数学与计算科学学院, 广东 广州 510275;

2. School of Informatics, University of Bradford, UK)

摘要: DCT 变换被广泛采用在图象和视频压缩标准中 (如 JPEG, MPEG, H.261/H.263)。而对于这些压缩图象的处理, 传统的手段是先解压到空间域, 再进行处理和识别, 因而增加了计算复杂性。针对这个问题, 运用压缩域上图像处理技术, 提出了人脸特征表达的方法, 并构造 DCT 域上人脸识别系统。在人脸标准库 FERET 上的测试, 提出的算法与经典的特征脸法和 PCA+LDA 法相比较, 不仅提高了精度, 而且降低了计算复杂性。

关键词: 人脸识别; 余弦变换 (DCT); 主成分分析 (PCA); 压缩域

中图分类号: O29; TN911.73 文献标识码: A 文章编号: 0529-6579 (2004) 05-0016-04

近 20 年来。由于计算机技术的发展和人脸识别的应用需求, 人脸识别变成是模式识别最活跃的领域之一。研究者们提出了许多的人脸识别方法, 主要有基于几何特征的人脸识别方法、基于模板匹配的人脸识别方法、特征脸方法、隐马尔科夫模型、神经网络方法、弹性图匹配方法、小波变换方法等^[1]。这些研究对提高人脸识别率起到了非常重要的作用, 使人脸识别技术日臻成熟。但是这些研究大多数是基于非压缩人脸图像数据的, 包括实验用的一些数据库, 如 Yale、ORL、MIT, 也大多数是 BMP 格式的图像。文献[1]综述了近几年人脸识别领域主要研究方法和进展。

随着社会管理的电子化、网络化, 海量的人脸图像须经数字化后归档到人像库, 如国家安全部门统一管理全国的人口档案信息, 便于全国乃至全世界各个准入场所调用。为了有效存贮海量的人脸图像, 必须对图像进行压缩编码后归档。JPEG 图像压缩标准就是海量人脸图像库提供了一种有效选择。而传统的人脸识别方法, 在对压缩的人脸图像进行识别时, 要先进行图像的“全解压”操作, 然后再进行识别, 这无疑大大限制了识别系统性能的提高。

近几年, 基于压缩域上的图像处理技术日益受到重视^[2-8], 文献[9]综述了压缩域图像特征提取的研究概况。针对 JPEG 图像库, 文献[4]综述了压缩域上的基于图像内容的检索算法。在人脸识别领

域, 针对传统的 PCA+LDA 方法, 文献[5]提出 DCT+LDA, 此方法不仅改善了识别精度, 而且只对 JPEG 图像部分解压, 从而降低了系统复杂性, 提高了识别速度。事实上, 与 PCA 比较, DCT 变换具有相近的信息压缩表现。DCT 系数不仅表达了图像的频率特性, 而且可直接从压缩域获得。鉴于此, 本文从压缩域上的图像特征提取技术出发, 提出一种基于 DCT 域上的人脸图像的特征提取方法, 并建立了人脸识别系统。通过在人脸标准库 FERET 测试, 与经典的特征脸 (Eigenface) 法, 及 PCA+LDA 法比较, 不仅提高了精度。而且降低了计算复杂性。

1 压缩域上图像特征提取技术

压缩域上的图像特征提取技术, 是指通过挖掘图像压缩时的中间结果或最终码流中包含的信息, 力争在不解码或部分解码的情况下提取表征图像内容的特征, 从而实现图像处理的目的。基于压缩域上的图像特征提取, 最理想的情况是直接最终码流上提取特征, 但是, 由于熵编码属于非结构化、非字节对齐编码, 在熵编码后进行这些操作十分困难, 通常都是在熵解码和反量化后进行。压缩域上图像特征提取技术大致上可以分为两类: 基于变换域和基于空域的技术。其中基于变换域的技术主要有离散傅立叶变换域 (DFT)、KL 变换域、离散余弦变换域 (DCT)、离散小波变换域 (DWT) 上的

* 收稿日期: 2004-04-01

基金项目: 教育部重点基金资助项目 (104145); 广东省自然科学基金资助项目 (031609)

作者简介: 廖红文 (1968 年生), 女, 在职硕士研究生, 现任广东女子职业技术学院讲师;

通讯联系人: 冯国灿; E-mail: mcsgc@zsu.edu.cn

图像特征提取技术, 基于空域的技术主要有矢量量化方法、分形/仿射变换方法。

DFT 是图像识别中的一个重要的数学变换。在图像压缩编码中, 利用 DFT 主要是由于它所具有的能量集中性; 在图像检索和模式匹配中, 利用 DFT 主要是因为它具有 DFT 系数幅值的平移不变性和 DFT 系数可用来有效地计算图像的空域相关性。据研究, 一般 DFT 系数半径的分布对纹理的粗度很敏感, 而相位的分布则对纹理的方向很敏感, 因而具有较好的分类性能^[7]。然而, 由于 DFT 本身有产生周期性结果的趋势^[8], 对于非周期性的人脸图像, 其压缩能力往往不够理想; 另外, DFT 是一种复数变换, 运算量大, 实用困难, 因此, DFT 很少被用于图像压缩, 从而使得基于 DFT 压缩域的图像处理能力也就无法很好体现。

特征脸方法是基于 KL 变换上的一种经典的人脸识别方法, 因而在 KL 变换域上进行特征提取也曾经受到关注。从能量压缩的角度来看, KL 变换是最优的。但是, 由于 KL 变换对外在因素带来的图像差异和人脸本身带来的差异是不加区分的, 另外, KL 变换的计算复杂度庞大, 所以实际应用中, 很少有图像用 KL 变换进行压缩, 使得基于 KL 变换域的图像特征提取不是一种实用的方法。

相对而言, 由于目前的编码标准中普遍采用的是基于 DCT 变换和 DWT 变换的压缩算法, 使得基于这两种压缩域上的图像特征提取技术成为压缩域上图像特征提取技术的主流^[4]。

对于 $M \times N$ 二维图像 $x(i, j)$, 其 DCT 变换 $C(u, v)$ 定义为

$$C(u, v) = \frac{2}{\sqrt{MN}} \alpha(u) \alpha(v) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x(i, j) \cdot \cos\left[\frac{(2i+1)u\pi}{2M}\right] \cos\left[\frac{(2j+1)v\pi}{2N}\right]$$
$$u = 0, 1, \dots, M-1, v = 0, 1, \dots, N-1 \quad (1)$$

这里, $\alpha(u) = \begin{cases} \sqrt{1/2} & \text{for } u = 0 \\ 1 & \text{otherwise} \end{cases}$, 相应的逆变换 (IDCT) 可表示为

$$x(i, j) = \frac{2}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u) \alpha(v) C(u, v) \cdot \cos\left[\frac{(2i+1)u\pi}{2M}\right] \cos\left[\frac{(2j+1)v\pi}{2N}\right] \quad (2)$$

根据 DCT 定义, 可导出其在图像压缩和特征表达方面的良好特性: ①KLT 是能量压缩上的最优者, 而 DCT 的能量集中效率与 KLT 几乎相同, 并且使用与数据无关的固定基, 还存在快速算法, 这使得

DCT 识别系统本身具有简单、高效的特点^[9]; ②在数据压缩领域, 由于 DCT 具有优良性能且未对数据结构作任何假定 (例如, DFT 假定所变换的数据是周期的), 因而 DCT 技术成为 JPEG、MPEG 系列的关键技术^[8]; ③由于 DCT 系数可以很好地描述图像特征, 因而被用于压缩域上的图像特征提取中; ④任何 JPEG 图像的 DCT 块可由 8×8 的 DCT 块直接生成, 这建起了 DCT 域与像素域之间的桥梁, 使运用 DCT 作为特征来进行人像快速识别成为可能性^[9]; ⑤如 FFT 一样具有尺度不变性^[9], 即对于 $M \times N$ 二维图像 $x(i, j)$, DCT 域的特征 $\frac{1}{\sqrt{MN}}C(u, v)$ 具有不变性。

另外, 由于人脸图像压缩大多数采用 JPEG 标准, 包括 FERET 人像数据库中也是 JPEG 图像, 而 JPEG 标准采用的是 DCT 技术, 因此研究 DCT 域上的人脸识别技术具有现实性。本文将着重讨论如何在 DCT 域上进行人脸识别的问题。

2 人脸识别算法设计

在 DCT 域上进行人脸识别就是在原始图像集的 DCT 频率域中提取图像的低频系数建立特征库; 再利用同样的方法, 提取待识别图像的低频系数, 然后将待识别图像已简化的信息 (系数) 与特征库中的信息 (系数) 进行比较, 计算欧氏距离, 得出识别结论的方法。下面, 我们从 JPEG 压缩模型出发, 阐述在 DCT 域上实现人脸识别的方法。

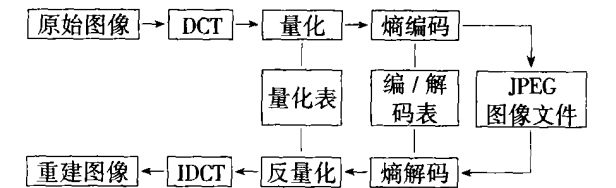


图1 JPEG 压缩模型编/解码过程
Fig.1 Diagram for JPEG coding/decoding

JPEG 是目前广泛采用的静态图像压缩的国际标准。其编/解码过程如图 1 所示^[10], 具体步骤如下: ① 输入图像被分成 8×8 像素块, 对每块进行二维正向离散余弦变换; ② 利用 64 个元素的量化表, 对 64 个 DCT 系数进行量化, 去掉人眼不敏感的高频成分; ③ 熵编码: 利用编码表对 64 个系数进行 Huffman 编码; ④ 压缩数据: 对经编码的数据块进一步压缩数据, 结果产生压缩的数据流; ⑤ 解压缩重建图像为上述过程的逆过程。

基于 DCT 系数的良好性态和频率特征, 建立的在 DCT 域上人脸特征的表达和识别的步骤表述如下: ① 先从人脸图像库中提出 JPEG 格式的人脸

压缩图像; ② 利用 Huffman 码表对压缩图像文件解码; ③ 利用量化表对 Huffman 解码文件反量化, 得到一系列 8×8 子块的 DCT 系数; ④ 再利用文献 [1] 的算法, 将该人像的所有 8×8 子块的 DCT 系数合成整幅人像的 DCT 表达, 其表达式为

$$C_B = \frac{1}{M} A \begin{bmatrix} C_{0,0} & C_{0,1} & \cdots & C_{1,k-1} \\ C_{1,0} & C_{1,1} & \cdots & C_{1,k-1} \\ \vdots & \vdots & \cdots & \vdots \\ C_{k-1,0} & C_{k-1,1} & \cdots & C_{k-1,k-1} \end{bmatrix} A^T \quad (3)$$

其中 $C_{i,j}$ ($i, j=0, 1, \cdots, k-1$), C_B , 分别为 8×8 块和整幅图像的 DCT 系数, A 是已知正交矩阵。其形式可以用图 2 表示。因为 A 是稀疏矩阵, 由 (3) 获得全局 DCT 比先实施 8×8 块 IDCT 再进行 DCT 变换极大降低计算复杂性;

⑤提取整幅人脸图像 DCT 域上的低频系数做为人脸的基本特征, 设第 P 人的特征为

$$X_P = (x_{p1}, x_{p2}, \cdots, x_{pr})^T$$

⑥用同样的方法提取待识别图像的低频系数

$$q = (q_1, q_2, \cdots, q_r)^T$$

⑦利用欧氏距离算法, 比较特征库与待识别图像的系数, 取得识别结果。

$$p = \arg \left(\min_p \sum_{i=1}^r (x_{pi} - q_i)^2 \right) \quad (4)$$

上面算法中, Huffman 解码和反量化的过程仅仅是一个查找表映射的算法, 所以时间花费很少, 而且, 据研究, 在传统图像处理场合中, DCT 与 IDCT 的计算量各占总计算量的 40%^[9], 这种算法省去了 IDCT 这一步, 可以明显提高系统效率。

在 DCT 域上的人脸识别过程可以用图 3 来描述。另外, 在 DCT 域上提取特征过程中, 有两点需要说明:

(1) 提取阶段, 选择低频系数作为图像的主要特征, 原因在于低频系数包含了图像更多的能量, 并且能够平滑由脸部表情或少许遮掩引起的变化^[9];

(2) 变换后的矩阵中第一个系数 $C(0, 0)$ 中包含了图像平均亮度的信息, 这个系数在提取特征时

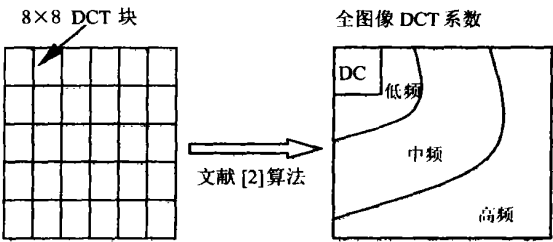


图 2 将所有 8×8 子块的 DCT 系数合成整幅 DCT 表达
Fig. 2 Merge all 8×8 DCT blocks into one global DCT

应予以剔除, 这样可以减少整个图像中光照的影响^[9]。

3 实验设计及其识别结果

实验数据来自的 FERET 人脸数据库。在所利用的 FERET 人像数据库中, 包含了 246 个人, 每个人两幅图像, 都是 JPEG 格式, 其中一幅图像用作训练, 另一幅用来作识别, 图 4 是其中一组图像。每个待识别图像都与训练库中的 246 个人进行了比较, 共进行了 246×246 次比较。每个人的图像上有表情、光照、轻度偏转等差异。该图像库是经过预处理、标准化后的图像。因此可以直接用来进行特征提取和识别工作。所有识别方法都使用欧氏距离进行最小距离分类。这组实验在迅驰 1.4G 计算机上用 Matlab 编程实现。提出的算法与特征脸法的比较见表 1。

实验表明: 利用 DCT 直接在压缩域上进行特征提取和识别, 不但有较高的识别率 (单幅训练图像识别率最高可达 75.6%), 而且系统消耗少。当人像库中的数目很大时, 利用 DCT 做识别具有很好的优势。

4 讨 论

目前, 压缩域上的人像识别研究比较少。文献 [3] 提出一种在 DCT 域上对 JPEG 人脸图像进行识别的算法。这是一种以 JPEG 文件经熵解码和反量化后得到的 DCT 系数为样本, 利用隐马尔科夫模

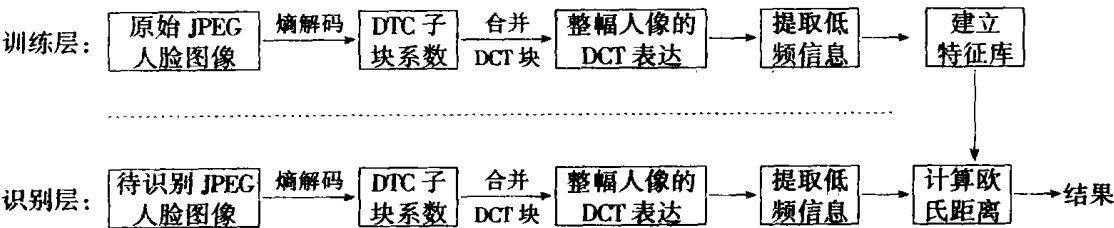


图 3 DCT 域上的人脸识别过程

Fig. 3 Diagram for face recognition in DCT Domain



图 4 来自于 FERET 人像库的一组图像

Fig 4 Face image samples from FERET database

表 1 PCA 与 DCT 的识别结果比较¹⁾

Tab 1 Comparison in recognition rate between PCA and DCT

算法	<i>r</i> = 20		<i>r</i> = 30		<i>r</i> = 40		<i>r</i> = 400	
	识别率 /%	时间 /s	识别率 /%	时间 /s	识别率 /%	时间 /s	识别率 /%	时间 /s
PCA 算法	68.7	18×60	69.1	18×60	68.7	18×60	—	—
DCT 算法	70.7	7	72.8	7	73.2	7	75.6	8

1) *r* 代表选取的特征值数目, 运行时间包括训练时间

型(HMM)进行特征提取并识别的算法。由于 HMM 包含迭代过程, 因此算法复杂性高。文[5] 中 DCT + LDA 方法, 能较好地解决 LDA 避开 LDA 中的小样本问题, 但比本文提出的算法多一道过程。本文提出的算法, 通过 FERET 测试, 与经典的特征脸 (Eigenface) 法比较, 和 PCA + LDA 比, 不仅提高了精度。而且降低了计算复杂性, 且容易实现。

参考文献:

[1] ZHAO W, CHELLAPPA R, ROSENFELD A, et al. Face recognition: a literature survey[C] . UMD CFAR Tech Report, 2002, 948.

[2] JIANG J, FENG G C. The spatial relationship of DCT coefficients between a block and its sub blocks[J] . IEEE Trans on Signal Processing, 2002, 50(5) : 1160—1169.

[3] EICKELER S, MULLER S, RIGOLL G. High quality face recognition in JPEG compressed images // Proc of the IEEE 1999 International Conference on Image Processing [C] . Kobe, Japan, 1999, 1: 672—676.

[4] 李晓华, 沈兰荪. 基于压缩域的图像检索技术[J] . 计算机学报, 2003, 26(9) : 1051—1059.

[5] FENG G C, JIANG J. Image extraction in DCT domain[J] . IEE Proc Vis Image Signal Process, 2003, 150(1) :

[6] 刘艳, 李宏东. DCT 域图象处理和特征提取技术[J] . 中国图像图形学报, 2003, 8(2) : 121—128, 499—508.

[7] AUGUSTEIJN M, CLEMENS L E, SHAW K A. Performance evaluation of texture measure for ground cover identification in satellite images by means of neural network classifier[J] . IEEE Transactions on Geoscience & Remote Sensing, 1995, 33(3) : 616—626.

[8] David Salomon. 数据压缩原理与应用[M] . 吴乐南, 等译. 北京: 电子工业出版社, 2003.

[9] MANDAL M K, IDRIS F, PANCHANATHA S. A critical evaluation of image and video indexing techniques in the compressed domain[J] . Image and Vision Computing, 1999, 17: 513—529.

[10] WALLACE G K. The JPEG still picture compression standard[J] . Communication of the ACM, 1991, 34(4) : 31—45.

Face Recognition in Compressed Domain

LIAO Hong wen¹, FENG Guo can¹, JIANG Jian min²

(1. School of Mathematics and Computational Sciences, Sun Yat sen University, Guangzhou 510275, China;
2. School of Informatics, University of Bradford, UK)

Abstract: The discrete cosine transform is widely used in digital compression techniques for images and videos. The conventional approaches for the compressed image processing is to transfer the image to the spatial domain first, then to employ the traditional image processing techniques in spatial domain. This is certainly computational cost. A method is proposed for representing face feature directly extracted in compressed domain of JPEG, and a face recognition system in DCT domain is established. The experiments on FERET face database show the proposed algorithm can achieve higher recognition rate and lower the computational cost in comparison with that of eigenface and PCA + LDA.

Key words: face recognition; discrete cosine transform; principal component analysis; compressed domain