

Statistical Analysis-based DCT Method for Face Recognition

Yang Xiaoshang Yang Cong

Ideal Institute of Information and Technology
Northeast Normal University, NENU
Changchun, China
yxs5669@163.com yangcong955@126.com

Wang Wenyong

Engineering & Research Center of E-learning
Changchun, China
wenyongw@yeah.net

Abstract—It's critical for feature recognition of DCT (Discrete Cosine Transform) coefficient selection in the new Data Field which commutated from the DCT; however, the DCT commutating does not have the capability for data compression. This paper proposes a new method combined with the DCT commutates and the statistical analysis to select DCT's coefficient. Firstly, calculate each capacity of DCT's coefficients. Secondly, select the data as a feature which has a higher discriminate ability, and then combined with the PCA method for data decorrelation. Finally, do the classification and recognition. Experiments show that this method of classification efficiency greatly improved compared to traditional methods.

Keywords—Face Recognition; DTC; Data Dispersion

I. 引言

人脸识别技术是通过计算机程序提取人脸的特征，并且根据这一特征进行身份识别的一种技术。它主要通过三个步骤完成识别任务：人脸检测，人脸特征提取以及分类。在过去几十年中，对该领域的研究方法层出不穷，如基于主成分分析法[1]，模板匹配识别算法，基于线性判别的方法[2,3]，神经网络算法等。离散余弦变换（Discrete Cosine Transform, 简称 DCT）是一种常用的图像数据压缩方法，它是具有近似于最佳 K-L 变换的数据压缩能力和类似傅立叶变换的快速算法，其数据压缩能力仅次于 K-L 变换，同时还可以保留对光照、表情和姿态等不敏感类别信息。在 DCT 变换之后，选取 DCT 系数对进行图像还原和类别识别至关重要。本文在 DCT 变换的基础上针对传统的 $m \times m$ 矩阵法和 Zig-Zag 法提出一种新的基于统计原理的方法提取具有最优分类能力的 DCT 系数。本文首先对所有图像进行 DCT 变换，提取图像的 DCT 系数，然后计算每个 DCT 系数的判别能力，最后选取判别能力高的系数进行图像还原以及识别。实验证明本方法在时间效率、分类能力上均优于传统方法。

II. 离散余弦变换

对于一副人脸图像，运用基于整体的思想，将其灰度分布作为一个特征模式来提取特征。我们试图通过对预处理之后的图像进行一系列的变换，以得到的系数来表征原图像信息，然后运用这些系数形成的向量来进行分类识别。

对于一副 $M \times N$ 的图像，求取 DCT 系数矩阵 F 的变换公式如下：

$$F(u, v) = \frac{1}{\sqrt{MN}} \alpha(u) a(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \times \cos\left(\frac{(2x+1)u\pi}{2M}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right) \\ (u = 0, 1, \dots, M, v = 0, 1, \dots, N). \quad (1)$$

其中 $\alpha(w) = \begin{cases} 1/\sqrt{2} & w=0 \\ 1 & \text{else} \end{cases}$ ， F 为图像的灰度矩阵。

选用 YALE 人脸数据库中的一副图片（大小为 100×100 ），经过 DCT 变换后的 DCT 人脸图像如图 1 所示：

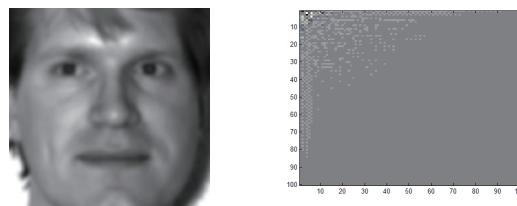


图 1. 原始图像（左）与 DCT 变换后的人脸图像（右）

之后通过 DCT 逆变换 Inverse Discrete Cosine Transformation (负离散余弦变换, 简称 IDCT) 来还原图像：

$$f(x, y) = \frac{1}{\sqrt{MN}} \sum_{u=1}^{M-1} \sum_{v=0}^{N-1} \alpha(u) a(v) F(u, v) \times \cos\left(\frac{(2x+1)u\pi}{2M}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad (2)$$

选用 YALE 人脸数据库中的一副图片（大小为 100×100 ），经过 DCT 变换后的人脸以及用 IDCT 重建的图像如图 2 所示：



图 2. IDCT 重建后图像

III. DCT 系数的选择

利用所有的 DCT 系数重建特征向量，显然会带来庞大的计算复杂度。实际上，不同的 DCT 系数的分类能力并不一样。从图 1 右边 DCT 变换后人脸图像可以看出分量的频率自左上角向右下角递减。通常低频区反映了光照影响，高频区反映数据库的噪声信息，中频区域则包含了大量有用的信息。所以通常在中频区选取系数用于构建特征向量从而重建图像。

A. 传统的系数选择法

利用所有的 DCT 系数重建特征向量，显然会带来庞大的计算复杂度。实际上，不同的 DCT 系数的分类能力并不一样。从图 1 右边 DCT 变换后人脸图像可以看出分量的频率自左上角向右下角递减。

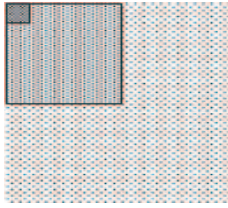


图 3. IDCT 系数频率分布

如图 3 所示，我们把 DCT 系数矩阵分为三个区域：低频区域，中频区域，高频区域。通常低频区反映了光照影响，高频区反映数据库的噪声信息，中频区域则包含了大量有用的信息。所以通常在中频区选取系数用于构建特征向量从而重建图像。传统的做法是使用一个 $m \times m$ 矩阵提取左上角的系数值，另一个被经常使用的方法是 Zig - Zag 方法，如图 4：

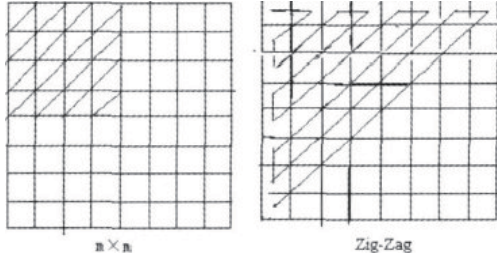


图 4. 两种提取 DCT 系数方法

以上方法虽然简单，但不严格，并且显然它们取得的实验效果因被训练数据库而异。于是我们尝试运用统计分析的方法使之通过训练适应于不同的数据库，从而达到相比于传统的方法提高识别率的目的。

B. 基于统计分析的 DCT 系数选择方法

在 DCT 系数选择初期，由于基于 DCT 系数在不同区域分类能力的认识，即低频反映了光照的变化，而高频则反映了噪音，为了回避光照影响并且降低噪音干扰，我们只选取中频区的 DCT 系数进行计算。

具有强分类能力的系数应该具备两个性质：尽量拉大类间距离，尽量缩小类内距离。首先我们对一幅图像产生其 DCT 矩阵，如公式 3 所示：

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ x_{M1} & x_{M2} & \cdots & x_{MN} \end{bmatrix}_{M \times N} \quad (3)$$

我们在一个含有 C 类人脸图像的数据库中，每类选取 S 幅图片作为测试图片，得到包含 $C \times S$ 幅图像的测试集合。然后选取每类的每张图像的 DCT 阵的第 (i, j) 位置上的值构建一个矩阵 X_{ij} ：

$$X_{ij} = \begin{bmatrix} x_{ij}(1,1) & x_{ij}(1,2) & \cdots & x_{ij}(1,C) \\ x_{ij}(2,1) & x_{ij}(2,2) & \cdots & x_{ij}(2,C) \\ \vdots & \vdots & \vdots & \vdots \\ x_{ij}(S,1) & x_{ij}(S,2) & \cdots & x_{ij}(S,C) \end{bmatrix}_{S \times C} \quad (4)$$

其中， $x_{ij}(S,C)$ 表示第 C 类第 S 副图片的 DCT 系数。

之后计算每类的平均值：

$$M_{ij}^c = \frac{1}{S} \sum_{s=1}^S X_{ij}(s,c), c=1,2,\dots,C. \quad (5)$$

进而计算得到样本的类间离散度 S_{ij}^B 和类内离散度 S_{ij}^W ：

$$S_{ij}^B = \sum_{s=1}^S (X_{ij}(s,c) - M_{ij}^c)^2, c=1,2,\dots,C. \quad (6)$$

$$S_{ij}^W = \frac{1}{C} \sum_{c=1}^C V_{ij}^c. \quad (7)$$

定义所有训练图片的 DCT 系数在位置 (i, j) 的平均值为：

$$M_{ij} = \frac{1}{S * C} \sum_{c=1}^C \sum_{s=1}^S X_{ij}(s,c). \quad (8)$$

进而计算所有训练图片 DCT 系数的方差：

$$V_{ij}^B = \sum_{C=1}^C \sum_{S=1}^S (X_{ij}(s,c) - M_{ij})^2. \quad (9)$$

最后得到对应点的分类能力 $J(i, j)$:

$$J(i, j) = \frac{V_{ij}^B}{V_{ij}^W}. \quad (10)$$

J 值越大意味着越好的分类能力。显然要使得 J 值越大, 要求 S_{ij}^B 越大且 S_{ij}^W 越小。

在 ORL 人脸数据库上采用本方法进行 DCT 系数的选择, 得到的如图 5 所示结果, 很明显, J 值越大的区域主要集中在低频区和高频区。



图 5. ORL 库的 DCT 系数 J 阵图像

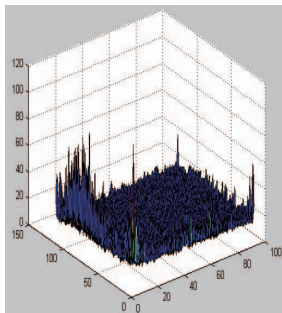


图 6. ORL 人脸库的 DCT 系数 J 值分布图

图 6 则显示了其 J 值分布情况, 与 DCT 三个区域具有对应关系。由图 6 说明 J 值高频区拥有较大的 J 值, 这是由于高频区反映的噪音的影响。为了回避光照影响并且降低噪音干扰, 我们只选取数值大小在中间部分的值重构特征向量。首先我们将所得到的 J 阵转换成列向量, 其分量按从大到小排序, 然后选择列向量中间段 $[m, n]$ 之间的数据进行重构得到新的向量 j' , 并且记录其对应于 J 矩阵的位置。对于测试样本集合, 在计算得到每个图片的 DCT 系数矩阵后选取对应位置的值形成特征向量。然后运用

PCA 方法形成新的特征空间, 去除数据的相关性并完成降维, 最后采用最近邻距离法实现分类。

IV. 结果及分析

本次实验在 ORL 人脸数据库上进行测试, 该数据库中有四十个人, 每个人十张不同的图片每幅图片的分辨率为 112×92 , 这些图片的差异主要表现在不同的拍摄时间、不同的拍摄光强以及不同的表情 (睁/闭眼睛、微笑/不微笑) 以及面部细节 (戴眼镜/不戴眼镜)。所有图像均采用黑色均匀背景、直立正面位置[4]。人脸姿态也有一定程度的变化, 如深度旋转和平面旋转可达 20° , 人脸的尺度也有多达 10% 的变化[5]。

每人分别选择 5 张训练图片, 5 张测试图片, 这样训练样本和测试样本分别都为 200 张。第一次运用 DCT 获取图像的 DCT 矩阵然后结合本文提出的方法提取人脸特征, 并且为了显示对比尽量选取相同的系数数量 (都为 75 个)。最后, 采用三阶最近邻距离进行分类识别。

表 I 是我们选择 J 阵转换后的向量在 $[12, 87]$ 区间上形成的向量构建的特征向量进行分类后所示本方法与传统的 $m \times m$ 矩阵法和 Zig-Zag 法在时间性能上的对比。

表格 I. ORL 上对时间效率实验结果

	所用方法		
	DCT + $m \times m$ 矩阵	DCT + Zig-Zag	DCT + 本文方法
时间	0.531000	0.610000	1.060000

时间单位是秒 (s)

该方法在选取不同 J 值域的情况下的识别率的对比如表 II 所示, 结果表明对系数的选择应尽量逼近中频区域。

表格 II. 选取不同 J 向量取值区域大小的实验结果

	J 向量取值区域大小					
	$[1, 10000]$	$[5, 81]$	$[10, 84]$	$[15, 86]$	$[20, 87]$	$[30, 91]$
精度	75.56	85.00	84.50	92.22	94.50	87.50

精度单位是百分数 (%)

表 III 则给出了本文方法与其他人脸识别方法识别效率的对比, 结果显示该方法在识别率上有明显提高。

表格 III. 不同方法在 ORL 人脸库上识别结果对比

	识别方法				
	PCA	2DPCA	DCT+PCA	DCT+Zig-Zag+PCA	本文方法
精度	82.22	87.78	84.44	85.56	94.50

精度单位是百分数 (%)

V. 结论

DCT 变换能有效的降低特征空间的维数, 同时具有对光照, 姿势, 表情, 姿态不敏感的特性。DCT 系数的选择

对于分类至关重要，但传统的方法如 $m \times m$ 矩阵法和 Zig-Zag 法均缺乏严谨性。基于每个 DCT 系数分类能力不同的认识，本文结合 DCT 和统计分析的方法计算出每个 DCT 系数的分类能力，然后选择更有效的系数进行分类。结合 PCA 方法，能够去除数据的相关性。实验表明它的识别率明显优越于传统的方法，并且对光照，姿态等影响也更具鲁棒。但同时由于本方法基于统计原理故而实验结果受样本数目的影响，一般来说样本数目越大误差越小。但是对于系数区间的选择仍然需要凭靠经验，这也是本文需要继续改进的地方。

REFERENCES

- [1] Kirby M, Sirovich L. Application of the KL procedure for the characterization of human faces[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1990, 12 (1) : 103-108.
- [2] 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.
- [3] 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.
- [4] Ferdinando Samaria, Andy Harter. Parameterisation of a stochastic model for human face identification. Proceedings of 2nd IEEE Workshop on Applications of Computer Vision, 1994, 12.
- [5] XIE Hong/a, NING Zhi-gang, ZHANG Lei. Face recognition system based on DCT and modular 2D-PCA, Applied Science and Technology, 2009, 6.