

一种有效的人脸识别方法

程永清 庄永明 汪华峰 杨静宇

(华东工学院计算机系, 南京 210014)

摘 要

人脸识别是模式识别领域中一个相当困难而又有重要理论价值的研究课题。本文首先引入表示人脸的特征图像和投影图像的概念, 提出利用投影图像在特征图像上投影的坐标作为描述人像本质属性的特征矢量, 并导出了基于投影图像的人脸特征抽取方法。最后构造了一个层次的距离分类器进行人脸的识别。实验结果表明基于投影图像的特征矢量具有很好的稳定性、鉴别能力和识别率。

关键词: 模式识别, 人脸识别, 特征抽取, 特征图像, 投影图像。

一、引 言

人脸自动识别的研究始于六十年代末。Bledsoe^[1] 首先提出人脸的半自动识别系统, 它用于分类的参数包括一些人脸脸部特征规一化的点间距离和比率, 如眼角, 嘴角, 鼻尖等部位。Kanade 的人脸鉴别系统^[2] 从单幅图像上计算一组脸部参数, 利用模式分类技术将此人脸与一组标准的人脸匹配。Bell 实验室早期的工作^[3-4] 集中于提供充分的人脸描述, 利用多达 21 个特征的特征矢量描述人脸, 然后利用标准的模式分类技术识别人脸。在识别方法上, Wong 等人^[5] 采用了模板匹配技术提取人脸的几何特征, 完成人脸的识别。Kaufman 等人^[6] 利用计算人脸侧面图像的自相关系数抽取 12 个特征, 然后根据最小距离准则进行识别。Craw 等人^[7] 又从人脸侧面上的 9 个特征按加权距离准则进行识别。

人脸的几何特征不足以描述人脸, 这类方法忽视了这样一个事实: 识别的对象是有感情的人脸, 而不是刚体。作者的基本观点是人脸图像的代数特征才是反映人脸的内在本质特征。

二、人脸的特征图像和投影图像

Golub 和 Reinch 于 1970 年提出计算矩阵的奇异值分解 (SVD) 算法。之后, SVD 算法已成为解决有关矩阵计算问题的一种有效工具, 并用于图像的压缩和存贮。SVD 算法是本文提出的矩阵秩分解 (RDM)^[8] 的一种特殊形式。若 $A \in R^{m \times n}$ 为任一幅人脸灰度图像, 则 A 有下列 SVD 分解, 即

$$A = U\Sigma V = \sum_{i=1}^l \lambda_i u_i v_i^T, \quad (1)$$

其中 $U = (u_1, u_2, \dots, u_l, u_{l+1}, \dots, u_m)$, $V = (v_1, v_2, \dots, v_l, v_{l+1}, \dots, v_m)$, $\Sigma = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_l, 0, \dots, 0)$. 这里, $l = \text{rank}(A)$, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l > 0$, $\lambda_i^2 (i = 1, 2, \dots, l)$ 是 AA^T 和 $A^T A$ 的特征值, λ_i 是矩阵 A 的奇异值, $u_i, v_i (i = 1, 2, \dots, m; j = 1, \dots, n)$ 分别是 AA^T 和 $A^T A$ 对应于特征值 λ_i^2 的特征矢量.

对于人脸图像 A , 它的特征矩阵 $u_i v_i^T (i = 1, 2, \dots, l)$ 称为图像 A 的特征图像. 为了便于下面的讨论, 首先给出矩阵的 Frobenius 范数.

定义 1. 对于 $A = (a_{ij}) \in R^{m \times n}$, 令

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}, \quad (2)$$

则称 $\|A\|_F$ 为矩阵 A 的 Frobenius 范数.

若 A, B 表示采样后得到的两个人脸灰度图像, 那么 Frobenius 范数 $\|A\|_F, \|B\|_F$ 恰好分别表示人脸图像 A, B 的总能量, 而 $\|A - B\|_F$ 表示人脸图像之间差异程度的一种度量.

对于人脸图像 A , 它有(1)式的 SVD 分解. 一般情况下, $\lambda_i (i = 1, 2, \dots, l)$ 接近零的项较多, 如果去掉这些项, 就能得到 A 的近似图像 \tilde{A} :

$$\tilde{A} = \sum_{i=1}^r \lambda_i u_i v_i^T, \quad (3)$$

这里 $r < l$, 矩阵 \tilde{A} 是 A 的最小二乘逼近. 这时损失的能量为

$$\|A - \tilde{A}\|_F = \sqrt{\sum_{i=r+1}^l \lambda_i^2}. \quad (4)$$

因此, 图像 \tilde{A} 对于 A 来说, 其损失的能量最小. 事实上, 近似图像 \tilde{A} 可以看作是图像 A 在它的特征图像 $u_1 v_1^T, u_2 v_2^T, \dots, u_r v_r^T$ 上的投影的叠加.

定义 2. 若 $A, B \in R^{m \times n}$, 且 A 有(1)式的 SVD 分解, 令

$$\bar{B} = U\bar{\Sigma}V^T = \sum_{i=1}^l \bar{\lambda}_i u_i v_i^T,$$

其中 $\bar{\Sigma} = \text{diag}(\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_l, 0, \dots, 0)$, $\bar{\lambda}_i = u_i^T B v_i (i = 1, 2, \dots, l, l = \text{rank}(A))$, 则称 \bar{B} 为图像 B 在图像 A 上的投影图像.

如果用 B 在 A 上的投影图像 \bar{B} 来描述 B , 则损失的能量为 $\|B - \bar{B}\|_F$. 由 Frobenius 范数的性质, $\|B - \bar{B}\|_F$ 满足下列不等式:

$$0 \leq \|B - \bar{B}\|_F = \|B - A + A - \bar{B}\|_F \leq \|A - B\|_F + \|A - \bar{B}\|_F, \quad (5)$$

即

$$0 \leq \|B - \bar{B}\|_F \leq \|A - B\|_F + \sqrt{\sum_{i=1}^l (\lambda_i - \bar{\lambda}_i)^2}. \quad (5')$$

由(5)和(5')式可知, 图像 B 和它的投影图像 \bar{B} 的差异与图像 A, B 的差异密切相关. 从理论上进一步得到下列结论:

定理. 若 $A, B \in R^{m \times n}$, 且 $B - A = \Delta$, $\|\Delta\|_F = \delta$ (δ 为实数), 如果 \bar{B} 是在图像 A

上的投影图像,则 1) $\|A - \bar{B}\|_F \leq \delta$; 2) $\|B - \bar{B}\|_F \leq \delta$.

上述定理表明: 只要 A, B 有较好的稳定性, 则 A 与 \bar{B}, B 与 \bar{B} 均有较好的稳定性. 对于同类的人脸图像, 图像样本通常具有较好的稳定性, 即若 $\{A_i\}_{i=1}^k$ 表示同类人脸的图像样本, 则存在常数 ε , 使下式成立:

$$\|A_i - A_j\|_F \leq \varepsilon, \quad (i, j = 1, 2, \dots, k), \quad (6)$$

这时, 若将人脸图像样本 $A_i (i = 1, 2, \dots, k)$ 投影到这类的基准图像 A 上得到投影图像 \bar{A}_i , 则有

$$\begin{aligned} \|\bar{A}_i - \bar{A}_j\|_F &= \|u \Sigma_i v^T - u \Sigma_j v^T\|_F = \|\Sigma_i - \Sigma_j\|_F \\ &\leq \|\text{diag}(u_1^T(A_i - A_j)v_1, u_2^T(A_i - A_j)v_2, \dots, u_n^T(A_i - A_j)v_n)\|_F \\ &\leq \|A_i - A_j\|_F \leq \varepsilon. \end{aligned} \quad (7)$$

式(7)清楚地表明: 同类人脸图像经过投影后, 仍具有很好的稳定性. 对于同类的多个人脸图像样本, 若选择图像样本的其中之一作为基准图像, 则图像样本中有些图像与它的投影图像差别较大. 而选择“平均图像”作为基准图像, 则可以在一定程度上限制这种情况.

若 A 表示某类的平均人脸图像, 且有(1)式的 SVD 分解, B_1, B_2, \dots, B_m 表示这类的所有人脸图像, 则 B_1, B_2, \dots, B_m 在 A 上的投影图像 $\bar{B}_1, \bar{B}_2, \dots, \bar{B}_m$ 具有相同的特征图像, 即

$$\begin{cases} \bar{B}_i = \sum_{j=1}^l \bar{\lambda}_j^{(i)} u_j v_j^T, \\ \bar{\lambda}_j = u_j^T B_i v_j. \end{cases} \quad (8)$$

假设 $y \in \{B_1, B_2, \dots, B_m\}$, 则

$$\begin{cases} \bar{y} = \sum_{j=1}^l \bar{\lambda}_j u_j v_j^T, \\ \lambda_j = u_j^T y v_j. \end{cases} \quad (9)$$

令 $X = (\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_l)$, $l = \text{rank}(A)$, 则称 X 为描述人脸图像 Y 的坐标特征矢量. X 由投影坐标 $\bar{\lambda}_k (k = 1, 2, \dots, l)$ 组成, 它描述人脸图像 Y 特有的内在本质特征, 它包含更多的鉴别信息, 并且有很好的稳定性.

三、人脸图像的特征抽取

通过上面的讨论, 可以给出人脸图像的坐标特征矢量的特征抽取方法. 则基于投影图像的特征抽取过程如下:

$$1) \text{ 求平均图像 } A, A = \frac{1}{k} \sum_{i=1}^k A_i;$$

2) 将 A 进行 SVD 分解, 令

$$A = U \Sigma V^T = \sum_{i=1}^l \lambda_i u_i v_i^T,$$

其中

$$\begin{aligned}\Sigma &= \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_l, 0, \dots, 0), \\ l &= \text{rank}(A), \\ U &= (u_1, u_2, \dots, u_m), \\ V &= (v_1, v_2, \dots, v_n),\end{aligned}$$

则求得一组特征图像 $u_i v_i^T (i = 1, 2, \dots, l)$.

3) 求人脸图像 A_i 在 A 上的投影图像 \bar{A}_i , ($i = 1, 2, \dots, k$), 即

$$\bar{A}_i = \sum_{j=1}^l \bar{\lambda}_j^{(i)} u_j v_j^T, \quad \bar{\lambda}_j^{(i)} = u_j^T A_i v_j, \quad (j = 1, 2, \dots, l);$$

4) 求人脸图像 A_i 的坐标特征矢量 $X_i (i = 1, 2, \dots, k)$, $X_i = (\bar{\lambda}_1^{(i)}, \bar{\lambda}_2^{(i)}, \dots, \bar{\lambda}_l^{(i)})$.

四、人脸图像的识别

对于同类人脸图像, 它们的投影图像具有相同的特征图像 $u_i v_i^T (i = 1, 2, \dots, l)$. 上述定理从理论上保证特征抽取后的坐标特征矢量具有很好的稳定性. 从本质上讲, 经校正后的同类人脸图像的平均图像的特征图像反映了同类人脸图像的共同信息, 而抽取的坐标特征矢量反映了同类人脸图像的各自图像的差异信息.

设 $Q_i = \{A_j^{(i)}\}_{j=1}^{N_i}$ 为第 i 类人脸图像的训练样本集 ($i = 1, 2, \dots, M$), 则人脸图像的识别算法如下:

1) 求第 i 类人脸图像的平均图像 $A^{(i)}$, 即

$$A^{(i)} = \frac{1}{N_i} \sum_{j=1}^{N_i} A_j^{(i)};$$

2) 将 $A^{(i)}$ 进行 SVD 分解, 得到第 i 类的一组特征图像 $\{u_j^{(i)} v_j^{(i)T}\}_{j=1}^l$ ($i = 1, 2, \dots, M$, $l = \min(n, m)$);

3) 求 $A_j^{(i)}$ 的坐标特征矢量 $X_j^{(i)} (j = 1, 2, \dots, N_i, i = 1, 2, \dots, M)$, 令 $\bar{\lambda}_{jk}^{(i)} = (u_k^{(i)})^T A_j^{(i)} v_k^{(i)}$, ($k = 1, 2, \dots, l$), 则 $X_j^{(i)} = (\bar{\lambda}_{j1}^{(i)}, \bar{\lambda}_{j2}^{(i)}, \dots, \bar{\lambda}_{jl}^{(i)})$;

4) 求第 i 类人脸图像的特征矢量 $\{X_j^{(i)}\}_{j=1}^{N_i}$ 的中心 $X^{(i)} (i = 1, 2, \dots, M)$, 即

$$X^{(i)} = \frac{1}{N_i} \sum_{j=1}^{N_i} X_j^{(i)};$$

5) 进行层次判别, 即若 Y 为任一待识人脸图像, 则首先求 Y 在每类平均图像 $A^{(i)}$ 上的投影图像 $Y^{(i)} (i = 1, 2, \dots, M)$, 从而得到 M 个坐标特征矢量 $\bar{Y}^i (i = 1, 2, \dots, M)$. 若 $\bar{\lambda}_k^{(i)} = (u_k^{(i)})^T y v_k^{(i)}$, $k = 1, 2, \dots, l$, 则 $\bar{Y} = (\bar{\lambda}_1^{(i)}, \bar{\lambda}_2^{(i)}, \dots, \bar{\lambda}_l^{(i)})$. 最后建立如下的层次判别准则:

若

$$\min \left\{ \sqrt{\sum_{j=1}^l (\bar{Y}^{(i)} - X_j^{(i)})^2} \right\} = \sqrt{\sum_{j=1}^l (Y_j^{(r)} - X_j^{(r)})^2}, \quad r \in \{1, 2, \dots, M\},$$

则

$$Y \in Q_r.$$

五、实验结果和结论

为了进一步验证上述特征抽取和识别方法的有效性, 本文完成了 4 人 32 幅人脸图像

的人脸识别实验(校正后的人脸图像)。本实验的图像输入装置是小型图像处理系统 PC-VISION。通过 PC-VISION 摄取 4 人的人脸图像,得到了一个灰度范围在 0—255 的 64×64 阶实矩阵。对于每个人脸图像,用 64×64 的窗口摄取人脸的部位。在摄像机

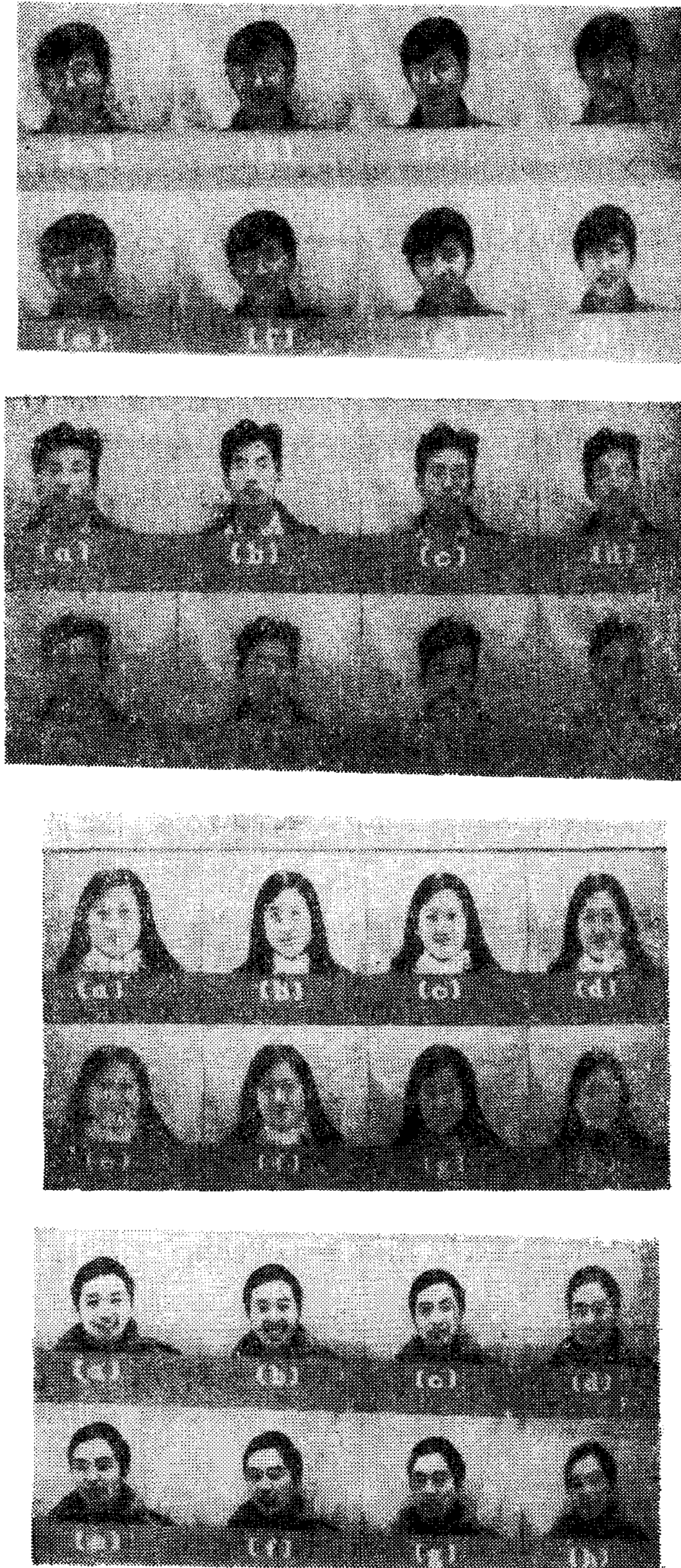


图 1 实验用的 32 张人脸图像

与人脸的相对位置略有变动,光圈和焦距,脸部表情,戴眼镜以及模糊程度有较大变动的条件下,每人重复采样 8 次,得到 32 个人脸图像,见图 1.

1. 训练人脸图像的识别

表 1 训练样本识别结果

类别	样本	分类器输出值				结果
		d_1	d_2	d_3	d_4	
No.1	<i>a</i>	0.76	201.22	2018.47	5018.71	1
	<i>b</i>	6.03	201.85	2218.78	5246.48	1
	<i>c</i>	9.38	195.98	1687.53	4551.09	1
No.2	<i>a</i>	356.40	10.40	2192.86	4848.58	2
	<i>b</i>	356.62	3.95	1882.45	4524.91	2
	<i>c</i>	400.26	8.73	1666.72	4204.79	2
No.3	<i>a</i>	2239.90	1928.07	19.47	1390.78	3
	<i>b</i>	2146.43	1865.62	26.54	2359.25	3
	<i>c</i>	1834.75	1518.52	15.82	1146.64	3
No.4	<i>a</i>	4599.43	4209.14	981.30	9.26	4
	<i>b</i>	4966.80	4445.78	1360.45	2.56	4
	<i>c</i>	4733.07	4321.16	1088.69	12.51	4

表 2 测试样本识别结果

类别	样本	分类器输出值				结果
		d_1	d_2	d_3	d_4	
No.1	<i>d</i>	5.74	216.93	2143.42	5270.92	1
	<i>e</i>	6.30	211.20	2139.27	5284.68	1
	<i>f</i>	13.31	216.57	1862.68	5028.79	1
	<i>g</i>	2.17	220.33	1926.13	5053.10	1
	<i>h</i>	119.66	262.33	1600.11	4708.75	1
No.2	<i>d</i>	400.51	21.05	1645.33	3922.17	2
	<i>e</i>	371.13	6.13	1708.54	4193.86	2
	<i>f</i>	373.03	24.71	1632.94	4091.93	2
	<i>g</i>	390.55	19.06	1638.17	4001.62	2
	<i>h</i>	163.72	13.71	1906.05	4502.83	2
No.3	<i>d</i>	1947.11	1729.99	4.96	1568.31	3
	<i>e</i>	1910.89	1624.15	17.33	1093.68	3
	<i>f</i>	1916.84	1555.06	25.15	1012.04	3
	<i>g</i>	1505.67	1193.19	121.18	1175.31	3
	<i>h</i>	1499.50	1360.71	34.96	1632.96	3
No.4	<i>d</i>	4555.39	4189.89	943.97	189.97	4
	<i>e</i>	4235.97	3664.23	1029.96	57.97	4
	<i>f</i>	4916.19	4372.90	1524.12	182.27	4
	<i>g</i>	4540.34	3969.91	1239.00	77.00	4
	<i>h</i>	3434.08	3109.68	619.69	194.12	4

取图 1 中每类人脸的(a), (b), (c)三个图像作为训练图像,求得 4 人平均图像的特征图像,然后对每类训练图像进行特征抽取,得到相应的坐标特征矢量,最后得到 4 人的坐标特征矢量中心 $X^{(i)} (i = 1, 2, \dots, 4)$ 。实验结果表明,坐标特征矢量的维数只需取 6,即 $t = 6$,就可以完成实验。4 人的训练图像识别结果如表 1。

2. 测试人脸图像的识别

取图 1 中每类人脸的(d)–(h)五个人脸图像作为测试图像,总试 5×4 个测试图像。其中每类人脸的(d), (e), (f)图像是在摄像机与人脸的相对位置、光圈和焦距、光照条件、脸部表情有变动的条件下得到的。每类人脸的(g), (h)图像是在摘取眼镜和较大模糊的情况下得到的。测试图像的识别结果如表 2 所示。

上述实验表明:由于坐标特征矢量的稳定性,使它具有对人脸微小的姿态变化,光圈和焦距,脸部表情,戴眼镜,模糊程度,图像噪声,光照条件等的不敏感性。坐标特征矢量的稳定性也使得对于同类的人脸图像有很好的聚类性。同类人脸图像的高聚类性和不同类人脸图像的高可分性保证了这种方法具有很高的识别率。值得提出的是若将抽取的坐标特征矢量单位化,则坐标特征矢量对于比例变化也不敏感。

由于坐标特征矢量的理论基础是同类图像在 Frobenius 范数意义下应具有相同的特征图像,因此基于坐标特征矢量的方法是一种识别人脸的好方法,但不是描述人脸的好方法,因而它不具有对图像的平移和旋转的不敏感性。

参 考 文 献

- [1] Bledsoe, W. W., Man-machine Facial Recognition, Panaramic Research Inc., Palo Alto. CA. Rep. PRI 22, Aug. (1966).
- [2] Kanade, T., Picture Processing System by Computer and Recognition of Human Face, Dept. of Information Science, Kyoto University, Nov. (1973).
- [3] Goldstein, Harmon and Lesk, Identification of Human Face, Proc. of IEEE, 59(1971), 748—754.
- [4] Harmon, L. D. and Hunt, W. F., Automatic Recognition of Human Face Profiles, *Computer Graphics and Image Processing*, 6(1977), 135—156.
- [5] Wong, K. H. et al., A Human Face Recognition System, Electronic Imaging'88-International Electronic Imaging Exposition, (1988).
- [6] Kaufman, G. J. and Breeding, K. J., The Automatic of Human Faces from Profile Silhouettes, *IEEE Trans. SMC*, SMC-6(1976), No. 2.
- [7] Craw, Ellis and Lishman, *Automatic Extraction of Face Feature*, *Pattern Recognition Letter*, 5(1987), 183—187.
- [8] Cheng, Y. Q., Zhuang, Y. M. and Yang, J. Y., Optimal Fisher Discriminant Analysis Using the Rank Decomposition, *Pattern Recognition* 25 (1992), 101—111.

AN EFFICIENT RECOGNITION APPROACH TO HUMAN FACES

CHENG YONGQING ZHUANG YONGMING WANGHUAFENG YANG JINGYU

(Dept. of Computer Science, East China Institute of Technology, Nanjing 210014)

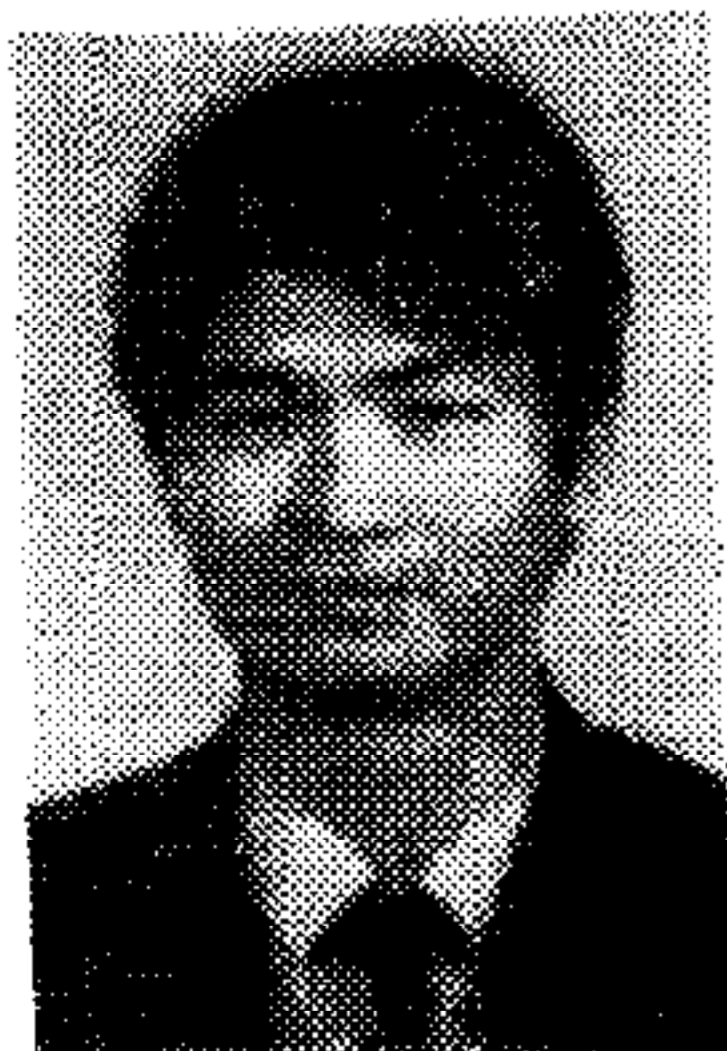
ABSTRACT

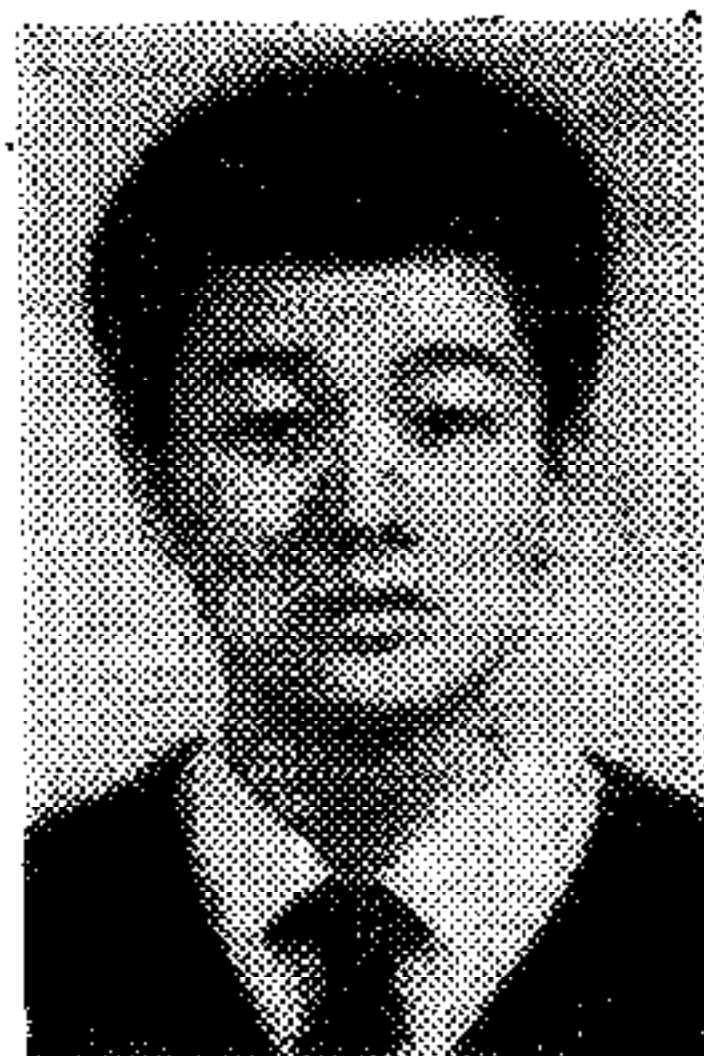
Human face recognition is a very difficult and important problem in the field of pattern recognition. In this paper, the characteristic images and projective images are first introduced. The projective coordinates of projective image on the characteristic images are presented to describe the intrinsic attributes of human faces. Then, a feature extraction method of human faces based on the projective images is proposed. Finally, a hierarchical distance classifier is designed to recognize human faces. Experimental results show that the feature vector based on the projective images has good stability, discriminant ability and recognition rate.

Key words: Pattern recognition; human face recognition; feature extraction; characteristic image; projective image.

程永清 江苏无锡人,生于1963年。曾就读于华东工学院计算机科学与工程系,分别在1984和1987年获得学士与硕士学位。现任该系讲师,已发表论文20余篇。主要研究方向是计算机视觉、模式识别、专家系统。

庄永明 江苏宜兴人,生于1965年。曾就读于徐州师范学院数学系、华东工学院计算机科学与工程系。分别于1987年、1990年获得学士及硕士学位。目前研究领域为模式识别及计算机视觉。





汪华峰 北京市人,生于1969年。1991年在华东工学院计算机科学与工程系获得学士学位。现在该系攻读硕士研究生。感兴趣的研究领域是:自然语言理解、模式识别及计算机视觉。



杨静宇 1941年生于长春市,1965年毕业于中国人民解放军炮兵工程学院仪器系。曾在美国伊利诺斯大学从事模式识别和机器视觉研究。现任华东工学院计算机科学与工程系主任、教授。先后主持和参加十多项国家和部委重点科研项目,多次获得各种奖励。已出版专著和译著三部,发表论文50余篇。目前主要研究领域是机器视觉、模式识别、智能机器人。