

SVD用于人脸识别存在的问题及解决方法

高全学¹⁾²⁾ 梁彦²⁾ 潘泉²⁾ 陈玉春²⁾ 张洪才²⁾

¹⁾(西安电子科技大学通信工程学院,西安 710071) ²⁾(西北工业大学自动化学院,西安 710072)

摘要 通过对人脸图像奇异值的分析,证实了图像奇异值是图像在特定基空间分解得到的,这个基空间是由图像本身决定的。进一步研究发现,导致基于奇异值向量人脸识别算法识别率低的根本原因是:不同人脸图像对应的奇异值向量所在的基空间不一致,奇异值向量与人脸图像之间并不存在一一对应关系,奇异值向量具有不可分割性。最后提出了类估计基空间识别算法。在 ORL、ORL-NWPU1 以及 ORL-NWPU2 数据库进行仿真,实验结果证实了分析和所提算法的正确性。

关键词 奇异值向量 人脸识别

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2006)12-1784-08

The Problem Existed in Face Recognition Using SVD and Its Solution

GAO Quan-xue¹⁾²⁾, LIANG Yan²⁾, PAN Quan²⁾, CHEN Yu-chun²⁾, ZHANG Hong-cai²⁾

¹⁾ School of Telecommunication Engineering, Xi Dian University, Xi'an 710071

²⁾ College of Automation, Northwestern Polytechnical University, Xi'an 710072

Abstract The singular value vector of face image is analyzed and its nature is revealed, i. e., the singular values of face image are obtained by decomposing the face in the special basis space that is determined by the face image. Based on this result, the intrinsic reasons that singular value vector-based recognition methods have the low recognition accuracy are follows: singular value vectors of arbitrary two face images have not the equivalent basis space in general; there is no one-to-one correspondence between singular value vector and face image; singular value vectors are not separated. Finally, a novel face recognition method coined class estimated basis space method is presented. ORL, ORL-NWPU1 and ORL-NWPU2 database are used to test, and the experimental results show the correctness of analysis and the proposed method.

Keywords singular value vector, face recognition

1 引言

SVD作为一种有效的代数特征提取方法,因其很好的属性已在数据压缩、信号处理和模式分析等许多领域中得到了广泛的应用^[1-5]。奇异值分解首次应用于人脸识别是由洪子泉^[1]提出的,与经典特征提取技术主成分分析(PCA)和Fisher判别分析(FLD)相结合,在ORL数据库上获得了不错的识别率。在此基础上,人们提出了许多基于奇异值分解

的人脸识别方法^[6-12]。这些方法首先利用奇异值分解提取图像奇异值向量,然后利用奇异值向量识别人脸。大量事实表明,基于奇异值向量人脸识别算法的识别率比较低,对此,目前有两种解释:一种认为是小样本对统计造成的影响^[1],另一种认为奇异值向量没有包含足够多的有利于识别的信息^[8]。但都没有从理论上给出解释,况且奇异值作为图像的一种代数特征,具有很好的稳定性以及位移、旋转等不变性性质^[13]。如果说这样的特征没有包含足够的信息或者说小样本问题造成奇异值向量人脸识

基金项目:国家自然科学基金项目(60372085,60172037),航空科学基金项目(03D53032)

收稿日期:2004-09-01,改回日期:2006-02-27

第一作者简介:高全学(1975~),男,2005年获西北工业大学控制理论与控制工程专业博士学位。主要研究方向为模式识别、计算机视觉及图像处理。E-mail: gaoqxnp@163.com

别算法识别率低,显然缺乏说服力。

通过理论分析发现,人脸图像奇异值其实是图像在特定基空间分解得到的,这个基空间是由人脸图像本身决定的。一般情况下,不同人脸图像对应的这个特定基空间是不相同的。进一步研究发现,基于奇异值向量人脸识别算法识别率低的原因是:不同人脸图像对应的奇异值向量所在的基空间不一致,奇异值向量与人脸图像之间并不存在一一对应关系,奇异值向量具有不可分割性。最后根据同类图像相似性比较大而非同类图像相似性比较小的假设,提出了类估计基空间人脸识别方法。并在 ORL、ORL-NWPU1 以及 ORL-NWPU2 数据库上进行了实验和分析。

2 奇异值向量识别人脸的分析

2.1 人脸图像奇异值向量的分析

定理^[14]1 设 A 为任意一个实矩阵,即 $A \in \mathbf{R}_k^{m \times n}$,其中 \mathbf{R} 表示实数域, $k = \text{rank}(A)$ 为矩阵的秩。则存在 m 阶正交阵 U 和 n 阶正交阵 V ,使得

$$A = UDV^T \quad (1)$$

其中, $D_{m \times n} = \begin{pmatrix} \sum_{k \times k} 0 & \\ & 0 \end{pmatrix}$, $\sum_{k \times k} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k)$, $\sigma_i (i = 1, 2, \dots, k)$ 称为矩阵 A 的奇异值,且 $\forall i > j, i = 1, 2, \dots, k, j = 1, 2, \dots, k$, 有 $\sigma_1 \leq \sigma_j$ 。

如果用 A 代表任意一幅人脸图像,则它的奇异值向量定义为 $x = (\sigma_1, \dots, \sigma_k)^T$

式(1)展开可写成:

$$\begin{aligned} A &= \sigma_1 u_1 v_1^T + \dots + \sigma_k u_k v_k^T \\ &= \alpha_{11} u_1 v_1^T + \dots + \alpha_{1n} u_1 v_n^T + \alpha_{21} u_2 v_1^T + \alpha_{22} u_2 v_2^T + \dots + \\ &\quad \alpha_{2n} u_2 v_n^T + \dots + \alpha_{m1} u_m v_1^T + \dots + \alpha_{mn} u_m v_n^T \end{aligned} \quad (2)$$

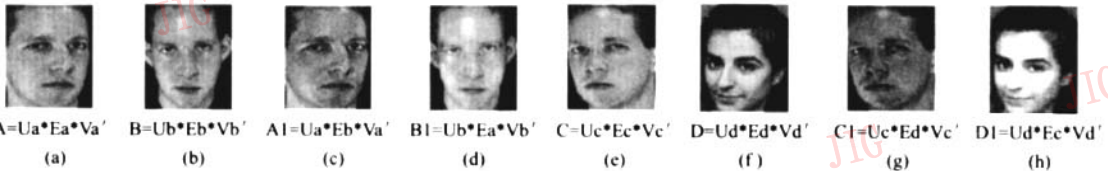


图 1 原图与重构后的图像

Fig. 1 Original images and reconstructed images

根据上述分析及奇异值分解定理,显然有以下两个结论:

结论 2 人脸图像与奇异值向量之间并不存一一对应关系。

结论 3 人脸图像的奇异值向量体现不出不同

$$\text{其中 } \alpha_{ij} = \begin{cases} \sigma_i & i=j=1, 2, \dots, k \\ 0 & \text{其他} \end{cases}$$

根据式(2)容易得出以下的定理 2。

定理 2 对任意一个人脸图像 $A \in \mathbf{R}^{m \times n}$, 设 $U \in \mathbf{R}^{m \times m}, V \in \mathbf{R}^{n \times n}$ 分别是图像 A 奇异值分解时对应的左右正交阵, 则矩阵 $u_1 v_1^T, \dots, u_m v_n^T$ 是矩阵空间 $\mathbf{R}^{m \times n}$ 中的一组最大线性无关组矩阵, 即 $\mathbf{R}^{m \times n}$ 中的一组基。其中 $u_i, v_j (i = 1, \dots, m; j = 1, \dots, n)$ 分别是矩阵 U 和 V 的第 i 列和第 j 列(证明见附录 A)。

根据奇异值分解定理及定理 2 可得结论 1。

结论 1 人脸图像的奇异值向量所在的基空间(矩阵)是由人脸图像本身决定的。

从定理 2 知,式(2)表明人脸图像 $A \in \mathbf{R}^{m \times n}$ 可以准确地通过基矩阵 $u_1 v_1^T, u_1 v_2^T, \dots, u_m v_n^T$ 的线性组合来表示,其中 $(\alpha_{11}, \alpha_{12}, \dots, \alpha_{mn})$ 是人脸图像 A 在这个基空间(坐标系)下的坐标。根据人脸图像奇异值的定义知,人脸图像 A 的奇异值 $\sigma_i (i = 1, \dots, k)$ 正是图像 A 在这个坐标系下对应于基矩阵 $u_i v_i^T (i = 1, \dots, n)$ 的坐标分量。由奇异值分解定理知,基空间 $u_1 v_1^T, u_1 v_2^T, \dots, u_m v_n^T$ 是由给定人脸图像本身决定的。因此,人脸图像的奇异值向量所在的基空间(矩阵)是由图像本身决定的。

此外,在 ORL 图库上通过大量实验发现:无论交换非同类图像还是同类图像相互间的奇异值向量,重构后的图像与原图像类别一致,相反,如果交换非同类图像间的左右正交阵,而奇异值不变,重构后的图像与原图像的类别不再一致,如图 1 所示。这表明,在不同或相同类别中,一定存在具有相同奇异值向量的人脸图像,即对给定的奇异值向量,它的类别具有不确定性。

类别图像之间的差异,具有不可分割性。

2.2 算法分析

利用图像奇异值向量识别模式的方法比较多,大致可分为两类:直接和组合。直接是指把提取的奇异值向量直接作为分类器的输入,然后度量每个

向量之间的距离,最后根据最近邻准则进行分类。组合是指首先把奇异值向量通过其他算法分析(如主成分分析或 Fisher 判别分析),得到新的特征,然后作为分类器的输入进行分类。

2.2.1 直接算法(SV 算法)

假定 $sv_i (i=1, \dots, N)$ 代表训练集中第 i 个图像的奇异值向量, N 代表总的训练图像数目。 tsv 代表测试图像的奇异值向量。则测试图像的奇异值向量 tsv 和训练集中第 i 个图像的奇异值向量之间的距离为

$$d(tsv, sv_i) = \|tsv - sv_i\|_2$$

如果 $d(tsv, sv_j) = \min_i d(tsv, sv_i)$, 且 sv_j 属于第 k ($k=1, \dots, C$) 类, 则认为测试图像属于第 k 类, 其中 C 为训练集类别数目。

大量实验发现,这种算法的识别率明显低于其他人脸识别算法(如 Eigenface、Fisher 人脸算法)的识别率。文献 [1] 的解释是小样本对统计造成的影响。但在训练样本多时,发现该类算法的识别率仍然比较低。因此这个解释不合理。

根据结论 1,并考虑到人脸图像获取受光照、视角、面部装饰、姿势以及人脸的非刚性变化等影响,对于任何两个不同人脸图像(它们既可以是属于同一类别,可以属于不同类别)的奇异值,它们所对应的基空间(坐标系)一般是不相同的。因此,一般情况下,对任何两个不同人脸图像的奇异值向量,它们所对应的基空间是不相同的。由 SV 算法识别过程可知,该类算法正好忽略了不同图像奇异值向量所在基空间不一致问题。换句话说,该类算法默认了所有人脸图像的奇异值向量所在的基空间是相同的,与实际不符。另外,直接度量不在同一坐标系下的任意两个向量之间的距离是没有任何意义的。因此,这类人脸识别算法存有本质上的缺陷。至于文献 [7] 提出的基于重构误差和 SV 算法相融合的人脸识别算法,这个思想本身是好的,因为同一特征不同算法相融合也是提高识别率的有效途径之一。但该算法同样存在基空间不一致问题。

文献 [7]、[8] 从信息量角度对这类算法进行了分析,认为引起低识别率的根本原因是人脸图像奇异值向量没有包含足够多的识别信息,但没有从理论上进行说明,也没有给出衡量一个特征包含信息量大小的标准。根据定理 2 可知,虽然人脸图像 $A \in \mathbf{R}^{m \times n}$ 在基空间下的坐标分量共有 $m \times n$ 个,而它的非零奇异值只有 k 个,但其余 $m \times n - k$ 个分量的

值都为零。而且对于给定的一幅人脸图像,它的奇异值向量是唯一的。由式(2)知,利用这 k 个奇异值完全可以准确无误地重构人脸图像。另外,如果把图像的奇异值重新按照升序排列,而对应的左右正交阵不变,发现重构后的图像发生了很大的变化,和原始图像不再属于同一类别(图 2)。结合图 1、图 2 及上述分析,在给定基空间下(即左右正交阵不变),奇异值向量包含了足够的分类信息,也可以体现出不同模式之间的差异。但在 ORL-NWPU1 和 ORL-NWPU2 图库上测试时发现,SV 算法的识别率为 0。对于上述两个事实,从信息量大小的角度是无法给出合理解释的,因此文献 [7] 的解释以及所提算法还有待进一步验证和研究。

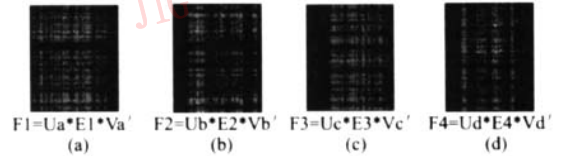


图 2 逆排序后构造的图像(对应于 A、B、C、D)

Fig. 2 Reconstructed images by ascend singular values

2.2.2 组合人脸识别方法

基于图像奇异值向量的组合方法比较多,典型的有基于奇异值向量的 PCA 识别算法以及基于奇异值向量的 FLD 识别算法。识别过程如下:

令 $TS = \{sv_i \in \mathbf{R}^k \mid i=1, 2, \dots, N\}$ 为训练集,假定 V_{PCA} (或 V_{FLD}) 代表 PCA (或 FLD) 根据 TS 得到的投影子空间。把训练集和测试集中所有人脸图像的奇异值向量向子空间 V_{PCA} (或 V_{FLD}) 投影,提取特征向量,最后根据特征向量间的距离进行分类。这种算法有较高的识别率,但通过大量实验发现,这种组合方法远没有直接利用 PCA 或 FLD 识别人脸的效果好^[2,3],文中对此并没有给出解释。

结合前面分析,认为这类组合算法识别率比较低的原因是(1)图像的奇异值向量与图像之间并不存在一一对应关系,也就是说,对于给定的奇异值向量,它既有可能属于同类模式中的两个不同图像,也有可能属于不同类别中的两个图像,即奇异值向量的类别具有不确定性;(2)奇异值向量具有不可分割性。而对没有确定类别的向量分类时,再好的分类器或识别算法也是无法把它分类出来的。此外,对具有不可分割性的向量分类时,PCA 和 FLD 算法就不适用了。

综上所述,基于奇异值向量人脸识别算法具有

低识别率的原因是 (1) 不同人脸图像的奇异值向量所在的基空间不一致 (2) 奇异值向量与人脸图像之间不存在一一对应关系 (3) 奇异值向量具有不可分割性。

3 类估计基空间的人脸识别方法

根据定理 2 以及结论 1 到结论 3, 显然有结论 4。

结论 4 对不同类别的图像 A 和 B , 不存在一个共同的左右奇异向量正交矩阵 U, V , 使得它们满足式 (1) 的分解。

根据物以类聚人以群分的思想, 认为同类图像间的左右奇异向量正交矩阵相似性比较大。因此用类估计基空间作为同类图像奇异值分解时对应的基空间, 带来的误差是比较小的(图 1 可以说明这一点)。而对于非同类模式之间的图像, 其结论是相反的。基于这样的假设, 在此提出类估计基空间的人脸识别方法。下面对两种典型代表算法进行分析。

算法 1 利用估计奇异值识别人脸

首先计算训练集每个类的图像均值, 得到对应类的模板。然后根据式 (1) 对每个类的模板进行分解, 获得每个类的估计基空间, 即左右正交阵 $U^{(j)}$ 和 $V^{(j)}$ ($j=1, \dots, C$), C 代表训练集类别数。然后利用式 (3) 提取训练图像的估计奇异值:

$$x_{p_i}^{(j)} = (u_i^{(j)})^T A_p^{(j)} v_i^{(j)} \quad (3)$$

$$i = 1 \dots k; p = 1 \dots N_j$$

其中 $x_{p_i}^{(j)}$ 代表第 j 类第 p 个图像的第 i 个估计奇异值, $u_i^{(j)}$ 和 $v_i^{(j)}$ 分别代表 $U^{(j)}$ 和 $V^{(j)}$ 的第 i 列, N_j 代表第 j 类的训练图像个数。分类待识别图像 A^* 时, 首先利用式 (3) 提取它在每个类的估计奇异值, 然后根据估计奇异值之间的关系进行分类。此算法与文献 [6] 有点相似。

算法 2 利用真实奇异值和估计正交阵构造的图像误差识别人脸

文献 [8] 提出把奇异值向量和重构误差匹配的结果进行融合, 取得了不错的分类效果。但如前所述, 该方法存在基空间不一致问题, 因此还值得进一步验证和研究。由于重构后的图像不存在基空间不一致问题, 因此提出利用构造误差识别人脸的算法, 这就是算法 2 的思想。识别过程如下:

首先计算每个类的左右正交阵 $U^{(j)}$ 和 $V^{(j)}$ ($j=1, \dots, C$), C 为训练集类别; 接着利用式 (1) 提取测

试图像 A^* 的实际奇异值 σ_i^* ; 根据式 (4) 计算 A^* 在第 j 类的构造图像 \hat{A}_j , 然后利用式 (5) 度量 \hat{A}_j 和 A^* 之间的误差, 最后根据最近准则进行分类。

$$\hat{A}_j = \sigma_1^* u_1^{(j)} (v_1^{(j)})^T + \dots + \sigma_k^* u_k^{(j)} (v_k^{(j)})^T \quad (4)$$

$$d(A^*, \hat{A}_j) = \sum_{i=1}^n \|\hat{a}_i^{(j)} - a_i^*\| \quad (5)$$

其中 $\hat{a}_i^{(j)}$ 和 a_i^* 分别代表 \hat{A}_j 和 A^* 的第 i 列, $\|\cdot\|$ 代表两个列向量之间的一种度量, 实验选用的是欧氏距离。

4 仿真结果及分析

为了证实上述分析及所提算法的有效性和正确性, 分别选用 ORL 数据库和根据 ORL 构造的图库进行实验仿真和分析。

4.1 ORL 图库及实验分析

ORL 数据库是剑桥 Olivetti 实验室拍摄的一系列灰度人脸图像, 具体包括 40 个人, 每个人含有不同表情、姿势及大小共 10 幅图像, 其大小为 112×92 。实验中的训练集分别是由每个类的前 2、3、4、5 图像组成, 共 4 个。测试集分别由训练集对应的每个类的余下图像组成, 共 4 个测试集。

图 3 和图 4 为 SV 算法和 SVD + PCA 分别利用 4 个训练集训练, 选用对应测试集测试时的分类结果。图 5 为 Eigenface 分别利用 4 个训练集训练, 选用对应测试集测试时的分类结果。图 6 和图 7 是算法 1 和算法 2 分别利用 4 个训练集训练, 选用对应测试集测试时的分类结果。

对比图 3、图 4 及图 5 可知, SV 算法和 SVD + PCA 算法的识别率都没有 Eigenface 的识别率高。根据前面分析知, 这两种算法假定不同图像的奇异值向量所在的基空间相同, 这与实际不符。另外, 人脸图像的奇异值向量具有不可分割性, 因此这两种算法的识别比较低。

对比图 3 到图 7 可知, 类估计基空间算法的识别率高于 SV 算法的识别率。根据类估计基空间算法的识别过程知, 这类方法不存在基空间不一致问题, 因此识别率比 SV 算法高。

对比图 6 和图 7 知, 算法 1 的识别率比算法 2 的识别率低。由于人脸的非刚性变化以及受外界因素的影响, 使得利用式 (3) 提取的估计奇异值具有很大的随机性, 不能保证同类模式间的估计奇异值满足某种排列(如降或升排列), 当然非同类模式间

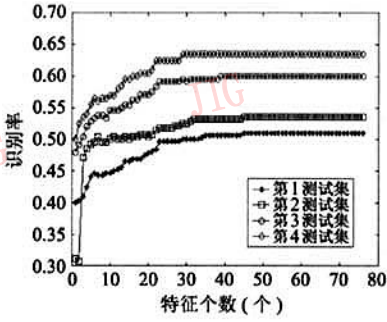


图 3 SV 算法的分类结果

Fig. 3 Recognition accuracy of SV

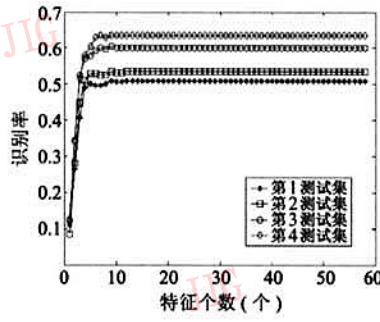


图 4 SVD + PCA 的分类结果

Fig. 4 Recognition accuracy of SVD plus PCA

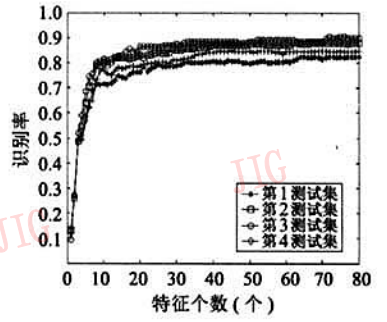


图 5 Eigenface 的分类结果

Fig. 5 Recognition accuracy of Eigenface

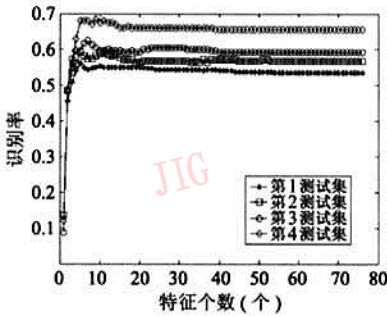


图 6 算法 1 的分类结果

Fig. 6 Recognition result of algorithm one

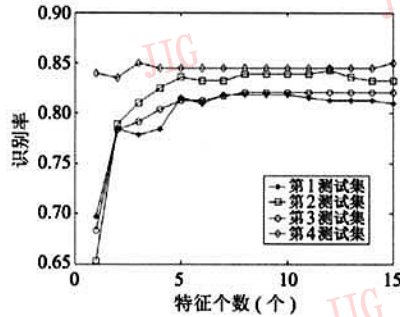


图 7 算法 2 的分类结果

Fig. 7 Recognition result of algorithm two

的估计奇异值更不用说了。因此,根据这些特征向量间的差异很难准确地反映出不同类别模式间的差异。即,一般情况下,不能保证误差最小的两个估计奇异值向量属于同一类别。因此文献[6]提出的算法(算法1)还有待进一步验证。至于文献[8]提出的基于估计奇异值和贝叶斯分类器的识别算法,同样存在这个问题。

算法2是根据构造图像误差来识别的,不存在基空间是否一致问题,而且还可以避免算法1存在的问题,因此具有较好的识别率。

4.2 构造图库及实验仿真分析

为了进一步证实算法及分析的正确性,根据 ORL 数据库分别构造了两个图库。第 1 个图库的构造过程如下:

假定选择该数据每个类的第 1 个图像构成新的训练集 $TC = \{A_i^j \mid i=1, 2, \dots, C\}$, 则根据奇异值分解定理有:

$$A_i^1 = (U_i^1)^T (D_i^1 \chi V_i^1) \quad i = 1, 2, \dots, 40 \quad (6)$$

其中 U_i^1 和 V_i^1 分别代表第 i 类第 1 个图像奇异值分解时的左右正交阵, D_i^1 代表第 i 类第 1 个图像奇

异值分解时的对角阵。

令 Z_j^i 代表第 i 类第 j 个新构造的图像, 则有:

$$Z_j^i = (U_i^j)^T (D_i^j \chi V_i^j) \quad (7)$$

$$i = 1, 2, \dots, 40; j = 1, 2, \dots, 40$$

则构造的第 1 个图库就是由所有图像 $Z_j^i (i=1, 2, \dots, 40; j=1, 2, \dots, 40)$ 组成, 其中上标 i 表示图像类别, 下标 j 表示图像个数。显然它包含 40 个人, 每个人 40 幅灰度图像, 其分辨率大小为 112×92 。图 8 为这个图库中的部分图像。可以看出, 同类人脸图像具有相同的姿势和大小。称这个图库为 ORL-NWPU1。

构造第 2 个图库的过程如下:

假定 A_j^i 代表 ORL 数据库中的第 i 类第 j 个图像, 则根据式(1)分解有:

$$A_j^i = (U_j^i)^T (D_j^i \chi V_j^i) \quad (8)$$

$$i = 1, 2, \dots, 40; j = 1, 2, \dots, 10$$

其中 U_j^i 、 V_j^i 及 D_j^i 分别代表第 i 类第 j 个图像矩阵 A_j^i 奇异值分解时的左右正交阵和对角阵。

如果令 Φ_j^i 代表第 2 个构造图库中的第 i 类第 j 个图像, 则有:



图 8 ORL-NWPU1 图库中的部分图像

Fig. 8 Original images of ORL-NWPU1

$$\Phi_j^i = (U_j^i)^T (D_1^i \chi V_j^i) \quad (9)$$

$$i = 1, 2, \dots, 40; j = 1, 2, \dots, 10$$

则这些图像 Φ_j^i 共同组成第 2 个构造图库,共包含 40 个人每人 10 幅灰度图像,其分辨率大小仍然为 112×92 。这个图库的特点是任何两个图像之间具有相同的奇异值向量。图 9 为这个图库中的部分图像,可以看出,同类人脸图像之间具有不同的姿势及大小。称这个图库为 ORL-NWPU2。



图 9 ORL-NWPU2 图库中的部分图像

Fig. 9 Some images of ORL-NWPU2

利用这两个图库验证上述分析及所提算法的正确性时,分别做了以下几个实验。第 1 个实验是选用第 1 个构造图库中的每个类的前 20 个图像组成训练集,称为第 5 训练集,然后利用余下图像进行测试(称为第 5 测试集)。第 2 个实验是把第 1 个构造图库中的所有图像看作训练集(称为第 6 训练集),然后利用 ORL 图库进行测试(称为第 6 测试集)。第 3 个实验是选用第 2 个图库中每个类的前 5 个图像组成训练集,称为第 7 训练集,然后利用余下的图像进行测试(称为第 7 测试集)。第 4 个实验是把整个第 2 个图库看作训练集(称为第 8 训练集),然后利用第 6 测试集测试。图 10 为算法 1 在 4 个实验中的分类结果,图 11 为算法 2 在 4 个实验中的分类结果。

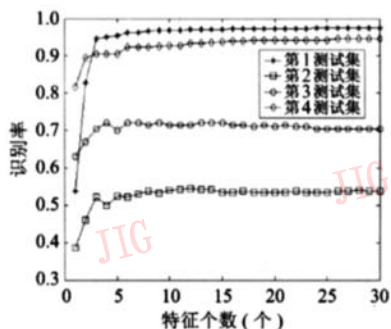


图 10 算法 1 的分类结果

Fig. 10 Recognition result of algorithm one

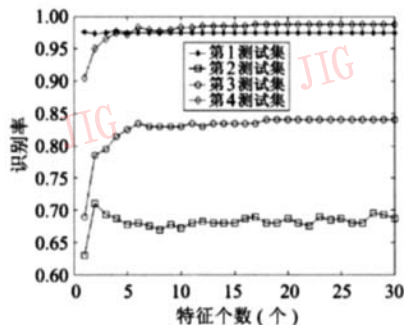


图 11 算法 2 的分类结果

Fig. 11 Recognition result of algorithm two

需要指出的是,SV 算法在 4 个实验中的识别率均为 0。根据 ORL-NWPU1 的构造过程知,对于第 5 测试集中任何一个测试图像的奇异值向量,总可以在第 5 训练集中找到一个具有相同奇异值向量的人脸图像,但它们不属于同一类别。这说明距离最近的两个奇异值向量不属于同一类别。而 SV 算法正是根据图像奇异值向量之间的距离来识别身份的。所以 SV 算法在第 1 实验中的正确识别率为 0。

根据 ORL-NWPU1 的构造过程知,在第 6 训练集中,至少可以找到两个具有相同奇异值向量的人脸图像,而且它们属于不同类别。这说明,对任意一个待识别人脸图像,与其奇异值向量距离最小的奇异值向量,在训练集中至少有两个,而且这两个不属于同一类别。此时 SV 算法无法对待识别图像进行分类,所以识别率为 0。

根据 ORL-NWPU2 的构造过程知,第 7 和第 8 训练集中的任何一个人脸图像都具有相同的奇异值向量,因此根据奇异值向量之间的距离分类时,必然存在不确定性问题,即无法分类。所以 SV 算法在第三和第四实验中的识别率也为 0。

上述实验进一步证实了,SV 算法具有低识别率的根本原因是不同人脸图像的奇异值向量所在的基

空间不一致以及奇异值向量与图像之间并不存在一一对应关系,而不是样本不足或奇异值所包含的信息不足造成的。

此外,基于奇异值向量的 PCA 以及 FLD 组合算法在第一和第二实验中的正确识别率为零。众所周知,对于确定的奇异值向量,经过 PCA 或 FLD 分析后,有唯一的特征与之对应。如上所述,在训练集中,对于任何一个图像的奇异值向量,至少可以找到另一个具有相同奇异值向量的人脸图像,但它们属于不同类别。因此经过 PCA 或 FLD 分析后,仍然不能够避免不同类别之间的图像具有相同的特征,即特征与图像不存在一一对应关系,因此这两种算法在第一和第二实验中的正确识别率也为 0。

ORL-NWPU2 图库的特点是所有图像的奇异值向量都相同。而 SVD + PCA 及 SVD + FLD 等组合算法均是把训练图像的奇异值向量看作后续 PCA 及 FLD 分析的训练集,因此在第七训练集和第八训练集中,显然每个类的均值向量和训练集均值向量相等,都等于任何一个图像的奇异值向量。此时,训练样本的协方差矩阵、类内散布矩阵以及类间散布矩阵恒为零。显然 PCA 和 FLD 不能使用。所以这两种算法在第三和第四实验中失去了作用。

结合上述两种情况分析,基于奇异值向量的 PCA 及 FLD 算法具有低识别率,甚至出现识别率为 0 或者不能使用的原因是:奇异值向量与人脸图像之间不存在一一对应关系、奇异值向量具有不可分割性。与前面分析一致。

对比图 10 和图 11 知,类估计基空间算法的识别率总体上比较好。在 4 个实验中,算法 1 和算法 2 识别率间的关系和 4.1 节的分析一致。算法 1 和算法 2 在第一和第四实验中具有很高的识别率,主要是因为第 5 训练集和第 5 测试集中,同一类别中的人脸图像具有相同的左右奇异向量正交矩阵,这使得每个类模板的左右正交阵,正好等于对应类中任何一个图像的左右正交阵,保证了测试图像在相同类别上的估计奇异值和同类图像的估计奇异值满足相同的排列,而测试图像其他类上的估计奇异值则具有随机性,不满足这个规则,保证了距离最小的两个特征向量,一般是相似基空间里的两个特征向量。

最后,根据 ORL-NWPU1 和 ORL-NWPU2 图库中人脸图像的视觉特性(图 8 和图 9)及特点知,实际中确实存在具有相同奇异值向量的不同人脸图像

(这些图像有可能属于同一类别,也可能属于不同类别)这进一步证明了第 2 节对人脸图像奇异值向量分析的正确性:即不同图像奇异值向量所在的基空间并不完全一致,且这个基空间由图像本身决定;人脸图像奇异值向量与人脸图像之间不存在严格上的一一对应关系;人脸图像的奇异值向量具有不可分割性。

5 结 论

从奇异值分解本身入手,证实了人脸图像的奇异值分解其实是人脸图像在特定基空间下的分解,而这个特定基空间是由人脸图像本身决定的。进一步研究指出,使用奇异值向量识别算法识别率低的根本原因是:不同图像的奇异值向量所在的基空间不一致,奇异值向量与人脸图像之间并不存在一一对应关系以及奇异值向量具有不可分割性。最后提出了类估计基空间人脸识别算法,并对其中的两种算法进行了分析,在 ORL、ORL-NWPU1 以及 ORL-NWPU2 数据库上的实验结果证实了分析及所提算法的正确性。

需要说明的是,在利用奇异值向量识别模式或其他应用领域中也存在着同样的问题。如何在模式识别中有效地利用奇异值是今后研究的重点。

参考文献 (References)

- 1 Hong Z Q. Algebraic feature extraction of image for Recognition [J]. Pattern Recognition, 1991, 24(3): 211 ~ 219.
- 2 Hong Zi-quan, Yang Jing-yu. Algebraic feature extraction of image for recognition [J]. Acta Automatic Sinica, 1992, 18(2): 233 ~ 237. [洪子泉 杨静宇. 用于图像识别的图像代数特征抽取 [J]. 自动化学报, 1992, 18(2): 233 ~ 237.]
- 3 Liu Rui-zhen, Tan Tie-niu. SVD based digital watermarking method [J]. Acta Electronica Sinica, 2001, 29(2): 168 ~ 171. [刘瑞祯, 谭铁牛. 基于奇异值分解的数字图像水印方法 [J]. 电子学报, 2001, 29(2): 168 ~ 171.]
- 4 Pan Quan, Zhang Min-gui, Zhou De-long, et al. Face Recognition Based on Singular Value Feature Vectors [J]. Optical Engineering, 2003, 42(8): 2368 ~ 2374.
- 5 Hong Zi-quan, Yang Jing-yu. Huamn facial image recognition algorithm based on singular value features and statistical model [J]. Computer Research and Development, 1994, 31(3): 60 ~ 65. [洪子泉 杨静宇. 基于奇异值特征和统计模型的人脸识别算法 [J]. 计算机研究与发展, 1994, 31(3): 60 ~ 65.]
- 6 Cheng Yong-qing, Liu Ke, Yang Jin-yu. A robust algebraic method for human face recognition [A]. In: International Conference on

Pattern Recognition[C]. Hague , Netherlands , 1992 : 221 ~ 224.

7 Wang Yun-hong , Tan Tie-niu , Zhu Yong. Face identification based on singular value decomposition and data fusion[J]. Chinese Journal of Computers , 2000 , 23(6) : 649 ~ 653. [王蕴红 , 谭铁牛 , 朱勇. 基于奇异值分解和数据融合的脸像鉴别 [J]. 计算机学报 , 2000 23(6) : 649 ~ 653.]

8 Tian Yuan , Tan Tie-niu , Wang Yun-hong. Do singular values contain adequate information for face recognition ? [J]. Pattern Recognition , 2003 , 36(6) : 649 ~ 655.

9 Cheng Yong-qing. Human face recognition method based on the statistical models of small samples size[A]. In : SPIE Proceedings on Intelligent Robots and Computer Vision[C] , Boston , Massachusetts , USA , 1991 : 85 ~ 95.

10 Swiniriaki R. Data mining methods in face recognition[A]. In : SPIE Proceedings of Applications of Artificial Neural Networks in Image Processing[C]. Cambridge UK , 2000 3962 : 52 ~ 59.

11 Liu Xiao-jun , Wang Dong-feng , Zhang Li-fei , et al. An approach for face recognition based on singular value decomposition and hidden Markov model[J]. Chinese Journal of Computers , 2003 , 26(3) : 340 ~ 344. [刘小军 , 王东峰 , 张丽飞等. 一种基于奇异值分解和隐马尔可夫模型的人脸识别方法 [J]. 计算机学报 , 2003 , 26(3) : 340 ~ 344.]

12 Gan Jun-ying , Zhang You-mei. A new approach for face recognition based on singular value features and neural networks[J]. Acta Electronica Sinaica , 2004 , 32(1) : 170 ~ 173. [甘俊英 , 张有为. 一种基于奇异值特征的神经网络人脸识别新途径 [J]. 电子学报 , 2004 , 32(1) : 170 ~ 173.]

13 Zhou De-long. A Study of Human Face Recognition Technology[D]. Xi 'an : Northwestern Polytechnical University , 2001. [周德龙. 人脸识别技术研究 D]. 西安 : 西北工业大学 2001.]

14 Cheng Yun-peng. Theorey of Matrix[M]. Xi 'an : Northwester Polytechnical University Preee , 2000. [程运鹏著. 矩阵论 [M]. 西安 : 西北工业大学出版社 2000.]

附录 A 证明定理 2 需要用到的有关定义及定理的证明

仿照线性基向量的定义^[12] , 有如下基矩阵的定义 :

定义 对于任何一个矩阵空间 $R^{n \times n}$, 矩阵 $W_i \in R^{n \times n} \quad i = 1, 2, \dots, r$ 满足 :

(1) W_1, W_2, \dots, W_r 线性无关 ;

(2) 矩阵空间 $R^{n \times n}$ 中的任何一个矩阵都可以通过 W_1, W_2, \dots, W_r 的线性组合来表示。则 W_1, W_2, \dots, W_r 为矩阵空间 $R^{n \times n}$ 的一组基 , 其中 $W_i (i = 1, \dots, r)$ 称为一个基矩阵。

根据线性空间基向量的个数等于线性空间的维数 , 显然有如下推论成立 :

推论 1 线性空间基向量的个数等于线性空间中任何一个向量所含有的分量个数。

根据定义及推论 1 , 容易验证下列定理成立 :

定理 A1 矩阵空间 $R^{n \times n}$ 中 , 基矩阵的个数等于矩阵空间中任何一个矩阵含有的元素个数 , 即 $r = n \times n$;

根据定义及定理 A1 , 有以下定理成立 :

定理 A2 在矩阵空间 $R^{n \times n}$ 中 , 如果 W_1, W_2, \dots, W_r 线性无关 , 且 $r = n \times n$, 则 W_1, W_2, \dots, W_r 矩阵空间 $R^{n \times n}$ 中的一组最大线性无关组 , 即矩阵基。

证明 设矩阵空间 $R^{n \times n}$ 的一个最大无关组中基矩阵的个数为 NUM , 根据定理 A1 有 :

$$NUM = n \times n \quad (A1)$$

假设 W_1, W_2, \dots, W_r 不是矩阵空间中的一组最大无关组 , 则还存在几个与 W_1, W_2, \dots, W_r 无关的基矩阵(不妨设一个) , 用 W_{r+1} 表示 , 则矩阵空间 $R^{n \times n}$ 的一个最大无关组含有的基矩阵个数为 $r + 1$:

$$r + 1 > r = n \times n \quad (A2)$$

这与式(1)相矛盾 , 故在矩阵空间中不再存在任何一个与 W_1, W_2, \dots, W_r 无关的矩阵 , 换句话说 , W_1, W_2, \dots, W_r 是矩阵空间中的一组最大线性无关组 , 即矩阵空间 $R^{n \times n}$ 中的一组基。

附录 B 定理 2 的证明

证明 假设 $u_1 v_1^T, u_1 v_2^T, \dots, u_m v_n^T$ 是线性相关的 , 则存在不全为零的常数使得下式成立 :

$$\alpha_{11} u_1 v_1^T + \alpha_{12} u_1 v_2^T + \dots + \alpha_{mn} u_m v_n^T = 0 \quad (B1)$$

不妨设 $\alpha_1 \neq 0$, 给方程两边左乘以 u_1^T 整理有 :

$$v_1^T = - \frac{\alpha_{12} v_2^T}{\alpha_{11}} + \dots - \frac{\alpha_{1n} v_n^T}{\alpha_{11}} \quad (B2)$$

式(B2)两边右乘 v_1 , 整理有 :

$$1 = 0 \quad (B3)$$

这与事实相矛盾 , 故假设错误 , 即 $u_1 v_1^T, u_1 v_2^T, \dots, u_m v_n^T$ 是一组线性无关组。

根据定理 A2 可知 $u_1 v_1^T, u_1 v_2^T, \dots, u_m v_n^T$ 是矩阵空间中的一个最大线性无关组 , 即一组矩阵基。