

基于组合特征判别分析的三维人脸识别算法

孙艳丰, 王 俊, 尹宝才

(北京工业大学 计算机学院 多媒体与智能软件技术北京市重点实验室, 北京 100124)

摘 要: 针对人脸识别中特征表示与提取问题, 提出了一种新颖的基于组合特征判别分析的三维人脸识别算法. 该算法首先使用基于非均匀网格重采样的方法对所有三维人脸做规格化处理, 使三维人脸具有统一的点数和拓扑结构; 其次, 以先分段、再重叠的形式将原本一维向量表示组织为二维矩阵表示, 然后使用二维线性判别分析方法(2DLDA)对获得的数据进行特征抽取. 这种方法在避免图像信息的丢失、增加组合特征的同时, 理论上也能避免单纯使用线性判别分析(LDA)进行特征抽取时容易出现的小样本问题. 在BJUT-3D大规模三维人脸数据库上的实验表明, 本方法取得了良好的识别效果.

关键词: 人脸识别; 特征抽取; 判别分析

中图分类号: TP 391.4

文献标志码: A

文章编号: 0254-0037(2010)01-0098-06

人脸识别技术因其广泛的应用前景得到了巨大的发展. 如在门禁系统、机场安检、出入境边检和电子商务等方面都能看到人脸识别技术的应用. 在过去的40年中, 经过研究人员的不懈努力, 基于图像的人脸识别技术日趋成熟, 在一定约束条件下已经取得较好的识别结果^[1]. 然而, 广泛的研究实验证明, 基于图像的人脸识别技术中的一些难点问题, 如光照、姿态、表情和年龄等问题并未得到良好的解决. 即便微小的光照、姿态和表情变化都会显著降低基于二维图像人脸识别算法的性能^[2].

近几年来, 研究人员开始探索利用图像之外的信息来提高人脸识别系统的识别性能, 而利用人脸的三维信息进行身份识别是其中最主要的探索方向. 随着计算、存储设备的飞速发展和三维数据获取技术的日趋成熟, 研究人员已经越来越多地将目光投向三维人脸识别研究. 三维人脸识别不同于二维人脸识别的关键在于所采用的数据不同, 因而也具有一些优势, 如: 1) 三维人脸数据被认为具有光照不变、姿态不变的特性^[3]; 2) 三维数据在信息量上比二维图像丰富, 三维数据包含人脸的空间信息, 有些三维人脸数据还有表面纹理信息, 而二维图像本质上只是三维物体在二维空间上的投影. 然而当经典的线性判别分析方法(LDA)^[4]应用于获取的三维人脸数据时, 一些问题依然存在, 如小样本问题等.

本文提出了基于组合特征判别分析的三维人脸识别方法, 该方法较好地解决了单纯使用线性判别分析可能出现的小样本问题. 在BJUT-3D大规模三维人脸数据库上分别对使用三维人脸的形状信息和纹理信息做处理得出判别结果, 并且在决策层上融合两者做出判别, 取得了良好的识别效果.

1 研究现状

产生小样本问题的原因在于, 在很多实际应用中, 用来估计散度矩阵的训练样本的数量远远小于散度矩阵的维数, 从而导致不够稳定的散度矩阵估计. 为了解决这个问题, 研究人员提出了多种方法. 一种方法是采用PCA降维, 使得降维后样本的类内散度矩阵满秩, 然后再用LDA进行处理, 称之为PCA+LDA^[4-5]. 零空间(null space)方法^[6]将所有的训练样本投影到类内散度矩阵的零空间中, 然后再对投影后的训练样本进行降维, 即可得到最优的投影向量. 类似的方法还有公共向量判别(discriminative common vector)方法^[7]等.

收稿日期: 2008-03-26.

基金项目: 国家自然科学基金项目(60533030, 60825203).

作者简介: 孙艳丰(1964—), 女, 黑龙江齐齐哈尔人, 教授.

基于二维图像矩阵的线性判别分析由于其良好的分类性能也得到了越来越多的关注. 2DFLD 方法^[8]直接采用原始图像矩阵计算图像类间散度矩阵和图像类内散度矩阵. 2DFisherface 方法^[9]先对原始图像矩阵进行 2DPCA 降维, 然后采用 IMLDA 的方法提取判别特征矩阵, A2DLDA 方法^[10]采用原始图像矩阵的转置作为输入特征. 2DLDA 方法先对原始图像矩阵采用 IMLDA 方法提取特征矩阵, 然后对特征矩阵采用 A2DLDA 方法降维, 再选择最具判别性的特征进行识别.

2 三维人脸数据的规格化

本文使用的三维人脸数据通过 CyberWare 3030RGB/PS 激光扫描仪获取, 该激光扫描仪可以一次扫描得到人头部的几何信息和彩色纹理信息. 最终获得的数据精度很高, 一个人的原始人脸数据大约由 20 万个点和 40 万个三角面组成.

对于扫描后的原始数据, 首先通过交互的方式, 使用插值、平滑的方法添补三维人脸数据的空洞, 去掉毛刺. 然后使用边缘切割的方法去除头发、肩部等部位的三维数据, 把人脸面部信息从整个头部扫描数据中分离出来. 最后对处理后的数据进行坐标矫正, 将所有人脸数据变换到朝向和姿态都相同的坐标系.

由于不同的人脸数据的点数和面数均不相同, 而且点、面的存储形式也非常松散. 而在使用本文所提出的基于组合数据的判别分析方法之前必须做到所有三维人脸数据的规格化. 规格化就是要将数据库中的人脸表示为统一的向量形式, 要建立三维人脸数据的稠密对应. 在建立了人脸数据的这种稠密对应关系后, 如果以某个人脸作为标准人脸或参考人脸, 就可以将人脸数据以标准人脸的点和面进行有序化, 从而可以表示为

$$\begin{aligned}
 \mathbf{S}_i &= (s_{i1}, s_{i2}, \dots, s_{in})^T \\
 \mathbf{T}_i &= (t_{i1}, t_{i2}, \dots, t_{in})^T
 \end{aligned}
 \quad 1 \leq i \leq N \tag{1}$$

其中, \mathbf{S}_i 是第 i 个人脸点的三维坐标组成的几何形状向量; \mathbf{T}_i 是对应于形状向量中点的 RGB 值构成的纹理向量; $s_{ij} = (x_{ij}, y_{ij}, z_{ij})$; $t_{ij} = (r_{ij}, g_{ij}, b_{ij})$; N 是人脸个数; n 是规格化后人脸的点数, 而且不同人脸向量中相同下标的点对应于相同人脸特征点.

本文使用一种基于分片的非均匀重采样对齐方法^[11], 根据人脸面部特征建立不同三维人脸数据间点到点之间的一一对应关系. 非均匀重采样对齐算法通过曲率分析考虑了人脸不同区域的形状复杂度, 重采样的疏密程度必须随着每片曲率的变化而变化, 曲率越大采样应该越密集, 曲率小的则采样稀疏. 图 1 显示了人脸的曲率分级表示图, 黑色表示曲率最高的区域, 灰色次之, 白色表示曲率最低的区域. 图 2 显示了采用不同重采样方式得到的三维人脸几何和纹理图像.

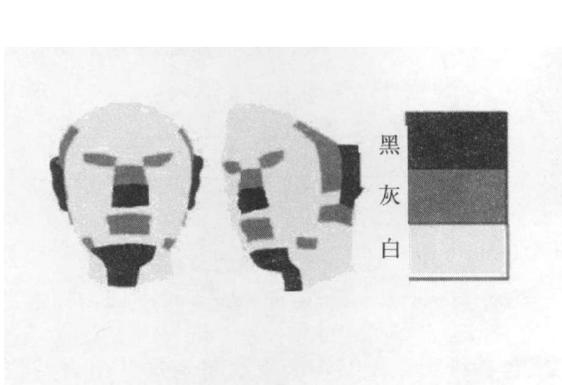


图 1 曲率分级表示图

Fig. 1 Hierarchical patches based on discrete curvature

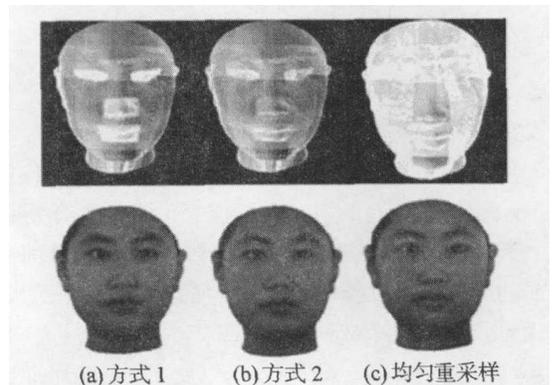


图 2 重采样后的几何和纹理图像

Fig. 2 Result of re-sampling

方式 1 的重采样^[11]得到的几何图像包含 2.605 8 万个点, 4.518 4 万个三角面片; 方式 2 的重采样^[11]得到的几何图像包含 1.345 6 万个点, 2.163 2 万个三角面片; 均匀重采样得到的几何图像包含 13.285 8

万个点, 24.9856万个三角面片. 在实验中, 本文选择采样后为2.6058万个点(方式1)的三维人脸数据作为实验数据.

3 基于特征组合的判别分析的识别算法

对于规格化后的三维人脸数据由于具有相同个数的点和统一的拓扑结构, 其形状信息以及相应的纹理信息都可以一维向量的形式来表示. 从基于二维图像人脸识别算法得到启发, 本文将三维数据应用于传统的人脸识别算法当中. 而线性判别分析作为一个优秀的分类算法, 在人脸识别中得到了广泛应用.

3.1 传统线性判别分析

线性判别分析采用对原始特征进行线性组合的方法来降维, 这一方法的主要目的是寻求在最小均方误差意义下, 可以将多个不同类别的样本分类得到效果最好的投影方法. 为了衡量不同类别样本的可分性, 需要定义一个可分性准则. 常用的准则函数为

$$J(W) = \frac{W^T S_B W}{W^T S_W W} \quad (2)$$

其中, S_B 是类间散度矩阵; S_W 是类内散度矩阵.

最优投影矩阵 W_{opt} 的列向量由 $S_W^{-1} S_B$ 的特征向量组成, 在多数人脸识别应用中, 因为训练样本的个数远远小于样本的维数, 所以样本类内散度矩阵 S_W 是奇异的, 无法直接求逆, 导致小样本问题产生. 另一方面, 特征向量的数量不能超过样本类间散度矩阵 S_B 的秩, 所以如果训练样本所包含的人数很少, 那么最终特征的维数也会很低, 不能包含足够的判别信息.

3.2 基于组合特征的判别分析

3.2.1 组合特征及其协方差

为了缓解小样本问题, 研究人员提出了很多方法, 主要有 IMLDA 和 2DFLD 等 2DLDA 算法. 在 2DLDA 中, 将图像表示为二维矩阵而不是展开的一维向量, 从而减缓了小样本问题. 受 2DLDA 的启发, 本文把经过规格化后所得到的一维向量以某种形式组合为一个二维矩阵. 由于一张三维人脸数据同时包括了形状信息向量 S 和纹理信息向量 T , 为了方便叙述, 本文统一以 U 表示任一向量进行公式推导. 即将 U 表示为

$$U = [u_1, u_2, u_3, \dots, u_n]$$

本文将此 n 维向量划分为 m 个大小为 p 的变量组, 并且组内变量依据一定次序形成一个变量, 这些向量的组合为 $X = \{x_1, x_2, \dots, x_m\}$, $n = m \times p$, $x_i = [u_{i1}, u_{i2}, \dots, u_{ij}, \dots, u_{ip}]$. 其中的 x_i 称之为组合特征. 那么为了获得更多的组合特征, 本文进行原始特征之间的叠加. 叠加的方法如图 3 所示.

从图 3 可知, 1 个组合特征由若干个原始特征点(三维人脸数据上的点)所组成, 而每个组合特征之间又有 j 个原始特征的重叠, 本文将最终获得的组合特征的个数定义为 m ($m \leq n$). 所以组合特征的集合为 $X = \{x_1, x_2, \dots, x_m\}$, 其中, $x_1 = [u_1, u_2, \dots, u_p]^T$; $x_2 = [u_{p-j+1}, u_{p-j+2}, \dots, u_{p+j}]^T$, 以此类推, 直到得到 m 个组合特征. 本文用 S 来表示基于这些组合特征而得到的散度矩阵, 那么矩阵 S 中的元素 s_{ij} 可以表示为

$$s_{ij} = E[(x_i - \bar{x}_i)^T (x_j - \bar{x}_j)], \quad i, j = 1, 2, \dots, m \quad (3)$$

其中, \bar{x}_i 表示向量 x_i 的均值; \bar{x}_j 表示向量 x_j 的均值. 那么散度矩阵 S 可以表示为

$$S = \frac{1}{N} \sum_{k=1}^N (X^{(k)} - M)(X^{(k)} - M)^T \quad (4)$$

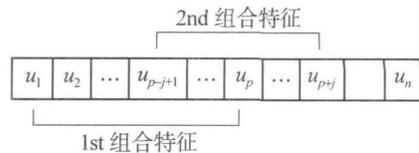


图 3 在一维向量中的组合特征

Fig. 3 Composite feature in a one-dimensional pattern

得到的散度矩阵, 那么矩阵 S 中的元素 s_{ij} 可以表示为

其中, $\mathbf{X}^{(k)} = [x_1^{(k)}, \dots, x_m^{(k)}]^T$ 是第 k 个人脸样本的组合特征集; $\mathbf{M} = [\bar{x}_1, \dots, \bar{x}_m]^T$; N 是所有样本的个数. 可得知, $\mathbf{X}^{(k)} \in \mathbf{R}^{m \times p}$, $\mathbf{S} \in \mathbf{R}^{m \times m}$.

考虑矩阵 \mathbf{S} 的秩, 分别用 $x_i^{(k)}$, $\bar{x}_i \in \mathbf{R}^m$ 表示 $\mathbf{X}^{(k)}$ 和 \mathbf{M} 的列向量. 那么 $\mathbf{X}^{(k)}$ 可以表示为 $\mathbf{X}^{(k)} = [x_1^{(k)}, \dots, x_p^{(k)}]$, \mathbf{M} 可以表示为 $\mathbf{M} = [\bar{x}_1, \dots, \bar{x}_p]$, 重写式(4)为

$$\mathbf{S} = \frac{1}{N} \sum_{i=1}^p \sum_{k=1}^N (x_i^{(k)} - \bar{x}_i)(x_i^{(k)} - \bar{x}_i)^T \quad (5)$$

式(5)中, 最多有 N_p 个线性独立的向量, 所以矩阵 \mathbf{S} 的秩最大为 N_p . 而所有的 $(x_i^{(k)} - \bar{x}_i)$ 之间并不是线性独立的, 因为与 $\sum_{k=1}^N (x_i^{(k)} - \bar{x}_i) = 0, i = 1, \dots, p$ 有关, 因此, 矩阵 \mathbf{S} 的秩为

$$\text{rank}(\mathbf{S}) \leq \min(m, (N-1)p) \quad (6)$$

3.2.2 使用组合特征的判别分析

假设训练集中的人脸数据分别为 $X^{(1)}, X^{(2)}, \dots, X^{(N)}$, 分别属于 D 个类别为 c_1, c_2, \dots, c_D . 每个类别的样本数为 $N_i, i = 1, 2, \dots, D$. 正如散度矩阵一样, 类内散度矩阵 $\mathbf{S}_W \in \mathbf{R}^{m \times m}$ 可定义为

$$\mathbf{S}_W = \sum_{i=1}^D p_i \left\{ \frac{1}{N_i} \sum_{k \in I_i} (\mathbf{X}^{(k)} - \mathbf{M}_i)(\mathbf{X}^{(k)} - \mathbf{M}_i)^T \right\} \quad (7)$$

其中, p_i 为人脸样本属于某类 c_i 的先验概率; I_i 为所有数据类 c_i 人脸训练样本的集合. 类间散度矩阵 $\mathbf{S}_B \in \mathbf{R}^{m \times m}$ 也可定义为

$$\mathbf{S}_B = \sum_{i=1}^D p_i (\mathbf{M}_i - \mathbf{M})(\mathbf{M}_i - \mathbf{M})^T. \quad (8)$$

由式(6)~(8)可知, \mathbf{S}_W 的秩为

$$\text{rank}(\mathbf{S}_W) \leq \min(m, (N-D)p). \quad (9)$$

\mathbf{S}_B 的秩为

$$\text{rank}(\mathbf{S}_B) \leq \min(m, (D-1)p). \quad (10)$$

在实际应用中, 和传统线性判别分析相比, 基于组合特征判别分析在采用较大的 n/p 值时, 即使训练样本数有限, 样本类内散度矩阵也是满秩的. 因为这种方法下的散度矩阵维数大于 n/p 维, 同时 \mathbf{S}_W 是一个 $(N-D) \times p$ 个秩为 1 且互相独立的矩阵的和, 当 p 较大时, $((N-D) \times p) > (n/p)$, 所以 \mathbf{S}_W 是满秩的, 不需要采用其他方法(如 PCA)对训练样本进行降维. 而且, 最终所提取特征的维数最多可以达到 $(D-1) \times p$, 也远远大于传统线性判别分析所能提取的特征维数. 在上述情况下, \mathbf{S}_W 的逆一般是可求的, 能有效避免小样本问题.

在组合特征判别分析中, 选择使所提取的向量特征可分性最大的变换矩阵 \mathbf{W}_{opt} , 需要最大化的准则函数为

$$\mathbf{W} = \arg \max_{\mathbf{W}} \frac{|\mathbf{W}^T \mathbf{S}_B \mathbf{W}|}{|\mathbf{W}^T \mathbf{S}_W \mathbf{W}|} \quad (11)$$

在此准则函数下最后的变换矩阵为

$$\mathbf{W}_{\text{opt}} = \arg \max_{\mathbf{W}} \text{tr}((\mathbf{W}^T \mathbf{S}_W \mathbf{W})^{-1} (\mathbf{W}^T \mathbf{S}_B \mathbf{W})) = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_l] \quad (12)$$

其中 $\{\mathbf{w}_i | i = 1, 2, \dots, l\}$ 是由前 l 个最大特征值所对应的特征向量, 即

$$\mathbf{S}_B \mathbf{w}_i = \lambda_i \mathbf{S}_W \mathbf{w}_i. \quad (13)$$

对任意的原始人脸输入数据 $\mathbf{X}^{(k)}$, 通过变换矩阵 \mathbf{W}_{opt} 所提取的向量特征集合为

$$\mathbf{Y}^{(k)} = \mathbf{W}_{\text{opt}}^T \mathbf{X}^{(k)}, \quad k = 1, 2, \dots, N. \quad (14)$$

式中 $\mathbf{Y}^{(k)} \in \mathbf{R}^{l \times p}$ 由 l 个特征向量 $[\mathbf{y}_1^{(k)}, \mathbf{y}_2^{(k)}, \dots, \mathbf{y}_l^{(k)}]^T$ 所构成.

4 实验结果与分析

为测试本文所提出算法的有效性, 在 BJUT-3D 三维人脸数据库上进行了实验. BJUT-3D 三维人脸数

数据库是北京工业大学多媒体与智能软件实验室于2005年正式公开发布的中国人的三维人脸数据库. 本文选取数据库中的105人作为实验数据, 每人3或4组数据, 共计340组三维人脸数据. 105幅人脸数据中每人的2幅人脸数据作为训练数据, 其余作为测试数据进行实验. 所有三维人脸数据都经过前述的规格化统一处理, 表示为统一的向量形式. 本文采用采样后为2.6058万个点的三维数据进行实验. 将之分为 101×258 块的组合特征块(即 $p=101$). 由文献[12]可知, 当2个组合特征重叠个数比率为50%的时候(即近似重叠50个特征点), 识别效果最为理想.

为了显示识别效果, 本文利用3种信息进行识别判断: 1) 形状信息; 2) 纹理信息; 3) 形状信息和纹理信息的结合. 形状和纹理信息分别提供了2种不同的判断值, 本文定义为 S_{shape} 和 S_{texture} , 那么两者结合得到的判断结果 S_{comb} 定义为

$$S_{\text{comb}} = (1 - \alpha)S_{\text{shape}} + \alpha S_{\text{texture}}, \quad 0 \leq \alpha \leq 1. \quad (15)$$

对于上述算法, 采用了基于欧氏距离的最近邻分类器, 以尽量减少分类器本身对识别率的影响. 本文选择了性能相对较好的2种基准人脸识别算法与本文所提算法对比在数据库上进行测试: PCA、PCA + LDA. 在形状信息与纹理信息上的识别效果分别如图4、图5所示.

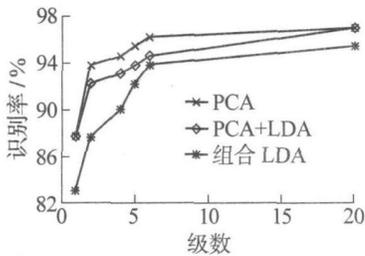


图4 基于形状信息的CMC曲线图

Fig. 4 CMC rate using shape information

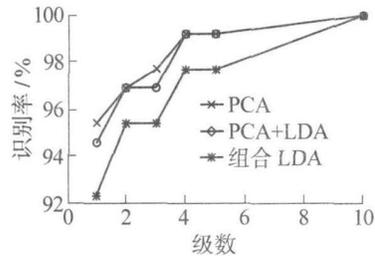


图5 基于纹理信息的CMC曲线图

Fig. 5 CMC rate using texture information

本文采用的是CMC(cumulative match characteristics)曲线图(也称之为累积识别率), CMC(r)表示在第 r 个备选结果才是正确的识别结果. 显而易见, CMC(1)即为首选识别率. 图4中, 使用形状信息的PCA算法的首选识别率为83.1%, 而使用PCA + LDA和本文提出的组合LDA则达到了87.7%, 效果明显比PCA算法好. 在纹理信息上也是类似, 图5中PCA算法的首选识别率为92.3%, PCA + LDA算法为94.6%, 而组合LDA则达到了95.4%.

图6是融合形状信息和纹理信息随着 α 值变化的识别率情况. 从图中可以得知, 融合形状和纹理信息的识别率比单纯使用形状或者纹理信息的识别效果要好. 本文提出的使用组合特征的判别分析所取得的识别效果与PCA + LDA算法不相上下, 远比使用PCA算法的效果好, 同时也能避免单纯使用LDA时出现的小样本问题.

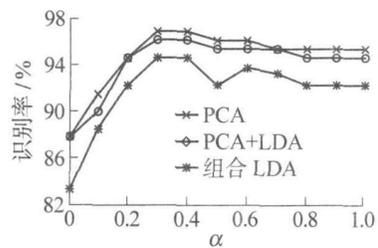


图6 随着 α 值变化的融合形状和纹理信息的识别率

Fig. 6 Recognition rate based on the fusion strategy with respect to α

5 结束语

本文提出了一种使用组合特征判别分析的三维人脸识别算法. 1个组合特征由若干个原始相邻的人脸数据点所构成. 为了避免信息的丢失并获得更多的组合特征, 采取了数据特征点之间互相重叠的方式, 使一维向量转换为二维矩阵的形式. 然后使用2DLDA对获得的数据进行特征抽取. 此算法也很好避免了小样本问题的出现. 该算法在BJUT-3D大规模三维人脸数据库进行了实验. 首先采取非均匀网格重采样的方式对所有数据进行规格化处理, 然后分别对形状、纹理信息和融合两者信息使用了文中提出的算法, 得到了较好的识别结果.

参考文献:

- [1] ZHAO W, CHELLAPPA R, ROSENFELD A. Face recognition: a literature survey [C] // Proceeding of ACM Computing Surveys. New York: ACM Press, 2003: 399-458.
- [2] MOSES Y, ADINI Y, ULLMAN S. Face recognition: the problem of compensating for illumination changes [J]. IEEE Transactions on PAMI, 1997, 19(7): 721-732.
- [3] BOWYER K W, CHANG K, FLYNN P. A survey of approaches and challenges in 3D and multi-modal 2D + 3D face recognition [J]. Computer Vision and Image Understanding, 2006, 101(1): 1-15.
- [4] BELHUMEUR P N, HESPAHNA J P, KRIEGMAN D J. Eigenfaces vs. fisherfaces; recognition using class specific linear projection [J]. IEEE Transactions on PAMI, 1997, 19(7): 711-720.
- [5] SUN Yan-feng, WANG Jun, YIN Bao-cai. A 3D face recognition algorithm based on nonuniform re-sampling correspondence [C] // Advances in Visual Computing-Third International Symposium, ISVC 2007, proceedings. Lake Tahoe: Springer Verlag, 2007, 407-416.
- [6] CHEN L, LIAO H, KO M, et al. A new LDA-based face recognition system which can solve the small sample size problem [J]. Pattern Recognition, 2000, 33(10): 1713-1726.
- [7] CEVIKAPL H, NEAMTU M, WILKE S, et al. Discriminative common vectors for face recognition [J]. IEEE Transactions on PAMI, 2005, 27(1): 4-13.
- [8] XIONG H, SWAMY M N S, AHMAD M O. Two-dimensional FLD for face recognition [J]. Pattern Recognition, 2005, 38(7): 121-1124.
- [9] JING X Y, WONG X S, ZHANG D. Face recognition based on 2D fisherface approach [J]. Pattern Recognition, 2006, 39(4): 707-710.
- [10] NOUSHATH S, KUMAR G H, SHIVAKUMARA P. (2D)2LDA: An efficient approach for face recognition [J]. Pattern Recognition, 2006, 39: 1396-1400.
- [11] 尹宝才, 何晏晏, 孙艳丰, 等. 三维人脸的非均匀重采样对齐技术 [J]. 北京工业大学学报, 2007, 33(2): 213-218. YIN Bao-cai, HE Yan-yan, SUN Yan-feng, et al. Nonuniform resampling based on method for pixel-wise correspondence between 3D faces [J]. Journal of Beijing University of Technology, 2007, 33(2): 213-218. (in Chinese)
- [12] KIM C, CHOI C H. Image covariance-based subspace method for face recognition [J]. Pattern Recognition, 2007, 40(5): 1592-1604.

3D Face Recognition Algorithm Based on Discriminant Analysis Using Composite Features

SUN Yan-feng, WANG Jun, YIN Bao-cai

(Beijing Municipal Key Laboratory of Multimedia and Intelligent Software Technology, College of Computer Science, Beijing University of Technology, Beijing 100124, China)

Abstract: This paper proposes an approach of face recognition using composite features. This approach first aligned 3D faces based on non-uniform mesh re-sampling which result the uniform number of vertex and topology structure and kept in a vector. Second, we changed the vector to the matrix by division and overlap, and then 2DLDA suggests a feature selection strategy to select the most discriminative features from the corner. This approach not only avoided face information missing, increased the number of composite feature, but also avoided the small sample size problem in theory. Experimental results for 3D face data set from BJUT-3D face database have demonstrated the performance of our algorithm.

Key words: face recognition; feature extraction; discriminant analysis

(责任编辑 苗艳玲)