

抗年龄干扰的人脸识别

吴长虹, 苏剑波, 陈叶飞

(上海交通大学自动化系, 系统控制与信息处理教育部重点实验室, 上海 200240)

摘要: 本文将人脸图像特征分解为身份特征部分和年龄干扰特征部分, 并分别投影到两个独立的子空间. 为了提高由字典张成的特征子空间的表征能力和区分能力, 字典学习过程中同时引入了人脸图像的重构误差约束项和类别约束项. 因此, 任意的人脸特征都可以由学习到的身份字典和年龄字典投影到对应的身份子空间和年龄子空间, 然后再基于身份子空间进行人脸识别, 从而使年龄的干扰得到了有效的抑制. 通过在 MORPH 和 FGNET 数据库上的实验, 证实了基于年龄不变的身份特征子空间学习方法能提升人脸识别性能.

关键词: 人脸识别; 年龄干扰; 字典分解; 子空间学习

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112 (2018)07-1593-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2018.07.008

Age Invariant Face Recognition

WU Chang-hong, SU Jian-bo, CHEN Ye-fei

(Department of Automation, Shanghai Jiao Tong University, Ministry of Education Key Laboratory of System Control and Information Processing, Shanghai 200240, China)

Abstract: In this paper, the face image is divided into two separated parts: aging effect feature part and identity related feature part. The identity dictionary and the age dictionary are introduced to encode the two feature parts into two separated feature spaces. To make sure the learned dictionaries are discriminative for different classes, the reconstruction error and label matrices constraints are added in the training. Face features can be encoded into identity and age space with the learned identity and age dictionaries. The identity space can be used for further classification. Extensive experiments are conducted on the MORPH and FGNET dataset, illustrating a great improvement over the state-of-the-arts.

Key words: face recognition; age invariance; dictionary decomposition; space learning

1 引言

近年来, 人脸识别技术已经取得了很大进展, 但对于存在较大类内干扰 (如光照、姿态、表情、年龄等) 的情况人脸识别的效果依然不是很理想. 相比于对其他干扰情况的研究^[1-3], 抗年龄干扰的人脸识别研究还相对较少, 已有的关于年龄的研究主要集中在年龄估计和衰老过程模拟上. 然而在诸多应用场景中, 抗年龄干扰的人脸识别都有着广泛的需求, 如寻找失踪儿童、辨别潜逃罪犯、身份注册认证等. 考虑到随着年龄的增长, 人脸的形状、纹理等特征会随之发生较大的改变 (如图 1 所示), 传统的直接提取人脸图像特征做分类的方法并不能获得很好的效果. 现有的抗年龄干扰的人脸识别模型主要集中在以下两个方向: 生成模型和判别

模型.



图1 FGNET 数据库中同一个人不同年龄的人脸图像

生成模型^[4-6]充分利用年龄影响过程的渐进性, 构建一个年龄影响下的人脸特征变化模型. 该模型能够利用已有年龄的人脸图像生成出其他年龄段的人脸图像, 并在此基础上进行人脸识别. 文献[4,5]基于人脸形状和

纹理的 PCA 系数建立一个年龄变化模型来模拟年龄的影响,在人脸识别之前将待测人脸图像转变到与待比对手脸图像相同年龄下的人脸.生成模型构建的过程就是一个模拟年龄变化对人脸特征影响的过程,需要较为干净的训练数据和准确的年龄估计做支撑,且模型的生成过程依赖大量的参数估计,年龄影响的仿真结果容易失真,计算的复杂度高^[7,8].因此,近年来关于抗年龄干扰的人脸识别研究主要从判别模型的角度考虑的,本文提出的模型也是属于判别模型的范畴.

判别模型^[7-14]充分利用人脸图像中与身份相关的某些特征在年龄变化过程中的稳定性,从提取年龄不变的特征的角度来克服年龄带来的干扰.文献[15]使用梯度方向金字塔 GOP 作为年龄不变的特征.文献[10]结合 MLBP 和 SIFT 两种局部特征作为年龄不变的人脸特征,并引入基于随机采样的线性判别分类器(RS-LDA)来进一步提升分类的性能.文献[8]使用一个较大的参考数据集(Cross-Age Celebrity Dataset)将原始的人脸图像特征编码到一个年龄不变的参考空间,同时引入最大池化聚合的方法保证一个人不同年龄的两张人脸能获得相似的编码特征.

判别模型的核心是将年龄的干扰信息独立出来,只提取与身份信息相关的人脸特征.然而上述文献^[8-12]只是从提取人脸图像中具有判别能力的特征角度出发,并没有主动分析年龄干扰对人脸特征的影响,合理地将年龄的干扰信息与人脸身份特征进行区分,导致最终提取的特征中不可避免地还含有年龄的干扰信息^[7].同一个人随着年龄的变化,人脸的特征会发生很大改变,但是与之身份相关的某些特征是相对稳定的,一定程度上,人脸图像特征可以看成是年龄影响下的特征与身份相关的特征的组合.基于此,本文希望通过构建两个子空间分别描述人脸图像中的年龄信息与身份信息,然后在只包含身份信息的子空间中进行人脸识别,以根本消除年龄对人脸识别的影响.

对于存在较大类内干扰的人脸图像,传统的子空间学习算法都是基于原始图像特征构建一个整体的优化目标来求投影矩阵,如图像向量的去相关性(PCA),保证类间距离最大和类内距离最小(LDA)等,并没有考虑到图像特征的不同部分对于最终识别效果的影响.文献[7,13]在一个线性概率判别模型中同时引入了年龄和身份的潜在因子,学习到一个身份子空间和一个年龄子空间,但是他们都对模型中身份和年龄因子以及随机噪声做了高斯分布的假设,并未对其合理性加以说明.文献[15]先用稀疏表示将人脸特征划分为与类别相关和与类别无关的部分,再进行子空间学习,但稀疏表示与子空间的分步学习设计忽略了两个优化目标的相关性.

稀疏表示^[16]的核心思想是原始信号可以由一组过完备的字典和稀疏系数进行表示.对于存在较大干扰的情况,单一的字典学习方法^[17]并不能将干扰信息与类别信息有效地区分开.文献[18~20]通过对字典的转化和约束使得学习到的字典能够将与类别相关的特征很好地区分出来.文献[21]将信号分解为与类别是否相关的部分和噪声部分,从低秩矩阵分解的角度进行进一步求解.但都忽略了对表示系数分类能力的约束.此外,文献[22]提出当单个类别样本数目不足或含较大类内干扰的情况下,样本并不能完全由字典中某个特定类别的样本进行表示,使用联合表示(Collaboration Representation)比稀疏表示(Sparse Representation)更加合理.

本文提出基于年龄不变的身份特征子空间学习(Age Invariant Feature Space Learning)的方法,尽可能地年龄干扰信息从身份特征中分离开.该方法将人脸图像特征划分成四部分:平均脸部分,身份特征部分,年龄特征部分以及噪声部分;引入身份字典和年龄字典分别对身份特征和年龄特征进行表示.考虑到含年龄信息的单个类别的样本是有限的,待测样本应该由字典中的多个类别联合编码获得,因此使用联合表示学习两个子空间.在字典学习过程中同时引入重构误差约束和类标监督约束项^[23].识别阶段,可以通过计算两张图像在该空间的特征向量的距离进行分类.

2 年龄不变身份特征子空间学习

2.1 人脸图像分解

将人脸图像看成四个独立的部分:平均脸部分,身份特征部分,年龄特征部分及其他噪声干扰部分.身份特征部分表示人脸特征中相对稳定的与身份相关的部分,年龄特征部分代表年龄干扰影响下的特征,平均脸部分即为所有样本对应的平均特征.而噪声即不能由字典表示的一些随机噪声.人脸图像特征分解的效果如图2所示.

假设训练集由 c 类身份类别组成,第 i 类训练样本表示为 $F_i = [f_i^1, \dots, f_i^j, \dots, f_i^{n_i}] \in \mathbf{R}^{d \times n_i}$,其中 $f_i^j \in \mathbf{R}^{d \times 1}$ 表示第 i 类训练样本的第 j 个样本.令 $F = [F_1, \dots, F_c] \in \mathbf{R}^{d \times n}$ 对应训练集样本矩阵,其中 $n = n_1 + \dots + n_c$.训练集中每个训练样本的人脸特征 $f_i^j \in \mathbf{R}^{d \times 1}$; $i = 1, \dots, c$; $j = 1, \dots, n_i$ 都可以表示为以下四部分的线性组合:身份特征部分 $I_i^j \in \mathbf{R}^{d \times 1}$,年龄干扰部分 $A_i^j \in \mathbf{R}^{d \times 1}$,平均脸部分 $m \in \mathbf{R}^{d \times 1}$,噪声部分 $\varepsilon_i^j \in \mathbf{R}^{d \times 1}$.因此,训练样本集的每个人脸图像分解可以表示为:

$$f_i^j = m + I_i^j + A_i^j + \varepsilon_i^j; i = 1, \dots, c; j = 1, \dots, n_i \quad (1)$$

2.2 基于身份字典和年龄字典的联合表示

引入身份字典 $U = [U_1, \dots, U_i, \dots, U_c] \in \mathbf{R}^{d \times p}$, U_i

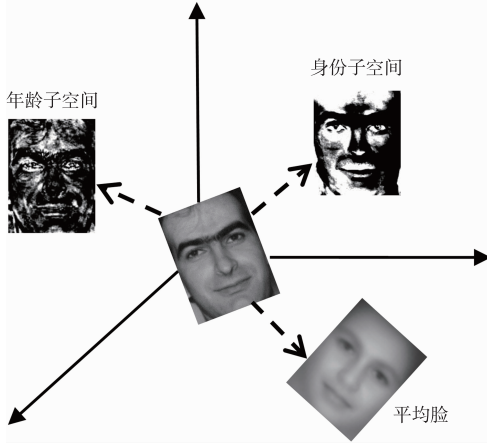


图2 人脸图像分解示意图

$\in \mathbf{R}^{d \times p}$ 和年龄字典 $\mathbf{V} = [V_1, \dots, V_s, \dots, V_t] \in \mathbf{R}^{d \times q}$, $V_s \in \mathbf{R}^{d \times q}$ 分别对训练样本矩阵 \mathbf{F} 的人脸特征的身份和年龄干扰部分进行编码. $\mathbf{X} = [X_1, \dots, X_c] \in \mathbf{R}^{p \times n}$ 为 \mathbf{F} 的身份部分对应的编码系数矩阵. $\mathbf{Y} = [Y_1, \dots, Y_t] \in \mathbf{R}^{q \times n}$ 对应 \mathbf{F} 的年龄部分的编码系数矩阵. $\mathbf{X}_i = [x_i^1, \dots, x_i^j, \dots, x_i^n] \in \mathbf{R}^{p \times n}$ 对应第 i 类样本 F_i 的身份部分的编码系数矩阵, 其中 $x_i^j \in \mathbf{R}^{p \times 1}$ 对应第 i 类样本中第 j 个样本 f_i^j 身份部分的编码系数. 同样, $\mathbf{Y}_i = [y_i^1, \dots, y_i^j, \dots, y_i^n] \in \mathbf{R}^{q \times n}$ 对应第 i 类样本 F_i 年龄部分的编码系数矩阵, 其中 $y_i^j \in \mathbf{R}^{q \times 1}$ 对应第 i 类样本中第 j 个样本 f_i^j 年龄部分的编码系数. 身份字典 \mathbf{U} 构成了与类别相关的子空间的一组基, 对应的编码系数 x_i^j 即为第 i 类第 j 个样本的身份特征部分 I_i^j 在该子空间的坐标. 同理, 年龄字典 \mathbf{V} 构成了类内年龄干扰信息子空间的一组基, 对应的编码系数 y_i^j 即为第 i 类第 j 个样本的年龄部分 A_i^j 在该子空间的坐标位置.

基于上述对人脸图像分解的分析, 本文结合重构误差约束和正则项约束构成子空间学习的目标函数. 文献[22]提出在训练样本单个类别样本不足或样本含较大类内干扰的情况下, 采用 l_2 范数比稀疏表示更加有效. 本文这里用 l_2 范数来对编码系数进行约束:

$$J_{U, V, X, Y} = \arg \min_{U, V, X, Y} \|\mathbf{F} - \mathbf{M} - \mathbf{UX} - \mathbf{VY}\|_F^2 + \lambda_1 \|\mathbf{X}\|_F^2 + \lambda_2 \|\mathbf{Y}\|_F^2 \quad (2)$$

其中 $\|\cdot\|_F$ 表示矩阵的 Frobenious 范数, 第一项 $\|\mathbf{F} - \mathbf{M} - \mathbf{UX} - \mathbf{VY}\|_F^2$ 对应信号的重构误差项, \mathbf{UX} 和 \mathbf{VY} 分别对应式(1)中人脸图像的身份特征部分 I 和年龄特征部分 A . 式(2)中对应的平均脸可以由训练集样本特征求平均得到: $\mathbf{M} = [m, \dots, m] \in \mathbf{R}^{d \times n}$, 其中

$$m = \frac{1}{n} \sum_{i=1}^c \sum_{j=1}^{n_i} f_i^j; f_i^j \in \mathbf{R}^{d \times 1}$$

n 表示训练样本集的样本数量. 为了平衡重构误差项和稀疏约束项, 引入 λ_1, λ_2 两个参数.

2.3 编码系数判别能力约束

身份字典和年龄字典张成了对应的身份子空间和年龄子空间, 为了保证子空间中身份特征 $x_i^j, i=1, \dots, c, j=1, \dots, n_i$ 和年龄特征 $y_i^j, i=1, \dots, c, j=1, \dots, n_i$ 具有足够的判别能力, 引入身份的类标约束和年龄的类标约束来保证编码系数的判别能力.

根据文献[23], 引入 $\mathbf{H}_1 \in \mathbf{R}^{c \times n}$ 作为身份特征的类标矩阵, 当第 l 个训练样本(字典)来自第 k 个人时, 该类标矩阵对应的第 k 行, 第 l 列的元素对应为 1, 该列的其他元素为 0; 引入 $\mathbf{H}_2 \in \mathbf{R}^{t \times n}$ 为年龄特征的类标矩阵, 当第 n 个训练样本来自第 m 个年龄阶段时, 该类标矩阵对应的第 m 行, 第 n 列的元素对应为 1, 该列的其他元素为 0. 对于整个训练集, 定义两个线性的转化函数 $g(\mathbf{X}; \mathbf{A}) = \mathbf{AX}, f(\mathbf{Y}; \mathbf{B}) = \mathbf{BY}$ 将编码得到的身份特征矩阵 \mathbf{X} 和年龄特征矩阵 \mathbf{Y} 转化到身份和年龄特征的判别空间, \mathbf{A} 和 \mathbf{B} 分别对应身份和年龄的转化矩阵. 由此, 加入类标矩阵约束的训练模型如下:

$$J = \arg \min_{U, V, A, B, X, Y} \|\mathbf{F} - \mathbf{M} - \mathbf{UX} - \mathbf{VY}\|_F^2 + \alpha \|\mathbf{H}_1 - \mathbf{AX}\|_F^2 + \beta \|\mathbf{H}_2 - \mathbf{BY}\|_F^2 + \lambda_1 \|\mathbf{X}\|_F^2 + \lambda_2 \|\mathbf{Y}\|_F^2 \quad (3)$$

其中 $\|\mathbf{H}_1 - \mathbf{AX}\|_F^2$ 对应身份特征的类标约束项, 它能保证同一身份类别的人编码系数 x_i^j 是相似的. 同样, $\|\mathbf{H}_2 - \mathbf{BY}\|_F^2$ 对应年龄特征的类标约束项, 它能保证同一年龄的人编码系数 y_i^j 是相似的. α, β 分别对应身份和年龄类标矩阵约束项的正则参数, 通过调整 α/β 的值, 可以调节两个子空间判别能力的权重比.

2.4 模型求解

年龄不变特征子空间的目标函数(式(3))对于字典 \mathbf{U}, \mathbf{V} , 类标转换矩阵 \mathbf{A}, \mathbf{B} 和编码系数 \mathbf{X}, \mathbf{Y} 整体的求解而言是非凸的, 然而当 \mathbf{X}, \mathbf{Y} 固定时, 只含 $\mathbf{U}, \mathbf{V}, \mathbf{A}, \mathbf{B}$ 变化量的目标函数是凸函数; 当 $\mathbf{U}, \mathbf{V}, \mathbf{A}, \mathbf{B}$ 固定时, 只含 \mathbf{X}, \mathbf{Y} 变化量的目标函数是凸函数. 因此, 模型的求解可以划分为两个子问题: 固定字典和类标矩阵, 更新编码系数; 固定编码系数, 更新字典和类标矩阵, 直至最终收敛.

首先, 固定 $\mathbf{V}, \mathbf{B}, \mathbf{X}, \mathbf{Y}$, 更新 \mathbf{U}, \mathbf{A} , 式(3)简化为:

$$J_{U, A} = \arg \min_{U, A} \|\mathbf{F} - \mathbf{M} - \mathbf{UX} - \mathbf{VY}\|_F^2 + \alpha \|\mathbf{H}_1 - \mathbf{AX}\|_F^2 \quad (4)$$

令 $\mathbf{S} = \begin{bmatrix} \mathbf{F} - \mathbf{M} - \mathbf{VY} \\ \sqrt{\alpha} \mathbf{H}_1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} \mathbf{U} \\ \sqrt{\alpha} \mathbf{A} \end{bmatrix}$, 式(4)可改为:

$$J_{D^*} = \arg \min_D \|\mathbf{S} - \mathbf{DX}\|_F^2 \quad (5)$$

\mathbf{D} 可以依据文献[24]的字典更新方法进行逐类更新, 字典 \mathbf{U} 和 \mathbf{V} 可根据行向量的维度从 \mathbf{D} 中分解得到.

其次, 固定 $\mathbf{U}, \mathbf{A}, \mathbf{X}, \mathbf{Y}$, 更新 \mathbf{V}, \mathbf{B} , 式(3)简化为:

$$J_{V^*, B^*} = \arg \min_{V, B} \|F - M - VY - UX\|_F^2 + \beta \|H_2 - BY\|_F^2 \quad (6)$$

类似地,式(6)可以转换成如下形式:

$$J_{W^*} = \arg \min_W \|Z - WY\|_F^2 \quad (7)$$

其中, $Z = \begin{bmatrix} F - M - UX \\ \sqrt{\beta}H_2 \end{bmatrix}$, $W = \begin{bmatrix} V \\ \sqrt{\beta}B \end{bmatrix}$, 同样采用文献[24]中字典更新的算法来求解式(7).

再次,固定 U, V, A, B, Y , 更新 X , 式(3)简化为:

$$J_{X^*} = \arg \min_X \|F - M - UX - VY\|_F^2 + \alpha \|H_1 - AX\|_F^2 + \lambda_1 \|X\|_F^2 \quad (8)$$

式(8)可以转换成如下形式:

$$J_{X^*} = \arg \min_X \left\| \begin{bmatrix} F - M - VY \\ \sqrt{\alpha}H_1 \end{bmatrix} - \begin{bmatrix} U \\ \sqrt{\alpha}A \end{bmatrix} X \right\|_F^2 + \lambda_1 \|X\|_F^2 \quad (9)$$

式(9)是一个凸函数,因此身份特征的编码系数可以直接计算得到:

$$X^* = \left(\begin{bmatrix} U^T & \sqrt{\alpha}A^T \end{bmatrix} \begin{bmatrix} U \\ \sqrt{\alpha}A \end{bmatrix} + \lambda_1 I \right)^{-1} * \begin{bmatrix} U^T & \sqrt{\alpha}A^T \end{bmatrix} \begin{bmatrix} F - M - VY \\ \sqrt{\alpha}H_1 \end{bmatrix} \quad (10)$$

最后,固定 U, V, A, B, X , 更新 Y , Y 的更新过程与 X 的近似.

$$Y^* = \left(\begin{bmatrix} V^T & \sqrt{\beta}B^T \end{bmatrix} \begin{bmatrix} V \\ \sqrt{\beta}B \end{bmatrix} + \lambda_2 I \right)^{-1} * \begin{bmatrix} V^T & \sqrt{\beta}B^T \end{bmatrix} \begin{bmatrix} F - M - UX \\ \sqrt{\beta}H_2 \end{bmatrix} \quad (11)$$

综上,身份、年龄字典的求解如算法 1 所示.

算法 1 身份字典和年龄字典求解算法

输入: c 个身份类别的训练样本矩阵 $F = [F_1, \dots, F_c] \in \mathbf{R}^{d \times n}$, 参数 $\alpha, \beta, \lambda_1, \lambda_2$.

输出: 身份字典 U 和年龄字典 V .

① 初始化

计算平均脸特征 M , 参照 2.2 节;

计算类标矩阵 H_1 和 H_2 , 参照 2.3 节;

对训练样本分别按照身份和年龄类别进行 PCA 降维, 初始化身份字典 U 和年龄字典 V ;

初始化编码系数矩阵 X, Y , 式(2) (假设 $\lambda_1 = \lambda_2$) 可简化为式

(12), $\begin{bmatrix} X^* \\ Y^* \end{bmatrix}$ 可直接求得.

$$J_{X^*, Y^*} = \arg \min_{X, Y} \|F - M - [U \ V] \begin{bmatrix} X \\ Y \end{bmatrix}\|_F^2 + \lambda_1 \left\| \begin{bmatrix} X \\ Y \end{bmatrix} \right\|_F^2 \quad (12)$$

初始化转换矩阵: $A = H_1 X^{-1}, B = H_2 Y^{-1}$.

② 固定 V, X, Y , 使用文献[24]中的字典更新方法更新式(4)中的 U, A .

③ 固定 U, X, Y , 使用文献[24]中的字典更新方法更新式(4)中的

V, B .

④ 固定 U, V, A, B, Y , 根据式(9)更新 X .

⑤ 固定 U, V, A, B, X , 根据式(11)更新 Y .

⑥ 迭代条件判断

达到预设的迭代次数或损失小于阈值后停止迭代, 否则转到②.

2.5 模型评价

(1) 模型如何确保字典 U 学到的是身份信息, V 学到的是年龄信息?

从特征描述角度, $\|F - M - UX - VY\|_F^2$ 项用人脸的平均脸信息、身份特征、年龄特征以及随机噪声来重构整个人脸图像特征, $UX + VY$ 对应于人脸中的身份信息和年龄信息的总和.

从任务完成的角度出发, 为了让 UX 包含完成人脸识别任务所需要的身份信息, 让 VY 包含年龄识别所需要的年龄信息, 引入了对编码系数 X 和 Y 的约束: $(\alpha \|H_1 - AX\|_F^2)$ 和 $(\beta \|H_2 - BY\|_F^2)$, 使得 X, Y 分别具有足够的判别能力来区分不同的身份类别和年龄类别, 由此, UX 对于身份信息的表示能力得到了保证, VY 对于年龄信息的表示能力得到了保证. 因此, 学到的字典 U 和 V 能够将人脸图像中的身份部分和年龄部分进行有效的区分, 使得编码的系数 X 和 Y 能够完成人脸识别和年龄识别的任务.

图 3 给出了在 FGNET 数据库上求解的身份子空间、年龄子空间的可视化结果. 由字典 U 表示的身份特征部分 UX 确实包含了人脸身份的相关信息, 由字典 V 表示的年龄特征部分 VY 包含了人脸年龄相关的纹理特征.

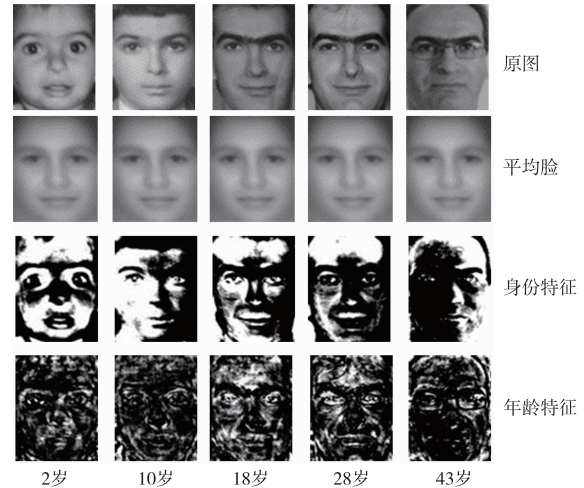


图3 子空间学习的可视化效果

由于实际的训练样本中不仅仅包含年龄的干扰信息, 还可能包含其他的干扰项. 如图 4 所示, 年龄子空间学到特征除了年龄的干扰部分, 还包括了眼镜、姿态等其他干扰信息. 本文考虑的是存在较大年龄干扰的人

脸识别任务,年龄字典和对应类标约束项的引入能够增强模型对年龄干扰的表示能力,保证与年龄识别相关的信息只由字典 V 张成,确保了身份字典 U 张成的子空间中不含年龄相关的信息.



图4 年龄子空间中包含眼镜等其他干扰信息

(2)在抗年龄干扰的人脸识别中,采用联合表示还是稀疏表示?

文献[25]证实了当训练数据在满足一定的条件时, l_1 范数的有效性是可以保证的,但是当训练样本存在很大干扰信息的时候,低维的线性假设便不再成立.文献[22]认为当字典不是过完备时,联合表示(使用 l_2 范数)比稀疏表示(使用 l_1 范数)的识别效果更好.在抗年龄干扰的人脸识别任务中,单个类别的样本数目是有限的且训练样本中包含较大的年龄干扰,过完备的字典是无法获得的,因此使用联合表示(使用 l_2 范数)更加合理.

3 抗年龄干扰的人脸识别框架

抗年龄干扰的人脸识别整体框架如图5所示.训练阶段,所有的训练样本根据身份类别和年龄类别进行分组,基于上述年龄不变的身份特征子空间学习算法获得身份字典和年龄字典张成的两个独立子空间.在测试阶段,任意待比对的图像都可以由上述的身份字典和年龄字典投影到对应的身份空间和年龄空间.然后基于身份子空间编码系数的距离进行人脸识别.

3.1 局部特征提取

相比于全局特征,局部特征在人脸图像特征多尺度、多方向的表示上更加有效,方向梯度直方图(HOG)特征是已有的局部特征描述算法中效果相对较好的一种^[7,13].为了便于与已有的抗年龄干扰的人脸识别方法进行比较,本文采用HOG特征作为特征描述子,具体的设置如表1所示.

考虑到提取的HOG特征维度很高,需要对特征进行降维,本文采用PCA来获得较低维度的特征.

表1 HOG特征描述算子参数设置

块的大小	重叠因子	方向数
10	0.9	9

3.2 人脸比对

对单个人脸图片提取完HOG特征并采用PCA降维后,获得表示人脸局部特征的较低维度的特征向量 $f \in \mathbf{R}^{d \times 1}$.该特征向量可以由学习到的身份字典 $U \in \mathbf{R}^{d \times p}$ 和

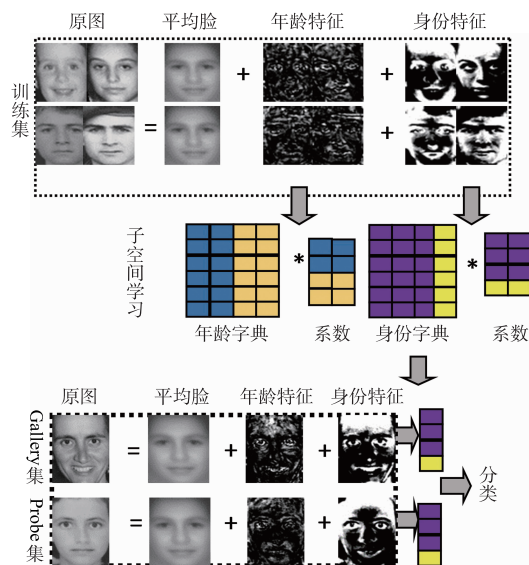


图5 年龄不变的身份特征子空间学习人脸识别框架

年龄字典 $V \in \mathbf{R}^{d \times p}$ 编码到对应的身份子空间和年龄子空间,编码系数 $x \in \mathbf{R}^{p \times 1}$ 即为年龄不变的身份特征.

$$J_{x^*, y^*} = \arg \min_{x, y} \left\| f - m - [U \quad V] \begin{bmatrix} x \\ y \end{bmatrix} \right\|_F^2 + \lambda_1 \left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_2 \quad (13)$$

由此, x^*, y^* 可直接由式(13)计算得到:

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = \left([U^T \quad V^T] \begin{bmatrix} U \\ V \end{bmatrix} + \lambda_1 I \right)^{-1} * [U^T \quad V^T] (f - m) \quad (14)$$

为了证明所提模型的有效性,本文直接计算样本年龄不变的身份特征 x 两两之间的距离,采用最近邻算法做分类.

4 实验与分析

4.1 数据库

包含年龄信息的人脸数据库主要有三个:FG-NET^[26]、MORPH^[27]和CACD^[28],具体信息如表2所示.

表2 含年龄信息的人脸数据库比对

数据库	样本数目	每个类别 样本数	数据库	年龄跨度
			年龄范围	
FGNET	1002	12.2	0-69	0-45
MORPH	55134	4.1	16-77	0-5
CACD	163446	81.7	10-60	0-10

FGNET数据库样本数相对较小,1002张图片,82个不同的身份类别,同一个人的年龄跨度相对较大,最大为从0岁到45岁.MORPH数据库学术版有55134张人脸图像,13618个人,平均每个人只有四张图片,每个人的年龄跨度相对较小,且包含较大的姿态、表情干扰.CACD数据库数据量最大,2000个人,163446张图片,平

均每个人有 82 张图片,但同类样本的年龄跨度也相对较小. 考虑到 CACD 包含一些重复和错误标记的图片, 本文在 MORPH 和 FGNET 数据库上验证所提模型的有效性, 所有训练集划分成不同的年龄组, 每 5 岁的年龄跨度作为一个年龄组.

4.2 参数设置

年龄不变的身份子空间学习算法中涉及到以下的参数: PCA 降维以后的维度 d , 类标约束的参数 α, β 以及正则项的参数 λ_1, λ_2 . 对于 PCA 降维维度, 本文将 d 从 100 提高到 2000, 当 $d=900$ 维时识别性能最好. 对于类标约束参数 α, β 和正则项参数 λ_1, λ_2 . 本文采用交叉验证的方法来寻找最优值. FGNET 数据库上 $(\alpha, \beta, \lambda_1, \lambda_2) = (2.3, 1.2, 0.005, 0.005)$, MORPH 库 $(\alpha, \beta, \lambda_1,$

$\lambda_2) = (2.3, 0.01, 0.002, 0.002)$.

4.3 FGNET 数据库上的实验

依据文献[7, 10]中训练集和测试集的划分, 本文采用留一法 (leave-one-person-out) 来验证模型的有效性. FGNET 中 82 个人的样本均参与实验. 81 个人的样本用来做训练, 另外一人的样本用来测试. 对每个待测试的类别随机挑选 10 对样本, 每对中的第一张图像与其他 81 个类别的样本组成 gallery 集, 另一张人脸图像作为 probe 集.

表 3 给出了实验的比较结果, 本文提出的方法识别率为 79.8%, 比其他已发表的方法效果更好些. 此外, 基于年龄不变的身份特征子空间学习模型的人脸识别比直接提 HOG 特征效果好很多.

表 3 抗年龄干扰的人脸识别算法在 FGNET 数据库上的比较

方法	数据库(类别数, 图像数) Probe 集和 Gallery 集中	识别率
Geng 等人(2008) ^[4]	生成模型: 构建人脸随年龄变化的 2D 模型, 依据该年龄影响模型将待测人脸图像转变到待比对图像对应年龄下的图像再做识别. FGNET(10, 10)	38.1%
Park 等人(2010) ^[5]	生成模型: 基于人脸的形状和纹理特征构建三维的人脸年龄变化模型来弥补年龄带来的影响. FGNET(82, 1002)	37.4%
Li 等人(2010) ^[14]	判别模型: 结合多个局部特征 LBP, SIFT 等作为年龄不变的特征, 采用基于随机采样的线性判别分类器 (RS-LDA) 构建判别模型. FGNET(82, 82)	47.5%
Gong 等人(2013) ^[7]	判别模型: 基于身份因子和年龄因子的隐因子分析概率判别模型. FGNET(82, 82)	69.0%
Gong 等人(2015) ^[13]	判别模型: 提出基于最大熵的特征算子 (MEFD) 并采用概率匹配模型直接计算两个人脸图像中身份特征部分相似的概率. FGNET(82, 82)	76.2%
本文方法	直接提取 HOG 特征, 采用最近邻算法做分类. 判别模型: 年龄不变的身份特征子空间学习算法. FGNET(82, 82)	53.7%
		79.8%

图 6 给出正确识别的部分人脸图像, 第一行为待测试的人脸图像, 第二行对应识别结果, 与正确的结果一致.

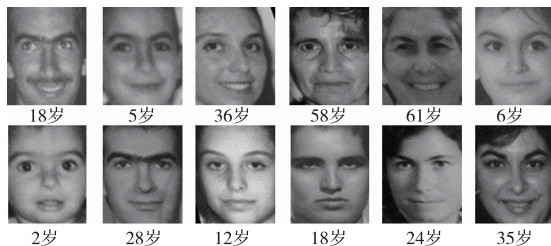


图6 FGNET 数据库中识别正确的部分人脸图像

图 7 给出错误识别的部分人脸图像. 第一行为待测试的人脸图像, 第二行对应识别结果, 第三行对应真实的结果. 可以发现, 由于年龄带来巨大的干扰, 识别结果比真实结果更加接近测试的人脸图像.

4.4 MORPH 数据库上的实验

本文在 MORPH Album2 数据库上进行了实验, 依据文献[7, 10]对训练集和测试集的划分, 随机选取 10000 人作为测试集, 每个人选取年龄跨度最大的两张图片, 年龄较小的 10000 张人脸图片作为 gallery 集, 年

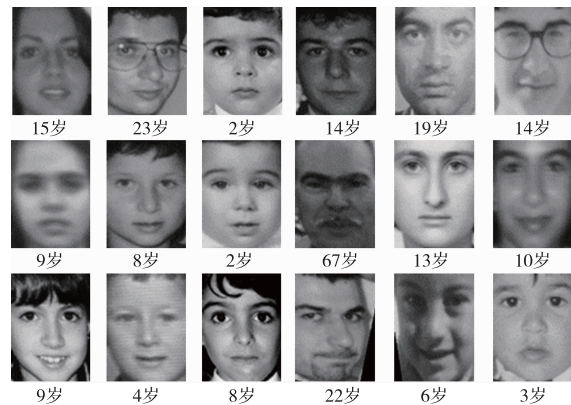


图7 FGNET 数据库中识别错误的部分人脸图像

龄较大的作为 probe 集, 再另外选取其他的 1000 人作为训练集, 实验结果如表 4 所示.

本文所提方法针对的是年龄跨度较大的人脸识别, 因为 MORPH 数据库单个人的训练样本较少, 年龄跨度很小, 且包含较大的姿态、表情的干扰, 并不能充分展现本文所提方法的优越性, 尽管如此, 本文所提方法依然有相对较好的识别效果, 比直接提 HOG 特征效

果好很多.

表 4 抗年龄干扰的人脸识别算法在 MORPH 数据库上的比较

方法	识别率
Park 等人(2010) ^[5]	79.80%
Li 等人(2010) ^[10]	83.90%
Du 等人(2012) ^[6]	79.24%
Klare 等人(2011) ^[9]	79.08%
Otto 等人(2012) ^[12]	81.27%
HOG + HFA (2013) ^[7]	91.14%
HLBP + CARC (2014) ^[8]	92.80%
HOG + IFA (2015) ^[13]	92.26%
文献[28]中的方法	87.13%
HOG	43.65%
本文方法(HOG + AIFSL + Euclidean)	88.43%
本文方法(HOG + AIFSL + Cosine)	89.57%

5 结论

本文提出了一种抗年龄干扰的人脸识别方法,该方法将人脸特征字典分解为身份字典和年龄字典,学习到对应的身份子空间和年龄子空间,再基于身份子空间进行人脸识别可有效地克服年龄干扰的影响.

在 MORPH 和 FGNET 数据库上的实验结果证明了本文方法的有效性.此外,本文引入的年龄字典可以被替换为表情、光照等其他干扰字典,有望用来解决类似的类内干扰较大的人脸识别问题.

参考文献

- [1] 杨利平,李武.光照健壮人脸识别的低秩相对梯度直方图特征[J].电子学报,2016,44(8):1940-1946.
YANG L P, LI W. Low-rank relative gradient histogram features for illumination-robust face recognition[J]. Acta Electronica Sinica, 2016, 44(8): 1940-1946. (in Chinese)
- [2] 平强,庄连生,俞能海.姿态和光照可变条件下的仿射最小线性重构误差人脸识别算法[J].电子学报,2012,40(10):1965-1970.
PING Q, ZHUANG L S, YU N H. Affine minimum linear reconstruction error face recognition under varying pose and illumination[J]. Acta Electronica Sinica, 2012, 40(10): 1965-1970. (in Chinese)
- [3] 辜小花.基于视觉感知和边缘保持的光照不变人脸识别[J].电子学报,2013,41(8):1500-1505.
GU X H. Visual perception and edge preserving illumination invariant face recognition[J]. Acta Electronica Sinica, 2013, 41(8): 1500-1505. (in Chinese)
- [4] GENG X, ZHOU Z H, SMITHMILES K. Automatic age estimation based on facial aging patterns[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 29(12): 2234.
- [5] PARK U, TONG Y, JAIN A K. Age-invariant face recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2010, 32(5): 947-954.
- [6] DU J X, ZHAI C M, YE Y Q. Face aging simulation and recognition based on NMF algorithm with sparseness constraints[J]. Neurocomputing, 2013, 116(10): 250-259.
- [7] GONG D, LI Z, LIN D, et al. Hidden factor analysis for age invariant face recognition[A]. Proceedings of the IEEE International Conference on Computer Vision [C]. USA: IEEE Computer Society, 2013. 2872-2879.
- [8] CHEN B C, CHEN C S, HSU W H. Cross-age reference coding for age-invariant face recognition and retrieval[A]. Proceedings of the European Conference on Computer Vision [C]. Cham: Springer, 2014. 768-783.
- [9] KLARE B, JAIN A K. Face recognition across time lapse: On learning feature subspaces[A]. Proceedings of the International Joint Conference on Biometrics [C]. USA: IEEE, 2011. 1-8.
- [10] LI Z, PARK U, JAIN A K. A discriminative model for age invariant face recognition[J]. IEEE Transactions on Information Forensics & Security, 2011, 6(3): 1028-1037.
- [11] LING H, SOATTO S, RAMANATHAN N, et al. Face verification across age progression using discriminative methods[J]. IEEE Transactions on Image Processing, 2006, 15(11): 3349.
- [12] OTTO C, HAN H, JAIN A. How does aging affect facial components? [A]. Proceedings of the International Conference on Computer Vision [C]. Berlin: Springer-Verlag, 2012. 189-198.
- [13] GONG D, LI Z, TAO D, et al. A maximum entropy feature descriptor for age invariant face recognition[A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. USA: IEEE, 2015. 5289-5297.
- [14] LI Z, GONG D, LI X, et al. Aging face recognition: A hierarchical learning model based on local patterns selection [J]. IEEE Transactions on Image Processing, 2016, 25(5): 2146.
- [15] ZHANG L, ZHU P, HU Q, et al. A linear subspace learning approach via sparse coding[A]. Proceedings of the IEEE International Conference on Computer Vision [C]. USA: IEEE, 2012. 755-761.
- [16] WRIGHT J, MA Y, MAIRAL J, et al. Sparse representation for computer vision and pattern recognition[J]. Proceedings of the IEEE, 2010, 98(6): 1031-1044.
- [17] AHARON M, ELAD M, BRUCKSTEIN A. ℓ_1/ℓ_2 -SVD: An algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311-4322.
- [18] LIU H D, YANG M, GAO Y, et al. Bilinear discriminative

- dictionary learning for face recognition [J]. *Pattern Recognition*, 2014, 47(5):1835–1845.
- [19] SUN Y, LIU Q, TANG J, et al. Learning discriminative dictionary for group sparse representation [J]. *IEEE Transactions on Image Processing*, 2014, 23(9):3816–3828.
- [20] ZHAO Y, SU J B. A structured sparse learning approach for efficient facial feature description [A]. *Proceedings of the IEEE International Conference on Information and Automation* [C]. USA: IEEE, 2013. 870–874.
- [21] JIANG X, LAI J. Sparse and dense hybrid representation via dictionary decomposition for face recognition [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 37(5):1067–1079.
- [22] ZHANG L, YANG M, FENG X. Sparse representation or collaborative representation: Which helps face recognition? [A]. *Proceedings of the International Conference on Computer Vision* [C]. USA: IEEE Computer Society, 2011. 471–478.
- [23] JIANG Z, LIN Z, DAVIS L S. Label consistent K-SVD: Learning a discriminative dictionary for recognition [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2013, 35(11):2651–2664.
- [24] YANG M, ZHANG L, YANG J, et al. Metaface learning for sparse representation based face recognition [A]. *Proceedings of the IEEE International Conference on Image Processing* [C]. USA: IEEE, 2013. 1601–1604.
- [25] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2009, 31(2):210.
- [26] FG-NET Aging Database [OL]. Available: <http://www.fgnet.rsunit.com/>. 2017-06-22.
- [27] JR K R, TESAFAYE T. MORPH: A longitudinal image database of normal adult age-progression [A]. *Proceedings of the International Conference on Automatic Face and Gesture Recognition* [C]. USA: IEEE, 2006. 341–345.
- [28] LI Z, GONG D, LI X, et al. Learning compact feature descriptor and adaptive matching framework for face recognition [J]. *IEEE Transactions on Image Processing*, 2015, 24(9):2736–2745.

作者简介



吴长虹 女, 1992 年 8 月生于江苏. 上海交通大学自动化系硕士研究生, 研究方向为人脸识别、图像处理.

E-mail: wuchanghong@sjtu.edu.cn



苏剑波 男, 1969 年 11 月生于江苏. 上海交通大学自动化系教授, 研究方向为机器视觉、机器学习与人机交互、多传感器信息融合、智能机器人等.



陈叶飞 男, 1988 年生于上海. 现为上海交通大学自动化系博士, 研究方向为人脸识别、深度学习.