

结合加权子空间和相似度量学习的人脸验证方法研究^{*}

汤红忠^{1,2}, 张小刚¹, 陈华^{3†}, 李骁², 王翔²

(1. 湖南大学 电气与信息工程学院, 湖南 长沙 410082; 2. 湘潭大学 信息工程学院, 湖南 湘潭 411105; 3. 湖南大学 信息科学与工程学院, 湖南 长沙 410082)

摘要:在无约束条件下,人脸表情、姿态、光照以及背景等复杂因素可能导致人脸图像的类内变化大于类间变化.针对如何降低较大的类内变化对人脸验证研究的影响,本文结合加权子空间,提出了一种带先验相似性和先验距离约束的相似度量学习方法.首先,利用类内人脸对样本,学习带权重的类内协方差矩阵,通过加权子空间的投影,从人脸图像中获得鲁棒性的人脸特征表达;其次,利用样本对的相似性与差异性,建立了带先验相似性和先验距离约束的相似度量学习模型,优化后的度量矩阵可以有效提高特征向量的类内鲁棒性和类间判别性;最后,利用优化的度量矩阵计算人脸对的相似度.在LFW(Labeled Faces in the Wild)数据集的实验验证了所提模型的有效性,与其它同类相似度量学习方法相比,优化的度量矩阵更能准确地评估人脸间的相似性,并在受限训练集上取得了91.2%的识别率.

关键词:类内变化;加权子空间;相似度量学习;人脸验证

中图分类号:TP391.4

文献标志码:A

Face Verification Based on Weighted Subspace and Similarity Metric Learning

TANG Hongzhong^{1,2}, ZHANG Xiaogang¹, CHEN Hua^{3†}, LI Xiao², WANG Xiang²

(1. College of Electrical and Information Engineering, Hunan University, Changsha 410082, China;

2. College of Information Engineering, Xiangtan University, Xiangtan 411105, China;

3. College of Information Science and Engineering, Hunan University, Changsha 410082, China)

Abstract: Under the unconstrained conditions, intra-personal variation is much larger than the inter-personal variation in face images due to the affecting factors such as expression, posture, illumination and background etc. To reduce the influence of larger intra-personal on face verification, we proposed a similarity metric learning method with priori similarity and priori distance constraint by combining weighted subspace. First, the weighted intra-personal covariance matrix is learned by employing intra-personal face samples. By projecting into the intra-subspace, robust face feature representations can be obtained from face images. Second, we set up the similarity metric learning model with priori similarity and priori distance

^{*} 收稿日期:2017-04-06

基金项目:国家自然科学基金资助项目(61573299, 61673162, 61672216), National Natural Science Foundation of China(61573299, 61673162, 61672216); 湖南省自然科学基金资助项目(2017JJ3315, 2017JJ2251, 2016JJ3125), National Natural Science Foundation of Hunan Province(2017JJ3315, 2017JJ2251, 2016JJ3125)

作者简介:汤红忠(1979—),女,湖南衡山人,湖南大学副教授

[†] 通讯联系人, E-mail: annychen@126.com

constraint, which effectively employs the similarity and discrimination information of samples that are in pairs, and the learned metric matrix can improve the robustness to intra-personal and discrimination to inter-personal. Finally, the updated metric matrix is used to compute the similarity scores of face-pairs. The experiments have been conducted on the Labeled Faces in the Wild (LFW) dataset, which shows the effectiveness of our proposed model. Compared with other metric learning methods, our learned metric matrix has higher accuracy rate for evaluating the face-pair similarity, and achieves a verification rate of 91.2% on the restricted setting.

Key words: intra-personal variations; weighted subspace; similarity metric learning; face verification

人脸验证是指预测一对人脸图像是否表示同一个人,是一个典型的二分类问题,可广泛应用于视频监控、公共安全、电子商务等生物特征识别、智能视频分析及计算机视觉系统。与传统的约束条件下的人脸验证问题相比,无约束条件下的人脸图像因受表情、姿态、光照以及背景等复杂因素的干扰,造成同一个人的不同人脸图像差异性较大,如何克服上述因素引起的人脸图像较大的类内变化,是当前人脸验证中亟需解决的重要问题之一,近几年来也引起了国内外广大学者的关注^[1-5]。

相似度度量指利用给定的人脸样本对,计算人脸对特征向量间的距离来判断人脸的身份,因其实现简单,计算代价小、可分性好等特点成为一种主要的人脸验证方法^[6-7]。因此,“特征提取+相似度度量”是解决人脸验证的有效方法,如 Liu 等^[8]利用欧氏距离或余弦距离计算一对人脸的 Gabor 特征向量间的相似度;然而,直接距离的度量没有考虑人脸表情、姿态、光照以及背景等影响,因此并不能准确地反映一对人脸之间的相似性。为了减少复杂干扰对相似度度量的影响,有较多研究者围绕相似度度量学习(Similarity Metric Learning, SML)方法开展工作并取得一定的效果^[9-20],主要分为以下三类。

1) 基于概率的相似度度量学习:如 Guillaumin 等^[9]提出一种逻辑判别度量学习(Logistic discriminant metric learning, LDML)方法,该方法利用概率估计样本对的相似度,建立最大似然估计模型,采用梯度上升方法优化求解。Mignon 等^[10]提出了一种成对约束成分分析(Pairwise constrained component analysis, PCCA)的相似度度量学习方法,该方法采用逻辑损失函数优化相似样本对的距离小于某个阈值,并保持非相似样本对的距离大于某个阈值。Köstinger 等^[11]基于高斯分布的假设,提出了保持简单直接的度量(Keep it simple and straight, KISS)方法,该方法利用似然比检验直接进行较大规模的距离度量,取得了较好的效果。2) 基于变体的

代价函数:如 Ying 等^[12]提出一种基于特征值优化框架的度量学习方法,通过迭代计算最大特征值来优化度量矩阵,最后采用最近邻实现分类。3) 基于正则项的相似度度量学习: Cao 等^[13]结合余弦相似度度量(Cosine similarity metric learning, CSML)^[14]和马氏距离,提出一种子空间相似度度量(Subspace similarity metric learning, Sub-SML)学习模型,但该模型中仅将度量矩阵与单位矩阵的差值作为目标函数的正则项,并不能准确表达样本对的先验信息,因此不能有效地抑制复杂干扰给分类带来的影响。随着稀疏性与低秩性在图像表示中的广泛应用, Fang 等^[15]引入群稀疏性作为正则项,提出一种稀疏相似度度量模型。文献[16]与[17]利用低秩性作为相似度度量模型的正则项。稀疏性与低秩性的约束能有效地解决噪声干扰问题。Hassner 等^[18]提出一种不同姿态下的人脸图像摆正方法,并结合 Sub-SML 方法实现人脸验证; Hu 等^[19]提出一种深度相似度度量学习(Discriminative deep metric learning, DDML)方法,该方法采用深度神经网络模型优化相似度度量矩阵,但由于神经网络中参数调整缺乏理论指导,较难寻优,鲁棒性也有待提高。Zhu 等^[20]利用 3D 形变模型提出一种姿态与表情标准化方法,同时结合 DDML 方法进行人脸验证,这两种方法人脸验证性能的提高主要依赖于前期对人脸图像姿态矫正的工作。

近年来,深度学习因具有主动的特征提取与较强分类能力在人脸验证领域中得到了广泛的应用,在深度模型中,通常将特征提取、编码与池化等嵌入到网络训练过程中。如 DeepFace^[21], DeepID^[22], DeepID2^[23], DeepID3^[24]都大大提高了人脸验证的分类性能,但大都采用了大规模的带类标户外数据集来训练深度模型,训练时间长。DeepID 因 LFW 数据集太小而引入了外部 CelebFaces 和 CelebFaces+数据集, DeepFace 分类性能也大大依赖了前期

图像预处理的人脸姿态矫正与对齐.而且,这些实验参数设置方法皆与本文不同,较多的参数设置调整都需经验指导.

子空间方法可将样本投影到某个最优子空间,实现降低输入特征维数和寻找分类特征的目的,因此被广泛应用于人脸识别^[13,25-27].文献[13]利用主成分分析(Principal component analysis, PCA),将人脸样本对投影到类内子空间,以期望降低图像类内变化产生的干扰.文献[25]采用线性判别分析(Linear Discriminant Analysis, LDA)的方法用于人脸识别,结合类内散度矩阵与类间散度矩阵,使得人脸特征在类间散度矩阵和类内散度矩阵上的比值达到最大从而实现分类.文献[26]在PCA的基础上进一步提出贝叶斯人脸识别,其目的也是通过子空间投影来降低特征维度与噪声干扰.文献[27]提出一种加权线性判别分析方法,该方法利用了加权成对Fisher准则约束,并在较小的人脸数据集取得了较好的效果.文献[13,25-27]的研究结果表明,类内子空间的投影可以有效降低类内变化的干扰,使得投影后的特征更利于后续的分类研究.

综上所述,为降低相似人脸对之间的类内差异的影响,同时提取较强的判别信息.本文围绕正则项相似度量学习问题,提出一种结合加权子空间和相似度量学习的人脸验证方法.一方面,基于PCA预处理后人脸对的类内样本,学习具有权重的类内协方差矩阵,通过加权子空间的投影,从人脸图像中学习更具有鲁棒性的人脸特征表达.与文献[13]相比,本文提出的“PCA+加权子空间”的特征投影方法能够有效地降低因人脸表情、姿态、光照以及背景产生的较大类内变化,获得判别性更强的分类特征(如表2、表3、表4的实验结果所示);另一方面,本文利用了样本对的相似性与差异性信息,结合先验相似性矩阵 G_0 与先验距离矩阵 M_0 ,建立了带先验相似性和先验距离约束的相似度量学习模型,该模型更能准确衡量样本对的相似性和差异性先验信息,并利用对偶函数求解,优化后的度量矩阵能有效提高特征向量的类内鲁棒性和类间辨别性.最后,利用优化的度量矩阵计算人脸对的相似度得分.实验结果表明,本文方法提出的相似度量学习模型,更能准确地评估人脸间的相似性.

1 加权子空间

给定人脸样本对 $X = \{(x_{i1}, x_{i2})\}, i = 1, \dots,$

N ,其中, $x_{i1}, x_{i2} \in R^d$ 表示一对人脸样本.若 $(x_{i1}, x_{i2}) \in S$ 表示两个样本对应同一个人,若 $(x_{i1}, x_{i2}) \in D$ 表示两个样本对应不同的人.人脸样本对的类内协方差矩阵定义如下^[13]:

$$C_S = \sum_{(x_{i1}, x_{i2}) \in S} (x_{i1} - x_{i2})(x_{i1} - x_{i2})^T \tag{1}$$

由式(1)可知,类内的人脸样本对之间的距离体现了样本对之间的可分性.然而,所有样本对的权值相同,会造成人脸样本对在投影后的空间产生特征信息的重叠.同时,人脸样本对 (x_{i1}, x_{i2}) 相似度越高,则 $\|x_{i1} - x_{i2}\|_2$ 值越小,对类内协方差矩阵的贡献度越少,为提高相似度较高的样本对协方差矩阵的特征值及其特征向量的影响,本文提出了一种加权协方差矩阵,其定义如下:

$$WC_S = \sum_{(x_{i1}, x_{i2}) \in S} w_i(x_{i1} - x_{i2})(x_{i1} - x_{i2})^T \tag{2}$$

其中, w_i 表示类内样本对 (x_{i1}, x_{i2}) 的权值,其定义为:

$$w_i = \begin{cases} \frac{1}{\|x_{i1} - x_{i2}\|_2^u}, & \|x_{i1} - x_{i2}\|_2 < a \\ 1, & \|x_{i1} - x_{i2}\|_2 \leq a \end{cases} \tag{3}$$

其中, $u=2, a=0.8$.由式(3)可知,类内样本对相似度越高,距离越小($\|x_{i1} - x_{i2}\|_2 < a$)则赋予较大权值,可以提高类内协方差矩阵对应的特征值及特征向量的影响,而类内样本对的相似度低于某个值($\|x_{i1} - x_{i2}\|_2 \geq a$)则将其权值赋为1,可以有效平衡类内样本对的相似度信息,从而降低因表情、姿态、光照以及背景产生的人脸样本对类内变化的干扰.

将式(2)进行特征值分解,可得其特征值 $\Lambda = \{\lambda_1, \dots, \lambda_d\}$ 及对应的特征向量 $P = (v_1, \dots, v_d)$,式(2)可转化为:

$$WC_S = P\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}P^T \tag{4}$$

令

$$L_S = P\Lambda^{\frac{1}{2}} \tag{5}$$

结合式(4)与式(5),可得:

$$WC_S = L_S L_S^T \tag{6}$$

因此,人脸样本对在类内子空间的投影过程定义如下:

$$\tilde{x} = L_S^{-1}x \tag{7}$$

由式(7)可知,人脸样本对 $X = \{(x_{i1}, x_{i2})\}, i = 1, \dots, N, x_{i1}, x_{i2} \in R^d$,通过加权子空间的投影,可得到投影后的人脸样本对 $\tilde{X} = \{(\tilde{x}_{i1}, \tilde{x}_{i2})\}, i = 1, \dots, N$,其中 $\tilde{x}_{i1}, \tilde{x}_{i2} \in R^d$.

2 相似度量学习

2.1 相似度量函数

人脸样本对 $(\tilde{x}_{i1}, \tilde{x}_{i2})$ 之间的相似度量,实际上要寻找一个度量矩阵 M 来测量样本对之间的距离,通常采用马氏距离进行距离度量,其定义如下:

$$d_M(\tilde{x}_{i1}, \tilde{x}_{i2}) = \frac{\| \tilde{x}_{i1} - \tilde{x}_{i2} \|^2_M}{(\tilde{x}_{i1} - \tilde{x}_{i2})^T M (\tilde{x}_{i1} - \tilde{x}_{i2})} \quad (8)$$

其中, $M \in R^{d \times d}$ 为半正定矩阵.

相似度量学习的目的是通过度量矩阵 M 可将原始特征空间映射到另一个更具判别性特征空间,从而较好地分类.文献[13]提出一种结合余弦相似度与马氏距离的相似度量模型,则人脸样本对 $(\tilde{x}_{i1}, \tilde{x}_{i2})$ 的相似度定义如下:

$$f_{(M,G)}(\tilde{x}_{i1}, \tilde{x}_{i2}) = s_G(\tilde{x}_{i1}, \tilde{x}_{i2}) - d_M(\tilde{x}_{i1}, \tilde{x}_{i2}) \quad (9)$$

其中, $s_G(\tilde{x}_{i1}, \tilde{x}_{i2}) = \tilde{x}_{i1}^T G \tilde{x}_{i2}$ 表示余弦相似度函数, $d_M(\tilde{x}_{i1}, \tilde{x}_{i2})$ 表示距离相似度函数.

若 $(x_{i1}, x_{i2}) \in S$, 则定义类标 $y_i = 1$; 反之, 若 $(x_{i1}, x_{i2}) \in D$, 则 $y_i = -1$. 因此, 当 $y_i = 1$ 时, 学习的度量矩阵 G 和 M 应使得式(9)的值较大; 反之, 当 $y_i = -1$ 时, 式(9)的值较小. 基于这一规则, hinge 损失函数定义如下:

$$\epsilon_{\text{emp}} = \sum_{(\tilde{x}_{i1}, \tilde{x}_{i2}) \in \tilde{x}} (1 - y_i f_{(M,G)}(\tilde{x}_{i1}, \tilde{x}_{i2})) \quad (10)$$

通过最小化式(10)的经验误差可以得到学习的 G 和 M . 但这种方法很容易造成函数的过拟合. 因

此增加相应的正则项约束, 式(10)可转化为:

$$\min_{M,G} \sum_{(\tilde{x}_{i1}, \tilde{x}_{i2}) \in \tilde{x}} (1 - y_i f_{(M,G)}(\tilde{x}_{i1}, \tilde{x}_{i2})) + \frac{\gamma}{2} (\|M - I\|_F^2 + \|G - I\|_F^2) \quad (11)$$

在式(11)中, 目标函数的正则项 $\gamma(\|M - I\|_F^2 + \|G - I\|_F^2)/2$ 保证了度量矩阵 (G, M) 在单位矩阵 I 的一定邻域内具有最优解. 很明显, 以单位矩阵 I 作为相似度量函数的约束条件不能准确衡量样本对的先验信息, 因此, 度量函数并不具备类内鲁棒性和类间判别性.

2.2 先验距离度量

Köstinger^[11]从统计学角度出发, 提出一种 KISSME 的算法. 其中, 样本对 $(\tilde{x}_{i1}, \tilde{x}_{i2})$ 的相似度定义如下:

$$\delta(\tilde{x}_{i1}, \tilde{x}_{i2}) = \log \frac{P[(\tilde{x}_{i1}, \tilde{x}_{i2}) \in D]}{P[(\tilde{x}_{i1}, \tilde{x}_{i2}) \in S]} \quad (12)$$

其中, $P[(\tilde{x}_{i1}, \tilde{x}_{i2}) \in S]$ 表示样本对属于同一个人的概率, $P[(\tilde{x}_{i1}, \tilde{x}_{i2}) \in D]$ 表示样本对属于不同人的概率. $\delta(\tilde{x}_{i1}, \tilde{x}_{i2})$ 值越小表示 $(\tilde{x}_{i1}, \tilde{x}_{i2})$ 对应同一个人的概率越大, $\delta(\tilde{x}_{i1}, \tilde{x}_{i2})$ 值越大则表示 $(\tilde{x}_{i1}, \tilde{x}_{i2})$ 对应不同人的概率越大. $P[\cdot]$ 是符合均值为 0 的单变量高斯分布概率密度函数, 其定义如下:

$$P[\cdot] = \frac{1}{\sqrt{2\pi} \left| \sum \right|} \exp \left(-\frac{1}{2} (\tilde{x}_{i1} - \tilde{x}_{i2})^T \times \sum^{-1} (\tilde{x}_{i1} - \tilde{x}_{i2}) \right) \quad (13)$$

将式(13)代入式(12), 可得:

$$\delta(\tilde{x}_{i1}, \tilde{x}_{i2}) = \log \frac{\frac{1}{\sqrt{2\pi} \left| \sum_D \right|} \exp \left(-\frac{1}{2} (\tilde{x}_{i1} - \tilde{x}_{i2})^T \sum_D^{-1} (\tilde{x}_{i1} - \tilde{x}_{i2}) \right)}{\frac{1}{\sqrt{2\pi} \left| \sum_S \right|} \exp \left(-\frac{1}{2} (\tilde{x}_{i1} - \tilde{x}_{i2})^T \sum_S^{-1} (\tilde{x}_{i1} - \tilde{x}_{i2}) \right)} \quad (14)$$

忽略式(14)的常数项, 进一步简化可得:

$$\delta(\tilde{x}_{i1}, \tilde{x}_{i2}) = (\tilde{x}_{i1} - \tilde{x}_{i2})^T (\sum_S^{-1} - \sum_D^{-1}) (\tilde{x}_{i1} - \tilde{x}_{i2}) \quad (15)$$

其中, $\sum_D = \sum_{(\tilde{x}_{i1}, \tilde{x}_{i2}) \in D} (\tilde{x}_{i1} - \tilde{x}_{i2}) (\tilde{x}_{i1} - \tilde{x}_{i2})^T$, $\sum_S = \sum_{(\tilde{x}_{i1}, \tilde{x}_{i2}) \in S} (\tilde{x}_{i1} - \tilde{x}_{i2}) (\tilde{x}_{i1} - \tilde{x}_{i2})^T$.

在本文中, 我们令 $G_0 = \sum_S^{-1}$, $M_0 = \sum_S^{-1} - \sum_D^{-1}$, 由式(15)可知, $\sum_S^{-1}$ 可有效衡量类内样本对的先验相似性, 而 $\sum_S^{-1} - \sum_D^{-1}$ 能有效衡量样本对类内分布与类间分布的先验距离, 结合式(11),

本文建立一种带先验相似性与先验距离约束的相似度量学习模型, 将先验相似性矩阵 G_0 和先验距离矩阵 M_0 作为目标函数的正则项约束条件, 其定义如下:

$$\min_{M,G} \sum_{(\tilde{x}_{i1}, \tilde{x}_{i2}) \in \tilde{x}} (1 - y_i f_{(M,G)}(\tilde{x}_{i1}, \tilde{x}_{i2})) + \frac{\gamma}{2} (\|M - M_0\|_F^2 + \|G - G_0\|_F^2) \quad (16)$$

由此可见, 式(16)中的正则项约束条件有效利用了样本对的先验相似性和差异性信息, 比式(11)预先假设的单位矩阵 I 更能体现样本的类内与类间分布, 因此度量函数具备类内鲁棒性和类间判别性.

2.3 相似度量学习

将式(16)引入松弛变量,可得:

min_{M,G} \zeta_i + \frac{\gamma}{2} (\| M - M_0 \|_F^2 + \| G - G_0 \|_F^2)

s. t. y_i [f_{(M,G)} (\tilde{x}_{i1}, \tilde{x}_{i2})] \geq 1 - \zeta_i

\zeta_i \geq 0, \forall i \in \{ 1, \cdots, N \}

(17)

目标函数式(17)是一个凸优化问题,进一步采用拉格朗日函数表示如下:

L(\alpha, \beta; G, M, \zeta) = \sum_{i=1}^N \zeta_i + \frac{\gamma}{2} (\| M - M_0 \|_F^2 + \| G - G_0 \|_F^2) - \sum_{i=1}^N \alpha_i (y_i [s_G (\tilde{x}_{i1}, \tilde{x}_{i2}) - d_M (\tilde{x}_{i1}, \tilde{x}_{i2})] - 1 + \zeta_i) - \sum_{i=1}^N \beta_i \zeta_i

(18)

将式(18)对 M, G, \zeta 分别进行求导:

\left\{ \begin{aligned} \frac{\partial L}{\partial M} &= \gamma (M - M_0) + \sum_{i=1}^N \alpha_i y_i (\tilde{x}_{i1} - \tilde{x}_{i2}) (\tilde{x}_{i1} - \tilde{x}_{i2})^T \\ \frac{\partial L}{\partial G} &= \gamma (G - G_0) - \sum_{i=1}^N \alpha_i y_i \tilde{x}_{i1} \tilde{x}_{i2}^T \\ \frac{\partial L}{\partial \zeta_i} &= \zeta_i - \alpha_i \zeta_i - \beta_i \zeta_i \end{aligned} \right.

(19)

并令 \frac{\partial L}{\partial M} = 0, \frac{\partial L}{\partial G} = 0, \frac{\partial L}{\partial \zeta_i} = 0, 可得:

\left\{ \begin{aligned} M &= M_0 - \frac{1}{\gamma} \sum_{i=1}^N \alpha_i y_i (\tilde{x}_{i1} - \tilde{x}_{i2}) (\tilde{x}_{i1} - \tilde{x}_{i2})^T \\ G &= G_0 + \frac{1}{\gamma} \sum_{i=1}^N \alpha_i y_i \tilde{x}_{i1} \tilde{x}_{i2}^T \\ \alpha_i + \beta_i &= 1 \end{aligned} \right.

(20)

将式(20)代入到式(18)得到相应的对偶函数为:

\max_{0 \leq \alpha_i \leq 1} \sum_{i=1}^N \alpha_i y_i (\| \tilde{x}_{i1} - \tilde{x}_{i2} \|^2 - \tilde{x}_{i1}^T \tilde{x}_{i2}) - \frac{1}{2\gamma} (\| \sum_{i=1}^N y_i \alpha_i (\tilde{x}_{i1} - \tilde{x}_{i2}) (\tilde{x}_{i1} - \tilde{x}_{i2})^T \|_F^2 + \| \sum_{i=1}^N y_i \alpha_i \tilde{x}_{i1} \tilde{x}_{i2}^T \|_F^2)

(21)

求解对偶函数式(21),得到解析解 \alpha^*, 则式

(17)的最优解 (M^*, G^*) 如下:

\left\{ \begin{aligned} M^* &= M_0 - \frac{1}{\gamma} \sum_{i=1}^N \alpha_i^* y_i (\tilde{x}_{i1} - \tilde{x}_{i2}) (\tilde{x}_{i1} - \tilde{x}_{i2})^T \\ G^* &= G_0 + \frac{1}{\gamma} \sum_{i=1}^N \alpha_i^* y_i \tilde{x}_{i1} \tilde{x}_{i2}^T \end{aligned} \right.

(22)

2.4 本文方法的具体操作步骤

本文采用PCA 对人脸特征进行预处理,学习具有权重的类内协方差矩阵,利用加权子空间的投影进一步降低人脸特征对类内差异的影响.并建立带先验相似性和先验距离约束的相似度量学习模型,将优化后的度量矩阵解决人脸验证问题.其方法流程如表 1 所示.

表 1 本文方法的具体操作步骤
Tab.1 Our proposed method

本文方法的具体操作步骤
输入: N 对人脸图像, 参数 u, a, \gamma
输出: N 对人脸图像的相似度得分
1. 提取每对人脸图像的特征(例如 LBP[28], TPLBP[29], SIFT[9]), 将每对人脸特征实现 PCA 变换, 得到人脸样本对的特征向量 X = \{ (x_{i1}, x_{i2}) \}, i = 1, \cdots, N, 其中, x_{i1}, x_{i2} \in R^d;
2. 结合式(2)和式(5)得到类内加权子空间 L_S;
3. 基于式(7)对 X = \{ (x_{i1}, x_{i2}) \}, i = 1, \cdots, N, 进行加权子空间的投影, 得到投影后的特征向量 \hat{x} = \{ (\hat{x}_{i1}, \hat{x}_{i2}) \}, i = 1, \cdots, N;
4. 基于式(15)得到先验相似性矩阵 G_0 和先验距离矩阵 M_0;
5. 基于式(16)~式(21)优化度量矩阵(G, M);
6. 将本文优化的度量矩阵(G^*, M^*) (式(22))代入式(9)计算人脸对的相似度得分.

3 实验内容与结果分析

3.1 实验数据

LFW(Labeled faces in the wild)数据集是目前人脸验证领域中最具有挑战性的数据集之一[30], 包括了来自 5 749 个人的 13 233 张人脸图像, 其中有 1 680 个人具有至少 2 张以上的人脸图像. 所有图像均采集于 Yahoo 网站, 如图 1 所示, 这些人脸图像在表情、姿态、光照、年龄以及背景存在较大的类内变化.



图 1 每列两张人脸图片对应于同一个人
Fig. 1 Same person with two images in each row

本文实验主要针对受限训练集,划分成 10 个子集,每个子集由 300 个匹配(matched)图像对和 300 个不匹配(mismatched)图像对组成,并且每个子集中的人都不重合. 并且,具有两种人脸对齐方法:“funneled”数据集^[30]与“aligned”数据集^[31]. 本文在“funneled”数据集上提取的是 SIFT 特征^[9],在“aligned”数据集上提取的是 LBP 特征^[28]和 TPLBP 特征^[29]. 这一特征选择方案设计主要参考了文献[9, 13, 19]的工作,文献[9]已证明在“funneled”数据集上, SIFT 特征的分类效果最优;基于文献[9],文献[13]进一步验证了在“funneled”数据集上提取 SIFT 特征^[9],在“aligned”数据集上提取的是 LBP 特征^[28]和 TPLBP 特征^[29]进行分类的有效性. 测试采用 10 折交叉验证方法,进行 10 次独立的实验,轮流选定 1 个子集作为测试集,而其余 9 个子集作为训练集.

3.2 本文提出方法的性能分析

为了有效地说明本文提出方法具有较好的特征投影与相似度量度的能力,本文利用 SIFT^[9]与 LBP^[28]特征,投影方法分别对比了 PCA^[32]、PCA+子空间^[13]与本文提出的“PCA+加权子空间”,其中 PCA 的维数取 300,采用 L₂范数对投影后的数据进行归一化处理,相似度量度分别对比了马氏距离度量、先验距离度量与本文优化的度量矩阵,从而证明本文提出的相似度量度学习模型的有效性,实验结果如表 2、表 3、表 4 所示. 其中,马氏距离度量、先验距离度量与本文优化的度量矩阵的定义如下.

1) 马氏距离度量

若 $M=G=I$ 时,式(9)则为直接的马氏距离度量,定义如下:

$$f_{G,M}(\tilde{x}_{i1}, \tilde{x}_{i2}) = (2 - 3\tilde{x}_{i1}^T \tilde{x}_{i2})/2 \tag{23}$$

2) 先验距离度量

若 $M=M_0, G=G_0$,式(9)则为非学习模型下的先验距离度量,定义如下:

$$f_{G,M}(\tilde{x}_{i1}, \tilde{x}_{i2}) = \tilde{x}_{i1}^T G_0 \tilde{x}_{i2} - (\tilde{x}_{i1} - \tilde{x}_{i2})^T M_0 (\tilde{x}_{i1} - \tilde{x}_{i2}) \tag{24}$$

3) 本文优化的度量矩阵

若 $M=M^*, G=G^*$,将式(22)代入式(9),则为本文优化的度量矩阵,定义如下:

$$f_{G,M}(\tilde{x}_{i1}, \tilde{x}_{i2}) = \tilde{x}_{i1}^T G^* \tilde{x}_{i2} - (\tilde{x}_{i1} - \tilde{x}_{i2})^T M^* (\tilde{x}_{i1} - \tilde{x}_{i2}). \tag{25}$$

当人脸特征为 SIFT 时,不同投影方法下不同度量矩阵的人脸验证结果如表 2 所示. 由表 2 可知,若特征投影方式相同(如“PCA+加权子空间”),本文优化的度量矩阵人脸验证率最高,原始特征与平方根特征分别为 87.12%与 87.62%. 传统的马氏距离度量人脸验证率最低,原始特征与平方根特征分别为 83.43%与 84.22%. 这一实验结果表明:带先验相似性与先验距离约束的相似度量度学习模型(式(16))能有效提高人脸验证性能.

采用相同的度量学习方法,进一步对比分析不同的投影方法,由表 2 可知,结合本文提出的“加权子空间”的分类性能最好,优于文献[13]提出的“未加权子空间”. 而且,表 3 与表 4 也给出了相同的实验结果. 由此可知,本文提出的“PCA+加权子空间”的特征投影方法能够获得具有鲁棒性的类内子空间. 提取的人脸特征经过类内子空间投影后,可有效地降低因人脸表情、姿态、光照以及背景产生的较大类内变化,获得判别性更强的分类特征.

表 2 基于 SIFT 特征,不同投影方法下不同度量矩阵的人脸验证结果对比
Fig. 2 Based on SIFT features, the results of face verification with different metric matrix under different projection methods

不同的度量矩阵	不同投影方法	原始特征	平方根特征
马氏距离度量矩阵(式(23))	PCA ^[32]	0.772 3±0.005 3	0.785 5±0.003 5
	PCA+子空间 ^[13]	0.821 8±0.002 7	0.829 5±0.002 3
	PCA+加权子空间	0.834 3±0.005 2	0.842 2±0.002 5
先验距离度量矩阵(式(24))	PCA ^[32]	0.789 3±0.002 3	0.792 5±0.002 4
	PCA+子空间 ^[13]	0.829 4±0.004 2	0.834 5±0.001 5
	PCA+加权子空间	0.854 3±0.002 2	0.852 2±0.001 4
本文优化的度量矩阵(式(25))	PCA ^[32]	0.808 4±0.001 4	0.805 5±0.003 2
	PCA+子空间 ^[13]	0.869 8±0.002 7	0.872 5±0.002 3
	PCA+加权子空间	0.871 2±0.001 4	0.876 2±0.002 6

如表 3 所示,当人脸特征为 LBP 时,结合不同投影方法与不同度量矩阵的人脸验证结果. 为了进一步验证本文提出方法的普适性,利用已训练好深度学习模型 VGG-face^[33]提取的人脸特征,结合不

同投影方法与不同度量矩阵进行人脸验证,实验结果如表 4 所示. 表 4 的数据结果与表 3、表 2 大致相符,进一步说明本文提出的结合加权子空间和相似度量度学习的人脸验证方法的有效性,不仅适合于

传统的手工特征,也适用于学习型特征.

表 3 基于 LBP 特征,不同投影方法下不同度量矩阵的人脸验证结果对比
Fig. 3 Based on LBP features, the results of face verification with different metric matrix under different projection methods

不同的度量矩阵	不同投影方法	原始特征	平方根特征
马氏距离度量矩阵(式(23))	PCA ^[32]	0.804 7±0.005 1	0.809 8±0.003 8
	PCA+子空间 ^[13]	0.832 3±0.002 5	0.844 5±0.004 3
	PCA+加权子空间	0.851 8±0.005 0	0.855 0±0.002 3
先验距离度量矩阵(式(24))	PCA ^[32]	0.823 4±0.003 1	0.829 8±0.002 5
	PCA+子空间 ^[13]	0.842 3±0.005 5	0.847 4±0.002 3
	PCA+加权子空间	0.861 8±0.001 3	0.865 0±0.003 1
本文优化的度量矩阵(式(25))	PCA ^[32]	0.834 7±0.002 1	0.839 8±0.003 2
	PCA+子空间 ^[13]	0.846 7±0.003 4	0.854 5±0.002 3
	PCA+加权子空间	0.877 2±0.005 2	0.879 8±0.003 6

表 4 基于 VGG-face 特征,不同投影方法下不同度量矩阵的人脸验证结果对比
Fig. 4 Based on VGG-face features, the results of face verification with different metric matrix under different projection methods

不同的度量矩阵	不同投影方法	原始特征	平方根特征
马氏距离度量矩阵(式(23))	PCA ^[32]	0.805 8±0.004 2	0.811 2±0.003 3
	PCA+子空间 ^[13]	0.844 5±0.003 7	0.850 8±0.004 2
	PCA+加权子空间	0.859 8±0.004 6	0.861 8±0.001 5
先验距离度量矩阵(式(24))	PCA ^[32]	0.832 2±0.002 5	0.838 1±0.004 0
	PCA+子空间 ^[13]	0.850 2±0.002 8	0.852 9±0.005 0
	PCA+加权子空间	0.863 6±0.003 2	0.869 2±0.002 4
本文优化的度量矩阵(式(25))	PCA ^[32]	0.842 8±0.002 1	0.843 4±0.003 6
	PCA+子空间 ^[13]	0.861 9±0.005 3	0.872 6±0.003 4
	PCA+加权子空间	0.879 5±0.004 2	0.883 6±0.003 2

3.3 参数分析

参数 γ 平衡了经验误差与正则项两者的性能. γ 设置过小会导致模型的鲁棒性和判别性不强,泛化能力较弱; γ 设置过大会导致模型不易收敛. 本文中 γ 的设置参考了文献[13]的设置方法,并结合多组数据对比结果将其设置为 0.5.

为了分析参数 u, a 对本文方法性能的影响,图 2、图 3 分别给出在两种不同数据集上,不同 u, a 对本文方法性能的对比实验. 通过对比实验可知,本文方法性能在 $u=2, a=0.8$ 附近达到最优.

3.4 与其他的人脸验证方法的实验对比

根据实验 3.3 的结果,利用 SIFT^[9]、LBP^[28] 以及 TPLBP^[29] 的原始特征与平方根特征,采用“PCA+加权子空间”的投影方法,基于 2.4 的步骤,根据式(25)计算得到的六种特征的相似度得分,将其进行融合并采用 SVM 预测得到最终实验结果. 分类准确率和标准差如表 5 所示,ROC(Receiver Operating Characteristic)曲线如图 4 所示.

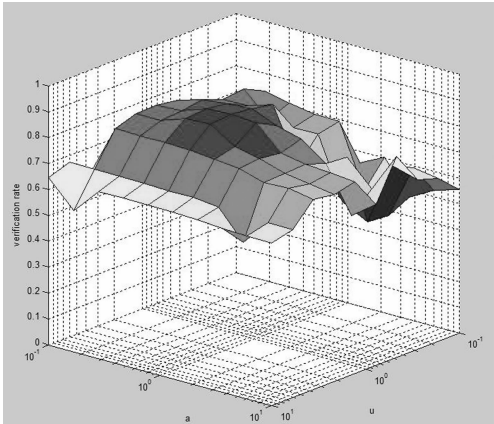


图 2 “funneled”数据集上不同 a, u 的分类性能
Fig. 2 Performance of face verification with different a, u on “funneled” dataset

由表 5 和图 4 可知,与 LDML^[9] 相比,本文提出的方法准确率提升了 12%,值得注意的是,本文提出基于“PCA+加权子空间”的特征投影,且利用带先验相似性与先验距离约束的相似度度量学习方法得到的度量矩阵,与 Sub-SM^[13] 方法相比,本文方法分类准确率提升了 1.6%. 与 DDML^[19] 相比,本文方法分类准确率也

有所提升. 特别在 $[0-0.12]$ 的 ROC 曲线区间, 本文方法的 ROC 曲线明显高于其他方法. 实验结果进一步说明了本文提出加权子空间的特征投影能有效降低样本对的类内变化, 建立的带先验相似性与先验距离约束的相似度量学习模型能有效提高特征向量的类内鲁棒性和类间判别性.

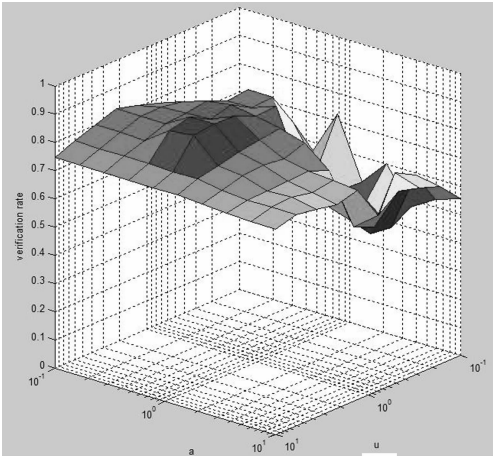


图 3 “Aligned”数据集上不同 a, u 的分类性能
Fig. 3 Performance of face verification with different a, u on “aligned” dataset

表 5 本文方法与其他人脸验证方法的结果对比
Tab. 5 Results of our proposed method with comparison to the other methods

方法	正确率
Combined b/g samples based methods, aligned ^[7]	0.868 3±0.003 4
LDML combined, funneled ^[9]	0.792 7±0.006 0
DML-eig combined, funneled + aligned ^[12]	0.856 5±0.005 6
HTBI Features, aligned ^[34]	0.881 3±0.005 8
CSML + SVM, aligned ^[14]	0.880 0±0.003 7
Sub-SML combined, funneled & aligned ^[13]	0.897 3±0.003 8
DDML combined, funneled & aligned ^[19]	0.906 8±0.014 1
本文方法, funneled + aligned	0.912 0±0.003 9

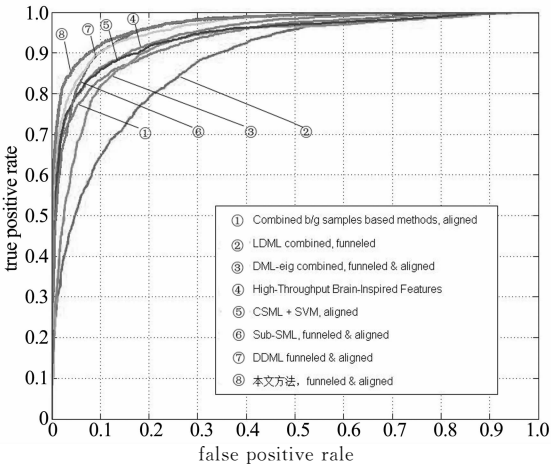


图 4 本文方法和其他人脸验证方法 ROC 特性曲线
Fig. 4 ROC curve of our proposed method with comparison to the other methods

4 结 论

针对无约束条件下的人脸验证问题, 本文提出了一种结合加权子空间和相似度量学习的人脸验证方法, 其主要思想在于抑制人脸样本对的类内变化及噪声的干扰, 同时增加类间距离. 首先, 引入人脸对类内样本的相似度信息作为权重, 提出一种加权子空间投影方法, 该方法可以获得具有鲁棒性的加权子空间, 利用“PCA+加权子空间”的特征投影方法能够有效降低因人脸表情、姿态、光照以及背景产生的较大类内变化对于后续分类的影响. 然后, 建立了带先验相似性和先验距离约束的相似度量学习模型, 该模型将先验相似性矩阵 G_0 和先验距离矩阵 M_0 作为相似度量学习的目标函数中的正则项约束, 更能准确衡量样本对的相似性和差异性先验信息, 优化后的度量矩阵更能有效提高特征向量的类内鲁棒性和类间判别性. 最后, 利用优化的度量矩阵获得人脸样本对的相似度得分. 从实验结果可知, 与其他同类的人脸验证方法相比, 本文方法具有更好的特征提取效果和分类性能, 在 LFW 的受限训练数据集上取得了 91.2% 的识别率, 取得了具有竞争性的效果.

参考文献

[1] XIANG Y, SU G. Multi-parts and multi-feature fusion in face verification[C]// Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE Computer Society Conference on. Anchorage AK US; IEEE Computer Society, 2008: 1–6.

[2] HUANG G B, LEE H, LEARNED-MILLER E. Learning hierarchical representations for face verification with convolutional deep belief networks[C]// Computer Vision and Pattern Recognition (CVPR), IEEE Conference on. Rhode Island, USA; IEEE Computer Society, 2012: 2518–2525.

[3] LI H, HUA G, LIN Z, *et al.* Probabilistic elastic matching for pose variant face verification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland Maine, USA; IEEE Computer Society, 2013: 3499–3506.

[4] HUANG R, LI T, YE M, *et al.* Unconstrained face verification by optimally organizing multiple classifiers[J]. International Journal of Control, Automation and Systems, 2014, 12 (4): 833–842.

[5] 沈媛媛, 严严, 王茜子. 有监督的距离度量学习算法研究进展[J]. 自动化学报, 2014, 40(12): 2673–2686.

SHEN Y Y, YAN Y, WANG H Z. Recent advances on supervised distance metric learning algorithms[J]. Acta Automatic Sinica, 2014, 40(12): 2673–2686. (In Chinese)

[6] WOLF L, HASSNER T, TAIGMAN Y. Similarity scores based on background samples[C]// Asian Conference on Computer Vision. Springer Berlin Heidelberg, 2009: 88–97.

[7] CHOPRA S, HADSELL R, LECUN Y. Learning a similarity

- metric discriminatively, with application to face verification [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). San Diego, CA, USA: IEEE Computer Society, 2005: 539—546.
- [8] LIU C, WECHSLER H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition[J]. IEEE Transactions on Image Processing, 2002, 11(4): 467—476.
- [9] GUILLAUMIN M, VERBEEK J, SCHMID C. Is that you? metric learning approaches for face identification[C]// IEEE 12th International Conference on Computer Vision. San Francisco, CA, USA: IEEE Computer Society, 2009: 498—505.
- [10] MIGNON A, JURIE F. Pcca: A new approach for distance learning from sparse pairwise constraints[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Rhode Island, USA: IEEE Computer Society, 2012: 2666—2672.
- [11] KÖSTINGER M, HIRZER M, WOHLHART P, *et al.* Large scale metric learning from equivalence constraints[C]// Computer Vision and Pattern Recognition (CVPR), IEEE Conference on. Rhode Island, USA: IEEE Computer Society, 2012: 2288—2295.
- [12] YING Y, LI P. Distance metric learning with eigenvalue optimization [J]. Journal of Machine Learning Research, 2012, 13(1): 1—26.
- [13] CAO Q, YING Y, LI P. Similarity metric learning for face recognition[C]// Proceedings of the IEEE International Conference on Computer Vision. Portland, Maine, USA: IEEE Computer Society, 2013: 2408—2415.
- [14] NGUYEN H V, BAI L. Cosine similarity metric learning for face verification[C]// Asian Conference on Computer Vision. Springer Berlin Heidelberg, 2010: 709—720.
- [15] FANG Y, CHEN Y Y S, WANG H, *et al.* Sparse similarity metric learning for kinship verification[C]// Visual Communications and Image Processing (VCIP). Chengdu, China: IEEE Computer Society, 2016: 1—4.
- [16] FANG Y, YAN Y, WANG H, *et al.* Adaptive metric learning with the low rank constraint[C]// Proceedings of the International Conference on Internet Multimedia Computing and Service. Xi'an, Chian: ACM, 2016: 61—65.
- [17] DING Z, SUH S, HAN J J, *et al.* Discriminative low-rank metric learning for face recognition[C]// Automatic Face and Gesture Recognition (FG), 11th IEEE International Conference and Workshops on. Ljubljana, Slovenia: IEEE Computer Society, 2015: 1—6.
- [18] HASSNER T, HAREL S, PAZ E, *et al.* Effective face frontalization in unconstrained images[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE Computer Society, 2015: 4295—4304.
- [19] HU J, LU J, TAN Y P. Discriminative deep metric learning for face verification in the wild[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE Computer Society, 2014: 1875—1882.
- [20] ZHU X, LEI Z, YAN J, *et al.* High-fidelity pose and expression normalization for face recognition in the wild[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE Computer Society, 2015: 787—796.
- [21] TAIGMAN Y, YANG M, RANZATO M A, *et al.* Deepface: Closing the gap to human—level performance in face verification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE Computer Society, 2014: 1701—1708.
- [22] SUN Y, WANG X, TANG X. Deep learning face representation from predicting 10,000 classes[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE Computer Society, 2014: 1891—1898.
- [23] SUN Y, CHEN Y, WANG X, *et al.* Deep learning face representation by joint identification—verification[C]// Advances in Neural Information Processing Systems. Palais des Congrès de Montréal, 2014: 1988—1996.
- [24] SUN Y, LIANG D, WANG X, *et al.* Deepid3: Face recognition with very deep neural networks[J]. arXiv preprint arXiv: 1502.00873, 2015.
- [25] LU J, PLATANIOU K N, VENETSANOPOULOS A N. Face recognition using LDA-based algorithms [J]. IEEE Transactions on Neural Networks, 2003, 14(1): 195—200.
- [26] MOGHADDAM B, JEBARA T, PENTLAND A. Bayesian face recognition[J]. Pattern Recognition, 2000, 33(11): 1771—1782.
- [27] PRICE J R, GEE T F. Face recognition using direct, weighted linear discriminant analysis and modular subspaces[J]. Pattern Recognition, 2005, 38(2): 209—219.
- [28] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971—987.
- [29] WOLF L, HASSNER T, TAIGMAN Y. Descriptor based methods in the wild[C]// Workshop on Faces in 'Real-life' Images: Detection, Alignment, and Recognition. Palais des Congrès Parc Chanot, 2008.
- [30] HUANG G B, MATTAR M, BERG T, *et al.* Labeled faces in the wild: A database for studying face recognition in unconstrained environments[C]// Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition. Palais des Congrès Parc Chanot, 2008.
- [31] TAIGMAN Y, WOLF L, HASSNER T. Multiple one-shots for utilizing class label information[C]// BMVC. London: 2009: 1—12.
- [32] BELHUMEUR P N, HESPANHA J P, KRIEGMAN D J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711—720.
- [33] http://www.robots.ox.ac.uk/~vgg/software/vgg_face/
- [34] COX D, PINTO N. Beyond simple features: A large—scale feature search approach to unconstrained face recognition [C]// IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG). Santa Barbara, CA, USA: IEEE Computer Society, 2011: 8—15.