

图嵌入投影非负矩阵分解图像特征提取方法

王娟 杜海顺 侯彦东 金勇

(河南大学图像处理与模式识别研究所 开封 475004)

摘要 针对投影非负矩阵分解(PNMF)不能揭示数据流形几何结构和判别信息的问题,提出了一种图嵌入投影非负矩阵分解(GEPNMF)特征提取方法。首先分别构建了描述数据流形几何结构和类间分离度的近邻图,然后采用它们的拉普拉斯矩阵设计了一个图嵌入正则项,并将其与PNMF的目标函数融合,以构造GEPNMF的目标函数。在GEPNMF目标函数中引入了图嵌入正则项,使求得的子空间能够在保持数据流形几何结构的同时,类间间距也最大。另外,还在目标函数中引入了一个正交正则项,以确保GEPNMF子空间基向量具有数据局部表示能力。对求解GEPNMF目标函数的累乘更新规则(MUR)进行了详细的推导。在Yale和CMU PIE人脸数据库上的实验结果表明,提出的图嵌入投影非负矩阵分解特征提取方法比PNMF更适用于解决分类问题。

关键词 人脸识别,特征提取,图嵌入,非负矩阵分解

中图分类号 TP391.4 文献标识码 A DOI 10.11896/j.issn.1002-137X.2014.08.066

Graph Embedding Projective Non-negative Matrix Factorization Method for Image Feature Extraction

WANG Juan DU Hai-shun HOU Yan-dong JIN Yong

(Institute of Image Processing and Pattern Recognition, Henan University, Kaifeng 475004, China)

Abstract To overcome the disadvantage that projective non-negative matrix factorization (PNMF) fails to discover the intrinsic geometrical and discriminating structure, a novel graph embedding projective non-negative matrix factorization (GEPNMF) was proposed for image feature extraction. The paper constructed two adjacent graphs that are separately used to characterize the intrinsic geometrical structure of data and interclass separability. Using the Laplacian matrices of the adjacent graphs, the paper designed a graph embedding regularization that incorporates with PNMF's objective function to construct the GEPNMF's objective function. Since the graph embedding regularization is adopted by the objective function, the learned subspace of GEPNMF can preserve the data geometrical structure while it maximizes the margins between different classes. That is to say, it has more discriminability. In addition, the paper introduced an orthogonal regularization into the objective function to ensure the learned bases to be parts-based. The paper deduced a multiplicative update rule (MUR) to optimize the objective function. The experimental results on Yale and CMU PIE face image datasets suggest the effectiveness of GEPNMF.

Keywords Face recognition, Feature extraction, Graph embedding, Non-negative matrix factorization

1 引言

近年来,非负矩阵分解(NMF)^[1]在图像处理、模式识别等领域得到了广泛的应用。由于对基向量矩阵和系数矩阵增加了非负约束,NMF能将一个数据表示为一组非负基向量只加不减的线性组合。从这个层面来说,NMF学习得到的基向量具有一定的数据局部表示能力^[2]。然而,NMF除了非负约束之外,并没有其它约束明确地保证其基向量具有数据局部表示能力。文献[3]对ORL人脸数据库中的人脸图像采用非负矩阵分解后发现,NMF学习得到的基向量并不能很好地表示人脸图像的局部特征。为了确保NMF具有数据局部表示能力,文献[3]在NMF的目标函数中引入了局部化约束,提

出了局部非负分解(LNMF)方法;文献[4]将稀疏性约束融合到NMF的目标函数中,提出了非负稀疏编码(NSC)方法。

NMF虽然具有一定的数据局部表示能力,但不能揭示数据的流形几何结构,而数据的流形几何结构在解决实际问题时的作用至关重要。文献[5]和文献[6]采用最近邻图表示数据的流形几何结构,并将其相关信息融入到NMF的目标函数中,分别提出了图非负矩阵分解(GNMF)方法和图保持稀疏非负矩阵分解(GSNMF)方法。为使NMF更适用于解决解决分类问题,文献[7]在NMF的目标函数中引入Fisher判别信息,提出判别非负矩阵分解(DNMF)方法;文献[8]将NMF与流形嵌入模型^[9]结合起来,提出了非负图嵌入(NGE)方法;文献[10]将局部判别嵌入模型^[11]引入到NMF

到稿日期:2013-09-13 返修日期:2014-03-21 本文受国家自然科学基金(U1204611),河南省基础与前沿技术研究计划项目(132300410474),河南省教育厅科学技术重点研究项目(12A520008)资助。

王娟(1977-),女,硕士,讲师,主要研究方向为图像处理、模式识别、数据融合等,E-mail:wj2001@henu.edu.cn;杜海顺(1977-),男,博士,副教授,硕士生导师,主要研究方向为图像处理、模式识别、数据融合等。

的目标函数中,提出一种流形正则化判别非负矩阵分解(MD-NMF)方法。

由于数据映射到基向量张成的子空间并不能直接得到系数向量,因此,NMF及上述扩展方法都不是严格意义上的子空间学习方法^[12-15]。为解决这一问题,文献[16]提出了投影非负矩阵分解(PNMF)方法,通过稍微改变NMF的目标函数,最终学习得到一个非负投影子空间。然而,与NMF一样,PNMF也没有考虑数据的流形几何结构和判别信息,因此它更适用于解决数据聚类问题而不是数据分类问题。为了得到一个能够揭示数据流形几何结构且更具判别力的子空间,本文将图嵌入模型与PNMF融合在一起,提出了一种图嵌入投影非负矩阵分解(GEPNMF)特征提取方法。由于GEPNMF在目标函数中引入了图嵌入正则项,其求得的子空间能够在保持数据流形几何结构的同时,类间间距也最大。因此,与PNMF相比,GEPNMF更适用于解决数据分类问题。

2 非负矩阵分解及投影非负矩阵分解概述

给定非负数据矩阵 $X = [x_1, \dots, x_n] \in \mathbb{R}^{m \times n}$, NMF^[1]的目的是寻找两个非负低秩矩阵 $W \in \mathbb{R}^{m \times r}$ 和 $H \in \mathbb{R}^{r \times n}$ 来逼近数据矩阵 X , 即 $X \approx WH$, 其中 $r < \min\{m, n\}$, $W = [w_1, w_2, \dots, w_r]$ 为基向量矩阵, $H = [h_1, \dots, h_n]$ 为系数向量矩阵。NMF采用欧氏距离衡量 WH 对 X 的逼近程度, 给出的目标函数如下:

$$J(W) = \min_{W \geq 0, H \geq 0} \|X - WH\|_F^2 \quad (1)$$

由目标函数式(1)可以看出,同时求解 W 和 H 是一个非凸问题,不仅得不到其全局最优解,而且求解困难。然而,固定 W 和 H 中的一个,求解另外一个,却是一个凸问题。因此,NMF采用迭代最优优化方法得到目标函数的最优解。Lee和Seung^[17]给出了 W 和 H 的累乘更新公式(Multiplicative Update Rule, MUR):

$$H_{au} \leftarrow H_{au} \frac{(W^T X)_{au}}{(W^T W H)_{au}} \quad (2)$$

$$W_{ia} \leftarrow W_{ia} \frac{(X H^T)_{ia}}{(W H H^T)_{ia}}$$

然而,NMF不是一个严格意义上的子空间学习方法,直接处理新的测试数据缺少必要的理论依据。为解决这一问题,Yang等人^[16]提出了投影非负矩阵分解(PNMF)方法。PNMF试图寻找一个低秩非负投影矩阵 $P \in \mathbb{R}^{m \times m}$, 使得 $X \approx PX$, 其中 $P = WW^T$, $W \in \mathbb{R}^{m \times r}$ 。与NMF相比,PNMF仅仅用 $W^T X$ 代替 H , 但这却使PNMF成为了一种子空间学习方法。新的数据向 W 张成的低维子空间投影,就能得到其低维表示。PNMF方法同样采用欧氏距离衡量 PX 对 X 的逼近程度,其目标函数如下:

$$J(W) = \min_{W \geq 0} \frac{1}{2} \|X - WW^T X\|_F^2 \quad (3)$$

文献[16]给出了 W 的MUR:

$$W_{*k} \leftarrow W_{*k} \frac{2(XX^T W)_{*k}}{(WW^T XX^T W)_{*k} + (XX^T WW^T W)_{*k}} \quad (4)$$

3 图嵌入正则化投影非负矩阵分解

3.1 目标函数

从子空间学习的角度来看,PNMF是将数据 $x_i \in \mathbb{R}^m$ 投影到基向量矩阵 W 张成的子空间,得到其低维数据表示。然而,PNMF没有考虑数据的流形几何结构和判别信息,因此

它更适用于解决聚类问题而不是分类问题。

假设数据集 $X = [x_1, \dots, x_n] \in \mathbb{R}^{m \times n}$ 是从嵌入到高维空间的一个低维流形 M 上采样得到的,流形学习算法采用 X 的最近邻图来描述流形 M 的流形几何结构。具体地,以 X 中的样本为顶点,以 S 为权重矩阵,创建了一个带权无向图 G , 其中权重定义如下:

$$S_{ij} = \begin{cases} 1, & \text{if } x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i) \\ 0, & \text{otherwise} \end{cases}$$

式中, $N_k(x_i)$ 表示数据 x_i 的 k 近邻集合。数据集 $X = [x_1, \dots, x_n] \in \mathbb{R}^{m \times n}$ 投影到子空间后,得到其低维表示 $H = [h_1, \dots, h_n] \in \mathbb{R}^{r \times n}$ 。为了使 H 保持 X 的流形几何结构,应使 $\text{tr}(HLH^T)$ 极小化。其中 $L = D - S$ 为图 G 的Laplacian矩阵,矩阵 D 为对角阵, $D_{ii} = \sum_j S_{ij}$ 。

为了在GEPNMF目标函数中融入数据流形几何结构和判别信息,与MFA^[9]、LDE^[11]类似,本文首先分别构建了描述数据流形几何结构和类间分离度的近邻图。为方便描述,训练样本 x_i 的类别标签记为 $c_i \in \{1, 2, \dots, N_c\}$, 其中 N_c 为训练样本的类别数。记第 c 类训练样本集合为 π_c , 第 c 类训练样本个数为 n_c 。在构造描述数据流形几何结构的最近邻图 G^p 时,对每个数据样本 x_i , 本文只让 x_i 与它同类的 k_1 个最近邻样本连接,并按照以下方式计算其权重:

$$S_{ij}^c = \begin{cases} 1, & \text{if } x_i \in N_{k_1}^+(x_j) \text{ or } x_j \in N_{k_1}^+(x_i) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

式中, $N_{k_1}^+(x_i)$ 表示与数据 x_i 同类的 k_1 个近邻组成的集合。为保证投影后的数据集 H 保持数据集 X 的流形几何结构,基于 G^p 图的Laplacian矩阵 $L_p = D_p - S_p$, 应使 $\text{tr}(HL_p H^T)$ 最小。

在构造描述类间分离度的最近邻图 G^c 时,本文考虑不同类数据间的近邻关系,并计算权重如下:

$$S_{ij}^c = \begin{cases} 1, & \text{if } (x_i, x_j) \in P_{k_2}(c_i) \text{ or } (x_i, x_j) \in P_{k_2}(c_j) \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

式中, $P_{k_2}(c_i)$ 表示样本对集合 $\{(x_i, x_j) | x_i \in \pi_{c_i}, x_j \notin \pi_{c_i}\}$ 的 k_2 个最近邻样本对子集。为了使投影后的数据集类间间距最大,基于 G^c 图的Laplacian矩阵 $L_c = D_c - S_c$, 应使 $\text{tr}(HL_c H^T)$ 最大。综上所述,本文给出如下图嵌入正则项^[10]:

$$\text{tr}(H(L_c^{-1/2})^T L_p L_c^{-1/2} H^T) \quad (7)$$

为确保 L_c 的正定性,本文给它的对角线元素加上微小的扰动,即 $\tilde{L}_c = L_c + \zeta I$, 本文 ζ 取 $10^{-4} \text{tr}(L_c)$ 。在本文的后续描述中,仍用记号 L_c 表示加扰动后的 \tilde{L}_c 。

将式(7)与PNMF的目标函数融合,并令 $H = W^T X$, 可以得到:

$$\min_{W \geq 0} \frac{1}{2} \|X - WW^T X\|_F^2 + \frac{\beta}{2} \text{tr}(W^T X(L_c^{-1/2})^T L_p L_c^{-1/2} X^T W) \quad (8)$$

式中, β 为正则项平衡参数。为确保学习得到的基向量具有更强的局部表示能力,本文在式(8)的基础上增加正交约束^[18]:

$$\sum_{i=1}^r w_i^T w_j = \text{tr}(WEW^T) \quad (9)$$

式中, E 表示对角线上元素为0、其他元素都为1的矩阵。

由式(8)和式(9),可得GEPNMF的目标函数如下:

$$J_F(W) = \min_{W \geq 0} \frac{1}{2} \|X - WW^T X\|_F^2 + \frac{\alpha}{2} \text{tr}(WEW^T) +$$

$$\frac{\beta}{2} \text{tr}(W^T X (L_c^{-1/2})^T L_p L_c^{-1/2} X^T W) \quad (10)$$

式中, α, β 为正则项平衡参数。

3.2 累乘更新规则(MUR)

为了求解使目标函数式(10)极小化的基向量矩阵 W , 本文首先对 W 的累乘更新规则进行推导。令:

$$J_{F1}(W) = \frac{1}{2} \|X - WW^T X\|_F^2 \quad (11)$$

$$J_{F2}(W) = \frac{\alpha}{2} \text{tr}(WEW^T) + \frac{\beta}{2} \text{tr}(W^T X (L_c^{-1/2})^T L_p L_c^{-1/2} X^T W) \quad (12)$$

由此可知 $J_F(W) = J_{F1}(W) + J_{F2}(W)$ 。记 $D \triangleq (L_c^{-1/2})^T D_p L_c^{-1/2}$, $S \triangleq (L_c^{-1/2})^T S_p L_c^{-1/2}$, 容易证明矩阵 D 和 S 非负, 具体参考文献[10]。记 $L \triangleq (L_c^{-1/2})^T L_p L_c^{-1/2} = D - S$ 。 $J_{F1}(W)$ 、 $J_{F2}(W)$ 对 W_{ab} 的一阶偏导为:

$$\frac{\partial J_{F1}(W)}{\partial W_{ab}} = -2(XX^T W)_{ab} + (WW^T XX^T W)_{ab} + (XX^T WW^T W)_{ab} \quad (13)$$

$$\frac{\partial J_{F2}(W)}{\partial W_{ab}} = \alpha(WE)_{ab} + \beta(XLX^T W)_{ab} \quad (14)$$

由上可得目标函数 $J_F(W)$ 对应 W_{ab} 的梯度为:

$$\frac{\partial J_F(W)}{\partial W_{ab}} = -2(XX^T W)_{ab} + (WW^T XX^T W)_{ab} + (XX^T WW^T W)_{ab} + \alpha(WE)_{ab} + \beta(XLX^T W)_{ab} \quad (15)$$

通过简单的代数运算, 式(15)的梯度可分解成正、负两部分, 即:

$$\nabla_{W_{ab}} = \nabla_{W_{ab}}^+ - \nabla_{W_{ab}}^- \quad (16)$$

其中,

$$\nabla_{W_{ab}}^+ = (WW^T XX^T W)_{ab} + (XX^T WW^T W)_{ab} + \alpha(WE)_{ab} + \beta(XDX^T W)_{ab} \quad (17)$$

$$\nabla_{W_{ab}}^- = 2(XX^T W)_{ab} + \beta(XSX^T W)_{ab} \quad (18)$$

在每一迭代步骤 t , 本文沿着负梯度方向更新矩阵 W , 即

$$W_{ab}^{t+1} = W_{ab}^t - \eta_{ab}^t (\nabla_{W_{ab}}^+ - \nabla_{W_{ab}}^-) \quad (19)$$

为保证每步迭代时 W 的非负性, 选择 $\eta_{ab}^t = \frac{W_{ab}^t}{\nabla_{W_{ab}}^+}$, 则

$$W_{ab}^{t+1} = W_{ab}^t \frac{\nabla_{W_{ab}}^-}{\nabla_{W_{ab}}^+} \quad (20)$$

可得 W 的 MUR 为:

$$W_{ab}^{t+1} = W_{ab}^t \frac{(2XX^T W^t + \beta XSX^T W^t)_{ab}}{(W^t W^{tT} XX^T W^t + XX^T W^t W^{tT} W^t + \alpha W^t E + \beta XDX^T W^t)_{ab}} \quad (21)$$

3.3 GEPNMF 求解及特征提取算法

假设已给定训练样本集 $X = [x_1, \dots, x_n] \in \mathbb{R}^{m \times n}$, 并按照前述公式 $D \triangleq (L_c^{-1/2})^T D_p L_c^{-1/2}$, $S \triangleq (L_c^{-1/2})^T S_p L_c^{-1/2}$ 计算得到 D 和 S 。本文给出 GEPNMF 计算基向量矩阵 W 的具体过程, 如算法 1 所示。

算法 1 GEPNMF 求解基向量矩阵 W 算法

输入: $X \in \mathbb{R}^{m \times n}$, D, S , $1 \leq r \leq \min\{m, n\}$

输出: $W \in \mathbb{R}^{m \times r}$

1. 初始化 $W^0, t=0$ 。

重复

2. 1 采用式(21)更新 W^{t+1} ;

2. 2 $t=t+1$ 。

直到达到最大迭代次数或 $|J_F(W^{t+1}) - J_F(W^t)| < \epsilon$ 。

GEPNMF 按算法 1 求解得到基向量矩阵 W 后, 则有数据 $x \in \mathbb{R}^m$ 在 W 张成的低维子空间中表示为 $y \in \mathbb{R}^r$, 即 $x \approx Wy$ 。由此, 对原始空间中的任一数据 x , 采用下式可得其低维数据特征 y :

$$y = W^\dagger x \quad (22)$$

式中, $W^\dagger = (W^T W)^{-1} W^T$, 为矩阵 W 的 pseudo 逆。

4 实验

为了评价 GEPNMF 特征提取方法的性能, 本文在 Yale 和 CMU PIE 人脸数据库上, 分别采用 GEPNMF 和 NMF^[1]、LNMF^[3]、PNMF^[16]、GNMF^[5]、MD_NMF^[10] 等方法提取人脸图像特征并用其进行人脸识别。实验时, 对任一人脸数据库, 本文从每个人所有图像中随机选择 k 幅图像构成训练样本集, 剩余图像作为测试样本集。训练样本集用于学习低维空间基向量矩阵 W , 然后采用 W 提取训练样本集和测试样本集中数据的特征, 并采用最近邻分类器 (Nearest Neighbor Classifier) 对测试样本集中的数据进行分类识别。为了减少随机选择样本对实验结果的影响, 对任一训练样本数 k , 上述实验过程独立进行 5 次, 最后取 5 次实验的平均识别率作为最终识别率。

4.1 Yale 数据库

Yale 数据库^[19] 包含 15 个人, 每人 11 张, 共 165 张人脸图像。每张图像为 256 个灰度级, 分辨率为 243×320 。该数据库中的人脸图像光照、表情以及拍摄角度都有较大程度的变化。本文对 Yale 数据库中的每张图像都以眼睛为中心进行了剪切和对齐处理, 并将其大小缩放为 32×32 像素。Yale 数据库中的部分人脸图像如图 1(a) 所示。在 Yale 数据库上, 本文分别从每个人所有图像中随机选择 3、5、7 张图像构成训练样本集进行人脸识别实验。图 2 给出了 GEPNMF 与其它 5 种方法随特征维数变化的识别率曲线比较。表 1 给出了 GEPNMF 与其它 5 种方法最高识别率及相应特征维数比较。由图 2 和表 1 可以看出, 在 Yale 人脸数据库上, 当从每个人所有图像中随机选择 3 张和 5 张图像构成训练样本集时, GEPNMF 提取的特征识别率远远高于其它 5 种特征提取方法; 当从每个人所有图像中随机选择 7 张图像构成训练样本集时, GEPNMF 提取的特征识别率仍远高于 NMF、GNMF、MD_NMF, 与 LNMF、PNMF 相比, GEPNMF 提取的特征识别率也较高。

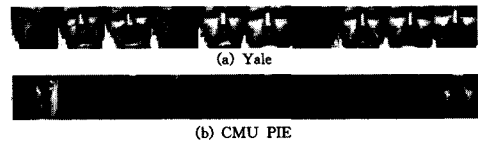
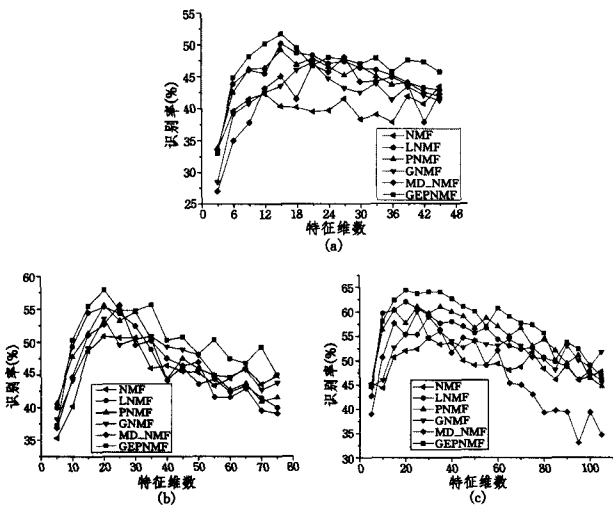


图 1 两个人脸数据库部分人脸图像

表 1 Yale 数据库上不同特征提取方法最高人脸识别率(%)比较

算法	3	5	7
NMF	43.33(45)	50.89(20)	54.67(30)
LNMF	50.17(15)	55.33(20)	62.00(20)
PNMF	49.17(15)	55.56(20)	61.00(25)
GNMF	47.17(21)	53.56(20)	60.00(25)
MD_NMF	48.00(27)	55.56(25)	59.00(30)
GEPNMF	51.67(15)	57.89(20)	64.33(20)

注: 括号中的数字代表该识别率所对应的特征维数



本文从每个人所有图像中随机选择(a) 3、(b) 5、(c) 7张图像构成训练样本集,其余图像构成测试样本集。实验独立进行5次,最后取5次实验的平均识别率为最终识别率

图2 Yale数据库上不同特征提取方法人脸识别率曲线比较

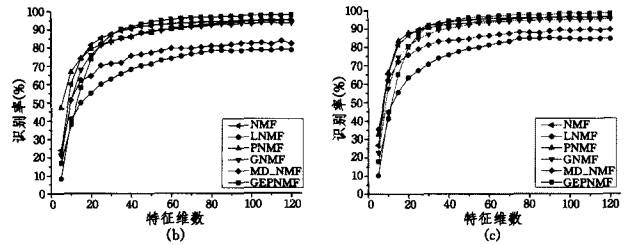
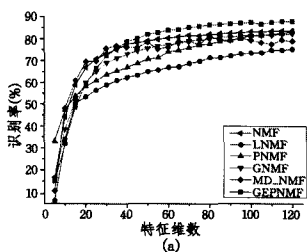
4.2 CMU PIE

CMU PIE 数据库^[20]共有68个人41368张图像,其中每个人的图像有13种姿态条件、43种光照条件和4种表情。本文选择姿态C27中21张不同光照下的图像构成CMU PIE数据库子集作为实验数据库。对CMU PIE数据库子集中的图像,本文以眼睛为中心进行了剪切和对齐处理,将其大小缩放为 32×32 像素。CMU PIE数据库子集中的部分人脸图像如图1(b)所示。在CMU PIE数据库子集上,本文分别从每个人所有图像中随机选择5、10、15张图像构成训练样本集进行人脸识别实验。图3给出了GEPNMF与其它5种方法随特征维数变化的识别率曲线比较。表2给出了GEPNMF与其它5种方法最高识别率及相应特征维数比较。由图3和表2可以看出,在CMU PIE数据库子集上,GEPNMF和NMF、GNMF、MD_NMF提取的特征都达到了比较高的识别率,其中GEPNMF的识别率最高;LNMF和PNMF提取的特征识别率远低于GEPNMF方法。

表2 CMU PIE数据库上不同特征提取方法最高人脸识别率(%)比较

算法	5	10	15
NMF	83.62(115)	95.20(120)	96.41(100)
LNMF	75.22(120)	78.95(115)	84.91(90)
PNMF	83.57(120)	95.00(120)	96.65(100)
GNMF	82.35(120)	93.58(110)	95.33(115)
MD_NMF	81.53(80)	83.48(115)	89.56(110)
GEPNMF	87.99(120)	97.80(120)	98.58(100)

注:括号中的数字代表该识别率所对应的特征维数



本文从每个人所有图像中随机选择(a) 5、(b) 10、(c) 15张图像构成训练样本集,其余图像构成测试样本集。实验独立进行5次,最后取5次实验的平均识别率为最终识别率

图3 CMU PIE数据库上不同特征提取方法人脸识别率曲线比较

结束语 本文提出的图嵌入投影非负矩阵分解(GEPNMF)方法通过将图嵌入模型与投影非负矩阵分解(PNMF)融合,使GEPNMF子空间在保持数据流形几何结构的同时,更具判别力。因此,与PNMF相比,GEPNMF更加适合用来解决数据的分类问题。具体地,本文首先分别构建了描述数据流形几何结构和类间分离度的近邻图;然后采用两近邻图的Laplacian矩阵,设计了一个既能保持数据流形几何结构,又能使不同类数据间间距最大化的图嵌入正则项,并将其与PNMF目标函数融合,建立了GEPNMF的目标函数。另外,本文还在GEPNMF的目标函数中引入了正交约束正则项,以确保GEPNMF学习得到的子空间基向量具有数据局部表示能力。为求解GEPNMF的目标函数,本文推导出基向量矩阵 W 的累乘更新规则(MUR)。最后,在ORL、Yale、CMU PIE人脸数据库上的大量实验结果表明,与NMF、LNMF、PNMF、GNMF和MD_NMF相比,GEPNMF提取的图像特征在用于人脸识别时,具有更高的识别率。

参考文献

- [1] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature, 1999, 401(6755): 788-791
- [2] 赵海峰, 李其超, 罗斌, 等. 基于局部Walsh变换和非负矩阵分解的脑白质图像分割[J]. 光电子·激光, 2012, 23(7): 1425-1430
- [3] Li S Z, Hou X, Zhang H, et al. Learning spatially localized, parts-based representation[C]// Proceedings of IEEE International conference on Computer Vision and Pattern Recognition, 2001: 207-212
- [4] Hoyer P O. Non-negative matrix factorization with sparseness constraints[J]. Journal of Machine Learning Research, 2004(5): 1457-1469
- [5] Cai D, He X F, Han J W, et al. Graph Regularized Non-negative matrix factorization for Data Representation[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1548-1560
- [6] Zhi R C, Flierl M, Ruan Q Q, et al. Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition[J]. IEEE Trans. on Systems, Man, and Cybernetics-PART B: Cybernetics, 2011, 41(1): 38-52
- [7] Zafeiriou S, Tefas A, Buciu I, et al. Exploiting discriminant information in nonnegative matrix factorization with application to frontal face verification [J]. IEEE Trans. on Neural Networks, 2006, 17(3): 683-695

- [8] Yang J, Yan S, Fu Y, et al. Non-negative graph embedding [C]// Proceedings of IEEE International conference on Computer Vision and Pattern Recognition, 2008:1-8
- [9] Yan S, Xu D, Zhang B, et al. Graph embedding and extensions: a general framework for dimensionality reduction [J]. IEEE Trans. on Pattern analysis and machine intelligence, 2007, 29(1):40-51
- [10] Guan N, Tao D, Luo Z, et al. Manifold regularized discriminative nonnegative matrix factorization with fast gradient descent [J]. IEEE Trans. on Image Processing, 2011, 20(7):2030-2048
- [11] Chen H T, Chang H W, Liu T L. Local discriminant embedding and its variants [J]. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2005(2):846-853
- [12] Ding C, Li T, Jordan M I. Convex and semi-nonnegative matrix factorizations [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2010, 32(1):45-55
- [13] 方蔚涛, 马鹏, 成正斌, 等. 二维投影非负矩阵分解算法及其在人脸识别中的应用[J]. 自动化学报, 2012, 38(9):1503-1512
- [14] 刘雪松, 王斌, 张立明. 基于非负矩阵分解的高光谱遥感图像混合像元分解[J]. 红外与毫米波学报, 2011, 30(1):1-5
- [15] 于红芸, 姜涛, 关键. SAR 图像目标检测的互信息非负矩阵分解算法[J]. 中国图象图形学报, 2011, 16(1):129-134
- [16] Yang Z, Oja E. Linear and nonlinear projective nonnegative matrix factorization [J]. IEEE Trans. on Neural Networks, 2010, 21(5):734-749
- [17] Lee D D, Seung H S. Algorithms for non-negative matrix factorization [J]. Advances in Neural Information Processing Systems, MIT Press, 2001, 13:556-562
- [18] Ding C, Li T, Peng W, et al. Orthogonal nonnegative matrix tri-factorizations for clustering [C]// Proceedings of 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2006:126-135
- [19] Georgiades A S, Belhumeur P N, Kriegman D J, et al. From few to many: illumination cone models for face recognition under variable lighting and pose [J]. IEEE Trans. Pattern Anal. Mach. Intelligence, 2001, 23(6):643-660
- [20] Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression database[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2003, 25(12):1615-1618

(上接第 288 页)

帧、16 个镜头。通过 K-mean 聚类关键帧提取算法提取的关键帧如图 9 所示。通过代表性关键帧算法提取的关键帧如图 10 所示。



图 9 K-mean 在神舟七号的新闻报道视频上提取的关键帧



图 10 本文算法在神舟七号的新闻报道视频上提取的关键帧

这段视频同样是富含短镜头的新闻报道。在这段视频中,我们的算法可以提取出更精简的极具代表性的关键帧。

从实验中我们可以看到,我们的算法不拘泥于按照镜头提取关键帧的方法,而是直接针对画面的代表性提取关键帧。这使我们的算法提取的关键帧不会随着视频的镜头数量增加而增加,使提取的关键帧具有真正代表视频内容的功能。在网络视频监控应用中,海量的视频信息需要通过最有效的方式进行摘要。这种需求使本文提出的算法可以在网监应用中更好地发挥作用。

结束语 我们在这篇文章中介绍了一种通过计算帧的代表性并基于代表性评价而实现的关键帧提取算法。在计算代表性的过程中我们提出了一种计算帧间相关度的算法,并利用了一种改进的 PageRank 算法完成代表性的计算。通过与传统的基于聚类的关键帧提取算法在 4 段视频上获得的关键帧进行对比,验证了我们的算法能够通过更少的极具代表性的关键帧对视频进行摘要。实验结果验证了算法的有效性。

参 考 文 献

- [1] Zhuang Y, Rui Y, Huang T S. Video key frame extraction by unsupervised clustering and feedback adjustment[J]. Journal of Computer Science and Technology, 1999, 14(3):283-287
- [2] Zhao L, Qi W, Li S Z, et al. Key-frame extraction and shot retrieval using nearest feature line (NFL)[C]// Proceedings of the 2000 ACM workshops on Multimedia. New York: ACM, 2000: 217-220
- [3] Pan R, Tian Y, Wang Z. Key-Frame Extraction Algorithm Based on Entropy[C]// 2010 International Conference on E-Product E-Service and E-Entertainment (ICEEE). IEEE, 2010: 1-4
- [4] Liu T, Zhang H J, Qi F. A novel video key-frame-extraction algorithm based on perceived motion energy model [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(10):1006-1013
- [5] Li C, Wu Y T, Yu S S, et al. Motion-focusing key frame extraction and video summarization for lane surveillance system[C]// 2009 16th IEEE International Conference on Image Processing (ICIP). Cairo: IEEE, 2009: 4329-4332
- [6] Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine [J]. Computer Networks and ISDN Systems, 1998, 30(1):107-117
- [7] Forsen P E. Maximally stable colour regions for recognition and matching[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2007(CVPR'07). IEEE, 2007: 1-8
- [8] Matas J, Chum O, Urban M, et al. Robust wide-baseline stereo from maximally stable extremal regions[J]. Image and Vision Computing, 2004, 22(10):761-767
- [9] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6): 381-395