

基于多视图核鉴别分析的图像识别

朱震宇,荆晓远

(南京邮电大学 自动化学院,江苏 南京 210003)

摘要:近年来多视图学习引起了研究者的广泛关注。在多视图学习中,数据主要来自于多个视图(或特征集)。多视图数据的最大优点是可以从不同视图之间提取互补信息。传统多视图学习方法是在不同视图上单独地训练分类器。这些方法利用了视图之间的互补信息,但是忽略了去除不同视图之间的冗余信息。为了解决上述问题,提出一种基于多视图核鉴别分析的识别方法。该方法通过基于核判别分析从各个视图中提取出相互正交的投影矩阵,从而能够提取出兼具互补和无冗余的特征。在 AR 和 Oxford Flowers17 公共数据库上的实验结果验证了所提算法的有效性。

关键词:多视图学习;互补信息;冗余信息;核鉴别分析

中图分类号:TP181

文献标识码:A

文章编号:1673-629X(2016)12-0092-04

doi:10.3969/j.issn.1673-629X.2016.12.020

Image Recognition Based on Multi-view Kernel Discriminant Analysis

ZHU Zhen-yu, JING Xiao-yuan

(College of Automation, Nanjing University of Posts and Telecommunications,
Nanjing 210003, China)

Abstract: Multi-view learning has caused wide public concern of researchers in recent years. In multi-view learning, data is mainly from many views (or feature set). The biggest advantage of multi-view data is that it can extract complementary information from different views. The traditional multi-view learning method learns classifiers in different views independently. These methods utilize the complementary information between views, but ignore the redundant information between different views. In order to solve the above problem, a recognition method based on multi view kernel discriminant analysis is proposed. It uses kernel discriminant analysis to extract projection matrix from each view and makes the transformations orthogonal, so that it can extract both complementary and non-redundant features. Experimental results on public database like AR and Oxford Flowers17 verify the effectiveness of the algorithm proposed.

Key words: multi-view learning; complementary information; redundant features; kernel discriminant analysis

0 引言

近年来,随着数据获取技术的不断发展,多视图学习已经在计算机视觉领域得到了广泛运用。例如,在传感器网络中,数据集来源于多个传感器,当前传感器缺失的数据可以通过其他不同的传感器弥补,因此通过多个传感器计算分类的效果要好于单一传感器^[1-2]。每个网页可以根据其所含文档和指向它的超链接进行表示;一个人可以根据与其相关的音频和视频内容来表示;一幅图像可由不同的特征来表示(如颜色和形状等)。这些不同视图包含着大量的特征数据,如何从各个视图获取更多信息,多视图学习已经成为机器学习的研究热点之一。

大部分关于多视图分类方法是基于特征融合(前期融合)的方法,然后找到一个综合的节点来表示多视图特征^[3-4];另一种是基于最终结果的融合(后期融合),通过结合了各个视图的结果来提升最终的表现^[5]。除非多视图融合的特征是进行优化过的,否则很难保证特征融合能带来好的分类结果。

Tang 等提出基于多视图的特征选择算法—MVFS^[6],分别对每个视图进行独立的特征选择,并通过谱分析对每个视图进行约束,使多视图学习满足一致性原则。该算法考虑各个视图之间的相互关系,但没有去除不同视图之间的冗余信息。Jing 等提出了基于鉴别变换的整体正交彩色图像识别方法—HOA^[7]。

收稿日期:2016-02-01

修回日期:2016-06-09

网络出版时间:2016-11-21

基金项目:国家自然科学基金资助项目(61272273)

作者简介:朱震宇(1989-),男,研究生,研究方向为生物特征识别;荆晓远,教授,博士生导师,研究方向为模式识别、图像与信号处理、信息安全、机器学习与数据挖掘。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20161121.1641.028.html>

该方法通过线性鉴别分析^[8-9] (LDA) 与整体正交分析相结合, 根据 Fisher 准则从彩色图像中分别抽取红、绿、蓝三种色彩的判别变换矩阵, 并使之相互正交。但该方法在选择正交时直接按照固定的正交顺序, 从而忽略了各个视图对最终识别效果的影响。

文中提出了基于多视图核鉴别的鉴别分析算法 (MVKDA), 利用核鉴别分析 (KDA)^[10-11] 得到各个视图对应的投影矩阵, 并使之按最优化的顺序进行相互正交。核鉴别分析是目前研究最多的非线性鉴别技术, 它将基于核函数的非线性映射方法与线性鉴别分析相结合, 将样本从原始特征空间映射到高维的核空间以提高数据的可分性。文中提出的方法提高了数据的可分性, 去除视图特征之间的冗余信息, 从而提高算法分类效果。在 AR^[12] 和 Oxford Flowers 17^[13] 公共数据库上的实验结果验证了所提算法的有效性。

1 核鉴别分析 (KDA)

假设样本矩阵是 $X = [x_1, x_2, \dots, x_N]$, x_i 表示 X 的第 i 个样本, 样本总数为 N , l_i 表示第 i 类样本的样本数。通过函数 φ , 将样本投影到新的空间 F 。假设数据一共有 c 个类别, 多类 KDA 的作用就是通过使用 $c-1$ 个判别函数将数据投影到 $c-1$ 维空间上。

$$y_i = w_i^T \varphi(x), i = 1, 2, \dots, c-1 \quad (1)$$

其中, w_i 是 W 的第 i 列数据。

类间协方差矩阵可以表示为:

$$S_B^\varphi = \sum_{i=1}^c l_i (m_i^\varphi - m^\varphi) (m_i^\varphi - m^\varphi)^T \quad (2)$$

其中, m_i^φ 表示第 i 类样本在新的投影空间内的均值; m^φ 表示所有样本在新的投影空间内的样本均值。

类内协方差矩阵可以表示成:

$$S_W^\varphi = \sum_{i=1}^c \sum_{n=1}^{l_i} (\varphi(x_n^i) - m_i^\varphi) (\varphi(x_n^i) - m_i^\varphi)^T \quad (3)$$

其中, x_n^i 表示 X 的第 i 类第 n 个样本。

基于费歇尔准则的目标函数最大化:

$$J(W) = \frac{|W^T S_B^\varphi W|}{|W^T S_W^\varphi W|} \quad (4)$$

根据再生核理论, 多类 KDA 的目标结果可以转化为:

$$A^* = \operatorname{argmax}_A \frac{|A^T S_B^\varphi A|}{|A^T S_W^\varphi A|} \quad (5)$$

根据 $S_W^\varphi \alpha = \lambda S_B^\varphi \alpha$, 求出 $S_W^{\varphi^{-1}} S_B^\varphi$ 的最大 $c-1$ 个特征对应的特征向量组成的 A^* , 这样输入一个向量 x_i , 通过投影矩阵 A^* , 输出:

$$y(x_i) = (A^*)^T K_i \quad (6)$$

其中, K_i 的第 i 个值表示成核函数 $k(x_i, x_i)$ 。

这样就完成了样本的降维, 将输入样本通过 KDA

算法维度下降到 $c-1$ 。

2 多视图判别特征变换正交

这里在 HOA^[7] 的算法基础上进行一部分改进, 对获得的投影变换矩阵进行正交变换。假设已经获取了三种视图的样本数据, 然后描述对 MVKDA 的具体实现过程。

2.1 获得所有视图核判别的变换 W_1, W_2, W_3

基于上述推导的费歇尔判别准则, 得到目标函数:

$$\max_W J(W) = \frac{|W^T S_B^\varphi W|}{|W^T S_W^\varphi W|} \quad (7)$$

因此通过解决下面的特征方程可以得到变换矩阵:

$$PW = \lambda W \quad (8)$$

其中, $P = S_W^{\varphi^{-1}} S_B^\varphi$ 。

根据式(8), 分别独立对每个视图求出相应的投影矩阵, 然后根据式(6)求出每个视图的特征, 再根据特征余弦最近邻算法分别求出每个视图的识别效果, 从高到低排序: $f(W_1) > f(W_2) > f(W_3)$ 。

2.2 更新 W_2

基于上述推导的核判别准则, 得到目标函数:

$$\max_{W_2} J(W_2) = \frac{|W_2^T S_{B_2}^\varphi W_2|}{|W_2^T S_{W_2}^\varphi W_2|} \quad (9)$$

$$\text{s. t. } W_2^T W_1 = 0$$

因此通过解决下面的特征方程对 W_2 进行更新。

$$P_2 W_2 = \lambda W_2 \quad (10)$$

其中, $P_2 = S_{W_2}^{\varphi^{-1}} (I - W_1 (W_1^T S_{W_2}^{\varphi^{-1}} W_1)^{-1} W_1^T - S_{W_2}^{\varphi^{-1}}) S_{B_2}^\varphi$ 。

因此投影矩阵 W_2 是由 P_2 的 $c-1$ 个最大非零特征对应的特征向量组成。

2.3 更新 W_3

基于上述推导的核判别准则, 得到目标函数:

$$\max_{W_3} J(W_3) = \frac{|W_3^T S_{B_3}^\varphi W_3|}{|W_3^T S_{W_3}^\varphi W_3|} \quad (11)$$

$$\text{s. t. } W_3^T W_1 = 0, W_3^T W_2 = 0$$

将 $W_3^T W_1 = 0, W_3^T W_2 = 0$ 转化成 $W_3^T [W_1, W_2] = 0$, 所以令 $W = [W_1, W_2]$ 。因此通过解决下面的特征方程对 W_3 进行更新。

$$P_3 W_3 = \lambda W_3 \quad (12)$$

其中, $P_3 = S_{W_3}^{\varphi^{-1}} (I - W (W^T S_{W_3}^{\varphi^{-1}} W)^{-1} W^T - S_{W_3}^{\varphi^{-1}}) S_{B_3}^\varphi$ 。

因此投影矩阵 W_3 是由 P_3 的 $c-1$ 个最大非零特征对应的特征向量组成。

MVKDA 算法描述:

步骤 1: 根据式(6)和式(8)计算出所有视图训练

样本的投影矩阵和特征向量,根据识别效果,得到 W_1 , W_2, W_3 ;

步骤 2:根据式(10)更新投影矩阵 W_2 ;

步骤 3:根据式(12)更新投影矩阵 W_3 ;

步骤 4:分别对 W_1, W_2, W_3 进行标准化;

步骤 5:利用投影矩阵 W_1, W_2, W_3 将各个视图的所有样本通过式(6)进行相应的投影,并将得到的每个样本的多视图特征融合在一起;

步骤 6:用余弦最近邻距离分类器^[14]对所有样本进行分类。

3 实验

为了验证文中算法的有效性,选择常用的公开人脸数据库 AR^[12] 和 Oxford Flowers17^[13] 作为实验数据库,以算法 HOA^[7]、KDA^[10]、MVFS^[5] 为对比方法,比较分类识别准确率。

3.1 数据库介绍

AR 公共彩色人脸数据库包含 102 类,每一类有 26 张图片,将图片处理成 60 * 60 的尺寸。考虑到表情、光照、位置、姿势等多种情况,为了有效地评价不同的变化对识别结果的影响,从每类样本中选择 8 个具有代表性的作为训练样本,剩余作为测试样本。图 1 为包含某类的部分图片。



图 1 AR 数据库的样本图像

Oxford Flowers17 数据库包含 17 种花,每种 80 张图片。图 2 为部分图片。选用 40 个训练样本,20 个测试样本,并从中提取出颜色、形状、纹理、HOG、HSV、SIFT 等特征作为多个视图特征。



图 2 Oxford Flowers17 数据库的样本图像

对 MVKDA 和相关算法进行实验对比。主要包括 HOA、KDA 和 MVFS。MVKDA 和三个比较方法在两个数据库上均做 20 次实验。

3.2 实验结果与分析

图 3 和图 4 分别给出了在 AR 和 Oxford Flowers17 两个数据库上所有对比方法随机 20 次的识别率波动图。表 1 给出所有方法在两个库上的平均识别率和方差。

从表 1 可以看出,MVKDA 有很好的分类性能。在 AR 公共彩色人脸数据库上,MVKDA 方法比 KDA、HOA 以及 MVFS 三种方法的平均识别率提高了至少 5.25% (93.90% - 88.65%);在 Oxford Flowers17 数据

库上,MVKDA 方法比三种对比方法的平均识别率提高了至少 5.35% (88.90% - 83.55%)。文中将原始样本通过高斯核函数投影映射到高维空间,从而解决低维非线性不可分问题;同时考虑到去除多视图之间的冗余信息,增加整体变换正交,从而提取出更加具有鉴别性的特征。实验结果充分证明 MVKDA 算法能够有效地提高识别率。

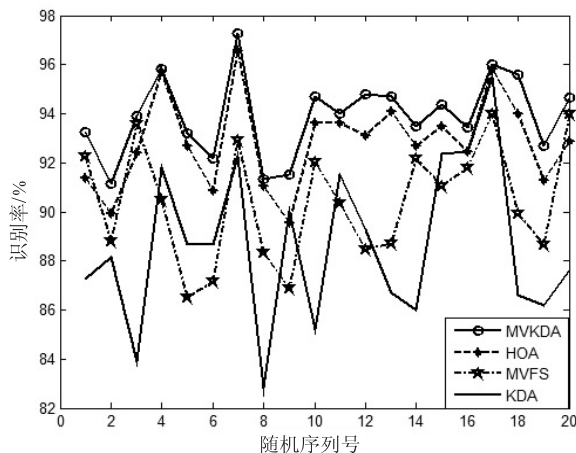


图 3 AR 数据库上的识别率

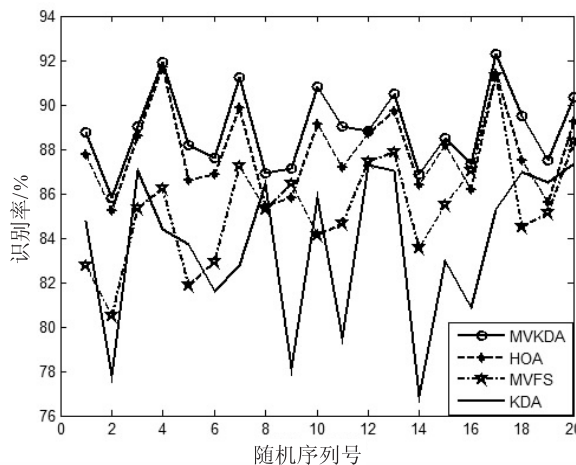


图 4 Oxford Flowers17 数据库上的识别率

表 1 所有方法在两个数据库上的识别率比较

方法	识别率(均值和方差)/%	
	AR 库	Oxford Flowers17
KDA	88.65 ± 8.25	83.55 ± 7.47
MVFS	90.04 ± 6.48	85.40 ± 5.50
HOA	92.14 ± 5.03	87.85 ± 4.84
MVKDA	93.90 ± 4.44	88.90 ± 3.88

4 结束语

结合核理论,并使各个视图的投影矩阵正交,文中提出一种基于多视图核鉴别分析的图像识别方法。该方法通过基于核的线性判别分析从各个视图中提取出相互正交的投影矩阵,从而能够提取出相应特征。在

AR 和 Oxford Flowers17 人脸数据库上的实验结果表明,MVKDA 与 MVFS、HOA 以及 KDA 相比,有效地提高了识别率。

参考文献:

[1] 朱长仁,王润生. 基于单视图的多姿态人脸识别算法[J]. 计算机学报,2003,26(1):104-109.

[2] Xiong N,Svensson P,Svensson P. Multi-sensor management for information fusion:issues and approaches[J]. Information Fusion,2002,3(2):163-186.

[3] Lai P L,Fyfe C. Kernel and nonlinear canonical correlation analysis[J]. International Journal of Neural Systems,2012,10(5):365-377.

[4] Shon A,Grochow K,Hertzmann A,et al. Learning shared latent structure for image synthesis and robotic imitation[C]//Advances in neural information processing systems. [s. l.]:[s. n.],2005:1233-1240.

[5] Li S Z,Zhu L,Zhang Z Q,et al. Statistical learning of multi-view face detection[C]//European conference on computer vision-part IV. [s. l.]:[s. n.],2002:67-81.

[6] Tang J,Hu X,Gao H,et al. Unsupervised feature selection for multi-view data in social media[C]//SDM. [s. l.]:[s. n.],2013:270-278.

[7] Jing X,Liu Q,Lan C,et al. Holistic orthogonal analysis of discriminant transforms for color face recognition [C]//17th IEEE international conference on image processing. [s. l.]:

IEEE,2010:3841-3844.

[8] 赵振勇,王保华,王力,等. 人脸图像的特征提取[J]. 计算机技术与发展,2007,17(5):221-224.

[9] Belhumeur P N,Hespanha J P,Kriegman D J. Eigenfaces vs. fisherfaces:recognition using class specific linear projection [J]. Pattern Analysis and Machine Intelligence,1997,19(7):711-720.

[10] Yang M H. Kernel eigenfaces vs. kernel fisherfaces:face recognition using kernel methods[C]//Proceeding of international conference on automatic face and gesture recognition. [s. l.]:[s. n.],2002:215.

[11] Mika S,Ratsch G,Weston J,et al. Fisher discriminant analysis with kernels[C]//Proceeding of IEEE international workshop on neural networks for signal processing IX. [s. l.]:IEEE,1999:41-48.

[12] Martinez A M,Benavente R. The AR face database[EB/OL]. 2009. http://cobweb.ecn.Purdue.edu/~aleix/aleix_face_DB.html.

[13] Nilsback M E,Zisserman A. A visual vocabulary for flower classification [C]//IEEE computer society conference on computer vision and pattern recognition. [s. l.]:IEEE,2006:1447-1454.

[14] Qian G. Similarity between Euclidean and cosine angle distance for nearest neighbor queries[C]//ACM symposium on applied computing. [s. l.]:ACM,2004:1232-1237.

(上接第 91 页)

applied to unsupervised classification of reviews[C]//Proc of the 40th annual meeting on association for computational linguistics. Stroudsburg,USA:Association for Computational Linguistics,2002:417-424.

[23] Goldberg A B,Zhu X. Seeing stars when there aren't many stars [C]//TextGraphs: the first workshop on graph based methods for natural language processing. [s. l.]:[s. n.],2006:45-52.

[24] Sindhvani V,Melville P. Document-word co-regularization for semi-supervised sentiment analysis[C]//Eighth IEEE international conference on data mining. [s. l.]:IEEE Computer Society,2008:1025-1030.

[25] Zhou S,Chen Q,Wang X. Active deep networks for semi-supervised sentiment classification [C]//International conference on computational linguistics. [s. l.]: Association for Computational Linguistics,2010:1515-1523.

[26] Zhou S,Chen Q,Wang X. Fuzzy deep belief networks for semi

-supervised sentiment classification [J]. Neurocomputing, 2014,131(9):312-322.

[27] Li S,Hao J. Spectral clustering-based semi-supervised sentiment classification[C]//Proc of the 8th advanced data mining and applications. Berlin:Springer,2012:271-283.

[28] Zhou Z H. Co-training paradigm in semi-supervised learning [C]//Proceeding of the Chinese workshop on machine learning and applications. Nanjing,China:[s. n.],2007.

[29] 郑文静,李雷. 基于图的组合半监督 SVM 聚类核算法研究[J]. 计算机技术与发展,2014,24(5):109-112.

[30] Weston J,Leslie C,Ie E,et al. Semi-supervised protein classification using cluster kernels [J]. Bioinformatics, 2005, 21(15):3241-3247.

[31] Szummer M. Partially labeled classification with Markov random walks [J]. Advances in Neural Information Processing Systems,2002(14):945-952.

[32] 李素科,蒋严冰. 基于情感特征聚类的半监督情感分类 [J]. 计算机研究与发展,2013,50(12):2570-2577.