

◎图形、图像、模式识别◎

融合 LDA 和多类 SVM 的图像语义映射研究

赵 炜,陈俊杰,李海芳

ZHAO Wei, CHEN Jun-jie, LI Hai-fang

太原理工大学 计算机与软件学院,太原 030024

College of Computer and Software, Taiyuan University of Technology, Taiyuan 030024, China

E-mail: zhaowei_02@126.com

ZHAO Wei, CHEN Jun-jie, LI Hai-fang. Research on image semantic mapping combining with multi-class SVM and LDA. *Computer Engineering and Applications*, 2009, 45(18): 164-166.

Abstract: Correlating image low-level feature with high-level semantic is one of the key questions of image semantic retrieval, SVM is an effective way. Multi-class SVM based on fuzzy C-means clustering is introduced to image semantic retrieval in order to facilitate the generation of rules. But because heterogeneous features of images are always motley, a great many branches of binary-tree are formed and mapping accuracy rate comes down obviously. Therefore, linear discriminant analysis is introduced to binary-tree to improve the algorithm's performance through pretreatment before clustering. The results show that the method builds up a more comprehensible tree configuration and improves the mapping correct rate due to LDA, and meets the require of image semantic mapping.

Key words: image semantic classification; support vector machine; fuzzy C-means clustering; linear discriminant analysis

摘 要: 建立图像低层特征到高层语义的映射是图像语义检索的关键问题之一, SVM 是其中行之有效的方法。为了便于规则生成, 将模糊 C 均值聚类 SVM 多类分类方法应用于图像语义映射。但由于异类图像特征常常混杂, 最终形成的二叉树分支一般很多, 映射准确率下降明显。为此, 将线性判别分析法引入二叉树建树过程中, 通过聚类之前先对特征优化处理来改进算法性能。实验结果表明该方法建立起了更便于理解的分树结构且 LDA 的引入使得映射准确率有所提高, 满足了图像语义映射的要求。

关键词: 图像语义分类; 支持向量机; 模糊 C 均值聚类; 线性判别分析

DOI: 10.3778/j.issn.1002-8331.2009.18.049 **文章编号:** 1002-8331(2009)18-0164-03 **文献标识码:** A **中图分类号:** TP391

1 引言

图像语义映射就是利用机器学习方法对图像低层特征进行学习, 挖掘图像特征与语义之间的关联, 从而建立起图像特征到高层语义的桥梁, 它包含两个关键步骤: 一是低层有效特征的提取; 二是映射算法的采用。目前, 应用于图像语义映射的技术已有很多, 包括: 贝叶斯、神经网络、遗传算法、聚类等, 但是这些方法只有在样本数目足够多的前提下, 其性能才有理论上的保证, 而在实际问题中, 样本数量往往是有限的^[1]。基于统计学习理论的支持向量机(Support Vector Machine, SVM)作为继神经网络之后模式识别领域新的研究热点, 以其在解决小样本、非线性以及高维模式识别中所具有的优势而成为图像语义映射中一个行之有效的方法。

实现图像语义映射的过程可以看作是一个根据高层语义的多类分类问题, 目前已有的基于 SVM 的研究主要集中于 1-a-r 多类分类策略^[2-3]。然而, 由于支持向量机的黑盒性和相应

规则抽取算法的不成熟, 由此建立起来的映射模型很难从中产生便于人类理解的规则。

通过基于聚类的 SVM 多类分类方法建立起决策二叉树^[4], 输入为图像的低层特征, 叶节点代表相应的语义类别, 树的不同分支就可以代表不同的规则, 使得识别过程更便于人类理解。针对异类图像特征可能相似的特点, 将线性判别分析法(Linear Discriminant Analysis, LDA)引入二叉树建树过程中, 每次通过寻找最优投影使得同类样本聚集, 异类样本松散, 之后再聚类。实验结果表明本文算法有效地提高了分类树图像语义识别的准确率, 满足图像语义映射的要求。

2 算法基本思想及要素

2.1 算法基本思想

对于图像来说, 具有相同语义的图像低层特征可能不同, 而不同语义类别的图像低层特征可能相近。这一特点决定了异

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60773004); 山西省自然科学基金(the Natural Science Foundation of Shanxi Province of China under Grant No.2007011050)。

作者简介: 赵炜(1984-), 男, 硕士, 主要研究领域为智能信息处理; 陈俊杰(1956-), 男, 博士, 教授, 主要研究领域为数据库与智能信息处理; 李海芳(1964-), 女, 博士, 教授, 主要研究领域为智能信息处理、情感计算。

收稿日期: 2008-09-25 **修回日期:** 2009-03-06

类图像常常是混杂在一起的,而聚类则是将具有相似特征的图像划分在同一分组,对于图像来说依靠聚类建立起的分类树分支必然很多,增大了识别错误的可能性。如果可以使在每次聚类之前异类图像特征尽可能的分离,这样必然会降低异类样本划分的难度,简化分类树结构,就有可能提高识别准确率。

因此本文的基本思想为:在形成每个节点的过程中,先将特征进行处理,使得同类样本特征聚集而异类样本特征松散,减小异类样本划分难度,再进行聚类,来优化分类树结构,改进识别效果。

2.2 基于模糊 C 均值聚类的多类 SVM(FCM-SVM)

FCM-SVM 算法的核心思想就是每次将待训练样本集通过聚类算法划分成两个分组,训练 SVM 分类器,再对每个分组分别聚类,直到每个分组只有一类样本为止,从而将多类问题转化为树形的多个二分类问题。模糊 C 均值聚类算法(Fuzzy C-means clustering-FCM)^[5]以其实现简单、收敛速度快等优点而成为最常被使用的聚类算法之一,本文利用它作为聚类的依据构建二叉树。

设 $X=\{x_i, i=1, 2, \dots, n\}$ 为待聚类样本集, n 为样本个数, $v_i(i=1, 2, \dots, C)$ 为各类的聚类中心,其中 C 为待聚类数目, $u_{ij}(i=1, 2, \dots, C, j=1, 2, \dots, n)$ 是第 j 个样本对第 i 类的隶属度函数。

FCM 的目标函数为:

$$J = \sum_{i=1}^C \sum_{j=1}^n u_{ij}^m \|x_j - v_i\|^2$$

约束条件为:

$$\sum_{i=1}^C u_{ij} = 1 (\forall j=1, 2, \dots, n)$$

其中 $m \in (1, \infty)$ 为加权指数。

FCM-SVM 算法步骤如下:

- (1) 确定聚类数目 $C=2$, 初始化加权指数 m 和聚类中心;
- (2) 根据下式更新隶属度函数和聚类中心:

$$u_{ij} = \frac{1}{\sum_{i=1}^C \left(\frac{\|x_j - v_i\|^2}{\|x_j - v_i\|^2} \right)^{\frac{1}{m-1}}}, \forall i=1, 2, \dots, C, j=1, 2, \dots, n$$

$$v_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m}, \forall i=1, 2, \dots, C$$

(3) 若满足终止条件 $|J^{(i)} - J^{(i-1)}| \leq \varepsilon$ (其中 i 表示第 i 次循环), 则聚类完成, 待分类集被划分为 A, B 两个分组, 转(4), 否则跳转(2)继续迭代;

(4) 利用 A, B 作为正负类训练 SVM 二分类器;

(5) 若每个分组都只包含一个类别则停止迭代, 否则对该分组继续上述步骤, 从而形成一棵二叉树。

2.3 线性判别分析

线性判别分析法基于 Fisher 准则, 基本原理为通过样本类间离散度和类内离散度比值的最大化, 将同类样本聚集在一起, 而异类样本相对松散, 从而提高模式分类的性能^[6]。它是一种考虑样本类别信息的特征降维方法。

设待训练样本集为: $X=\{x_i, i=1, 2, \dots, n\}, x_i \in R^N$ 。其中 n 为样本个数, N 为样本特征维数, 共分为 c 类, 降维后 $x_{Li} \in R^m$

($m < N$), 则 LDA 就是寻找线性变换矩阵 $W \in R^{N \times m}$ 使得样本类间离散度与类内离散度的比值最大化。

样本类间离散度矩阵定义为:

$$S_b = \sum_{i=1}^c p_i (\mu_i - \mu)(\mu_i - \mu)^T$$

样本类内离散度矩阵定义为:

$$S_w = \frac{1}{n} \sum_{i=1}^c \sum_{x \in c_i} (x - \mu_i)(x - \mu_i)^T$$

式中 $p_i = n_i/n$ 是每个类的先验概率, n_i 是 $c_i (i=1, 2, \dots, c)$ 类的训练样本数, $\mu_i = \frac{1}{n_i} \sum_{x \in c_i} x$ 为 c_i 的样本均值向量, $\mu = \frac{1}{n} \sum x$ 为所有样本的均值。

LDA 的目标就是找一个最佳投影 W_{opt} , 使得:

$$J(W_{opt}) = \arg \max \frac{W^T S_b W}{W^T S_w W}$$

则新样本点 x 在投影空间的映射即为: $x_L = W_{opt}^T x$ 。

3 融合 LDA 的 FCM-SVM 算法(LFCM-SVM)

针对图像特征维数高且异类特征混杂的特点, 本文利用 LDA 对于特征的优化处理来实现同类样本聚集、异类样本松散, 将其应用于 FCM-SVM 的每个节点之上, 每次首先对待分类样本集进行 LDA 投影, 再对投影后的特征集执行聚类从而将原特征集划分成两个分组, 利用降维后的特征训练 SVM 分类器, 再对每个分组的原特征集分别执行上述过程, 建立起图像语义识别的二叉树结构。

LFCM-SVM 算法流程如图 1 所示。

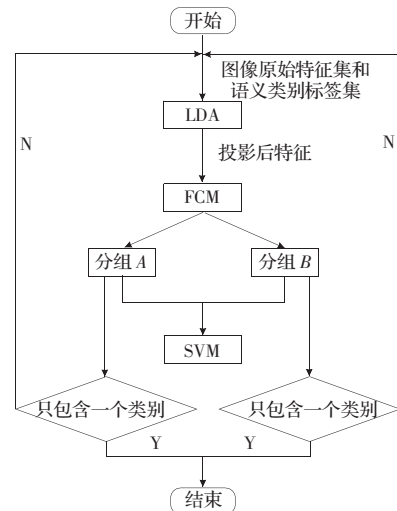


图 1 LFCM-SVM 算法流程图

算法具体描述如下:

设 $X=\{x_i, i=1, 2, \dots, n\}$ 为待训练图像低层特征集, 其中 $x_i \in R^N$, $L=\{l_i, i=1, 2, \dots, n\}$ 为图像对应语义类别标签集。

训练阶段:

(1) 根据当前待划分训练特征集 X 和对应标签集 L 执行 LDA, 找到最佳投影方向对图像特征进行预处理, 处理后数据集为 X_L , 在本结点内保存映射矩阵 W ;

(2) 对处理后的特征集 X_L 执行 FCM 聚类, 形成分组 X_{LA} 和

X_{L_B} , 对应原特征集分为两个分组 X_A, X_B , 标签相应分为 L_A, L_B ;

(3) 以 X_{L_A} 和 X_{L_B} 分别作为正负类训练 SVM 分类器 s 保存于本结点;

(4) 分别对 $X_A, L_A; X_B, L_B$ 执行步骤(1)~(3), X_A 的执行结果作为本结点的左子树, X_B 的执行结果作为当前结点的右子树, 直到每个分组只包含一个类别为止。

映射阶段:

(1) 从根结点开始, 对于待识别图像特征 x , 根据本结点映射矩阵 W 将其映射到目标子空间一点 x_L ;

(2) 将 x_L 送入该结点所对应的 SVM 分类器, 结果为 1 则深入左孩子结点, 若为 -1 则深入右孩子结点继续上述过程, 直到叶结点得到该叶节点对应的某一语义类别标签, 实现图像低层特征到高层语义的映射。

4 实验及结果分析

实验用图像选自 corel 图像库, 从中选择 200 幅图像作为研究对象, 包括蝴蝶 50 幅, 落日 50 幅, 花朵 50 幅和汽车 50 幅, 选取时尽量包含各种颜色, 不同形态的图像。将所选各类图像中 40 幅作为训练集, 10 幅作为测试集, 即训练集共 160 幅, 测试集共 40 幅。各类样图如图 2 所示。



图 2 实验用图样例

对于图像低层特征的提取基于 MPEG-7, 主要采用颜色和纹理特征。

颜色是图像最为重要的视觉特征之一, 也是基于内容图像检索中最常被使用的低层特征, 本文主要采用基于 HSV 颜色空间的 255 维颜色直方图来表示图像的颜色特征, 具体算法见文献[7]。

边缘对于图像感知起到重要作用, 实验提取的纹理特征为 MPEG-7 中定义的边缘直方图, 主要包括 5 种边缘 ($0^\circ, 45^\circ, 90^\circ, 135^\circ$ 和一种无方向边缘) 的空间分布统计, 共计 5 维的纹理特征, 具体算法见文献[7]。

实验环境选择 Matlab 7.0, 所用工具箱为: Maastricht 大学 Laurens van der Maaten 的特征降维工具箱和台湾 Chih-Jen Lin 的 LIBSVM 工具箱。

分别以颜色直方图 (colorhist, 256 维), 边缘直方图 (edgehist, 5 维) 和两种特征融合 (both, 261 维) 三种方式作为图像特

征进行实验, 比较了 FCM-SVM 和 LFCM-SVM 及传统的 $1-a-r$ 方法在同等条件下的准确率。选用径向基函数作为核函数, 参数设置为: $C=1\ 000, \sigma=0.07$ 。考虑到聚类算法初始聚类中心的不确定会使最终生成的树形结构不同, 最终导致映射准确率有所差异, 每组实验均重复 10 次, 映射准确率取 10 次的平均值, 实验对比结果如表 1 所示。

表 1 映射准确率对比结果

选用特征	$1-a-r$	FCM-SVM	LFCM-SVM
Colorhist	0.775	0.765	0.850
edgehist	0.725	0.640	0.655
both	0.850	0.860	0.925

由实验结果可以看出: 采用颜色与纹理的特征组合映射准确率要比单独采用其中一个有显著提高; 除边缘直方图外, $1-a-r$ 与 FCM-SVM 分类性能接近而 LFCM-SVM 要优于二者。总体上, 由于 LDA 的引入, 使得映射准确率较原算法有了一定程度的提高, 得到了满意的结果。

5 结束语

将线性判别分析法引入 SVM 二叉树建树过程之中, 通过寻找最佳投影来更好地实现异类样本分离, 可以有效改善基于聚类 SVM 图像语义映射的准确率。而二叉树的建立又使得规则提取更为便利。实验表明, 本方法在满足映射需求的同时提高了映射过程的可理解性。但是本文并未考虑聚类算法及 SVM 本身的优化, 使得算法并未达到最佳性能, 这将作为今后的研究方向, 以期进一步提高算法执行效果。

参考文献:

- [1] 黎曦, 谢毓湘, 吴玲达. 基于 SVM 的图像分类系统设计[J]. 系统仿真学报, 2006, 18(z1): 385-387.
- [2] 万华良, Chowdhury M U. 基于支持向量机的图像语义分类[J]. 软件学报, 2003, 14(11): 1891-1899.
- [3] 成洁, 石跃祥, 易璨. FSVM 在图像低层特征与高层语义关联中的应用[J]. 小型微型计算机系统, 2007, 28(6): 119-122.
- [4] Liu Song, Yi Hao-ran, Chia L T, et al. Adaptive hierarchical multi class svm classifier for texture based image classification[C]//IEEE International Conference on Multimedia and Expo, 2005: 1191-1194.
- [5] Stan D, Sethi I K. Mapping low-level image features to semantic concepts[C]//Proceedings of SPIE-The International Society for Optical Engineering, 2001: 172-179.
- [6] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces VS Fisherfaces: Recognition using class specific linear projection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.
- [7] Manjunath B S, Ohm J R, Vasudevan V V, et al. Color and texture descriptors[J]. IEEE Trans CSVT, 2001, 11(6): 703-715.