摘要

随着多媒体技术以及计算机互联网技术的迅速发展，人们生活中触手可及的图像数据呈现爆发式的增长趋势，这对机器自动分析和识别图像内容的技术提出了迫切的需求和挑战。一方面，受制于图像底层信息和高层语义之间的语义鸿沟，图像中的物体类别信息很难直接从底层特征获得，因此需要借助物体识别的方法建立从底层特征到高层语义的映射。另一方面，不同的底层特征对物体识别的贡献不尽相同，既存在类内的变化，也存在一定的类间差异及共性。因此，本文针对基于多核函数的物体识别方法进行深入研究，致力于设计不同的多核函数组合以及对应的多核学习方法，挖掘不同问题背景下各种底层特征和核函数对物体识别的贡献，从而建立从多种底层特征到物体类别的合理映射。本文的主要创新点包括：

第一，提出了一种多核主动学习方法，用于解决训练样本筛选及多核学习方法计算复杂度较高的问题。该方法将主动学习引入多核学习框架，根据无标签样本的分布情况，通过一种局部自适应的样本筛选策略，自顶向下地利用样本簇间竞争和簇内竞争筛选出一批最有信息量并兼顾代表性及差异性的样本，进行类别标注并参与分类器训练。实验表明，该方法在不影响识别性能的前提下，有效地去除了训练样本冗余，缓解了多核学习算法复杂度较高的问题，并取得了显著优于已有主动学习方法的结果。

第二，提出了一种样本特定的多核学习方法，用于解决类内变化较大的物体识别问题。该方法提出了样本特定的多核函数组合对样本间的相似性进行度量，挖掘了各个核函数在不同样本上对于识别的贡献。该方法对物体的类内多样性具有较好的适应性。多个公开数据集上的大量实验表明，相比已有多核学习方法，该方法能更好地适用于类内变化较大的物体识别问题，并取得了与目前国内外前沿的物体识别方法性能相当的结果。

第三，提出了一种簇敏感的多核学习方法，用于解决不同物体类别间数据混叠的物体识别问题。该方法将样本簇引入多核学习框架，对物体类内多样性及类间相关性进行建模，提出了簇敏感的多核函数组合方式，有效地挖掘了不同核函数在样本簇上对于识别的贡献，在捕捉物体类内变化的同时兼顾了物体类别的共

性。该方法在模型复杂度上对样本特定的多核学习方法进行了改善，降低了过拟合的风险。多个公开数据集上一系列的实验表明，该方法对于复杂数据中由类内多样性和类间相关性带来的数据混叠问题具有较强的适应能力，其识别能力显著优于已有的多核学习方法，并取得了与国际前沿的物体识别方法性能相当甚至更优的结果。

第四，提出了一种类共享的多核学习方法，用于解决多类物体识别问题中，多核函数组合的知识共享问题。该方法在物体识别领域将多核学习拓展到了多标签学习问题，并在统一的学习框架下，根据训练数据的充分程度，提供了多核函数组合在物体类别间相互独立、部分共享或者完全共享的多核学习方法。该方法在充分考虑物体类别在多核函数空间个性的同时，有效地挖掘了多个类别在多核函数空间的共性。实验表明，该方法能削弱训练数据不充分带来的不利影响，具有较强的识别能力和泛化性能。

综上所述，本文针对物体识别的多核学习问题进行了深入的研究。在充分考虑物体类别类内多样性、类间相关性以及类间共性等因素的基础上，挖掘了多种底层特征对于不同复杂程度的物体识别问题的贡献，提出了四种多核学习方法，并实现了基于视觉内容的图像分类系统，通过大量的实验验证了本文提出的方法的有效性。此外，本论文的研究成果将为物体识别领域开展更深入的多核学习方法研究奠定基础。

关键词：物体识别，图像分类，多核学习，主动学习

Multiple Kernel Learning for Object Recognition

Jingjing Yang (Computer Application)
Supervised by Professor Wen Gao

Abstract

With the advancement of multimedia and internet technologies, the amount of image data is growing explosively, leading to pressing needs and new challenges on the techniques of automatic content analysis and object recognition. On one hand, due to the well-known semantic gap between low-level features and high-level semantic concepts, object categories are difficult to be identified from low-level features directly. On the other hand, different low-level features vary in their discriminative abilities for different object categories, exhibiting intra-class variance and inter-class correlation. To approach these problems, this dissertation focuses on multiple kernel learning for object recognition and introduces new multiple kernel combinations and their corresponding learning methods. By investigating the contributions of different kernels, mappings from multiple low-level features to object categories can be established. The main contributions of this dissertation can be summarized as follows:

Firstly, a multiple kernel active learning method (MKL^{active}) is proposed to alleviate the burden of manually labeling and reduce the computation cost. Based on the distribution of unlabeled data, MKL^{active} adopts a local adaptive sampling strategy to actively select the most informative, representative and diverse training samples for manually labeling and MKL training. In this methodology, uncertain samples are firstly clustered into groups and the samples are consequently selected via inter-group and intra-group competitions. Experimental results show that by local adaptive sampling, MKL^{active} can significantly reduce the burden of labeling and the computation cost in MKL training, and outperforms the state-of-the-art active learning methods in terms of recognition rate.

Secondly, a per-sample multiple kernel learning method (PS-MKL) is proposed to take into account the intra-class variance. In PS-MKL, sample-specific multiple kernel combination, where the kernel weight depends not only on the corresponding kernel, but also on the samples to be compared, is introduced to investigate the contributions of different kernels on each training sample. Extensive experiments are carried out over three benchmarking datasets demonstrating that PS-MKL is able to achieve encouraging performance and outperform MKL.

Thirdly, a group-sensitive multiple kernel learning method (GS-MKL) is proposed to accommodate the intra-class diversity and the inter-class correlation. In GS-MKL, group is introduced between the object class and individual images as an intermediate representation and group-sensitive multiple kernel combinations are proposed to investigate the contributions of different kernels on each sample groups. Compared with PS-MKL, GS-MKL provides a tractable solution to obtain a trade-off between capturing the diversity and the invariance for each object class, leading to a simplified model complexity and lower risk of overfitting. On three benchmarking datasets, GS-MKL has achieved the encouraging performance comparable to the state-of-the-art and outperformed several existing MKL methods, including PS-MKL.

Fourthly, a class-shared multiple kernel learning method (CS-MKL) is proposed for multi-class object recognition with knowledge sharing. CS-MKL extends MKL to the task of multi-label learning and provides a uniform solution for different scenarios such as independent, partially shared, and fully shared multiple kernel learning among multiple object classes. Experiments substantiate that CS-MKL with the partially shared multiple kernel combination, which explores both the general and individual patterns in multi-kernel space among object classes, can effectively alleviate the negative effect of training data insufficiency and achieve a more powerful generalization ability.

In summary, this dissertation performs a comprehensive study on multiple kernel learning methods for object recognition. In particular, focusing on the intra-class diversity, inter-class correlation and inter-class invariance in object recognition, four multiple kernel learning methods are proposed for sample selection, class-independent and class-shared multiple kernel learning respectively. A content-based image classification system is implemented. Extensive experiments demonstrate the effectiveness of the proposed methods. Furthermore, the achieved results of this work lay the foundation of further research of multiple kernel learning for object recognition.

Keywords: object recognition, image classification, multiple kernel learning, active learning