

# 基于边际 Fisher 准则和迁移学习的小样本集分类器设计算法

舒醒<sup>1</sup> 于慧敏<sup>1,2</sup> 郑伟伟<sup>1</sup> 谢奕<sup>1</sup> 胡浩基<sup>1</sup> 唐慧明<sup>1</sup>

**摘要** 如何利用大量已有的同构标记数据(源域)设计小样本训练数据(目标域)的分类器是一个具有很强应用意义的研究问题. 由于不同域的数据特征分布有差异, 直接使用源域数据对目标域样本进行分类的效果并不理想. 针对上述问题, 本文提出了一种基于迁移学习的分类器设计算法. 首先, 本文利用内积度量的边际 Fisher 准则对源域进行特征映射, 提高源域中类内紧凑性和类间区分性. 其次, 为了筛选合理的训练样本对, 本文提出一种去除边界奇异点的算法来选择源域密集区域样本点, 与目标域中的标记样本点组成训练样本对. 在核化空间上, 本文学习了目标域特征到源域特征的非线性转换, 将目标域映射到源域. 最后, 利用邻近算法(k-nearest neighbor, kNN)分类器对映射后的目标域样本进行分类. 本文不仅改进了边际 Fisher 准则方法, 并且将基于自适应样本对筛选的迁移学习应用到小样本数据的分类器设计中, 提高域间适应性. 在通用数据集上的实验结果表明, 本文提出的方法能够有效提高小样本训练域的分类器性能.

**关键词** 小样本集分类器, 迁移学习, 边际 Fisher 准则, kNN 分类器, 域间转换

**引用格式** 舒醒, 于慧敏, 郑伟伟, 谢奕, 胡浩基, 唐慧明. 基于边际 Fisher 准则和迁移学习的小样本集分类器设计算法. 自动化学报, 2016, 42(9): 1313–1321

**DOI** 10.16383/j.aas.2016.c150560

## Classifier-designing Algorithm on a Small Dataset Based on Margin Fisher Criterion and Transfer Learning

SHU Xing<sup>1</sup> YU Hui-Min<sup>1,2</sup> ZHENG Wei-Wei<sup>1</sup> XIE Yi<sup>1</sup> HU Hao-Ji<sup>1</sup> TANG Hui-Ming<sup>1</sup>

**Abstract** It has great practical significance to design a classifier on a small dataset (target domain) with the help of a large dataset (source domain). Since feature distribution varies on different datasets, the classifiers trained on the source domain cannot perform well on a target domain. To solve the problem, we propose a novel classifier-designing algorithm based on transfer learning theory. Firstly, to improve the compactness of the same category and separateness of different categories in the source domain, this paper utilizes the extended margin Fisher criterion where the distance is measured by the inner product between data. Secondly, to select good sample pairs for transfer learning, this paper presents an algorithm to get rid of marginal singular points by selecting high-density samples in the source domain. The non-linear feature transformation mapping the target domain to the source domain is learned in the kernel space. Finally, k-nearest neighbor (kNN) classifiers are trained for classification. Compared with the existing works, this paper not only extends the margin Fisher criterion, but also applies the transfer learning theory based on the algorithm of selecting training sample pairs to design classifiers of a small dataset. We experimentally demonstrate the superiority of our method to effectively improve the performance of classifiers on the general datasets.

**Key words** Classifiers on a small dataset, transfer learning, margin Fisher criterion, k-nearest neighbor (kNN) classifiers, domain transformation

**Citation** Shu Xing, Yu Hui-Min, Zheng Wei-Wei, Xie Yi, Hu Hao-Ji, Tang Hui-Ming. Classifier-designing algorithm on a small dataset based on margin Fisher criterion and transfer learning. *Acta Automatica Sinica*, 2016, 42(9): 1313–1321

收稿日期 2015-09-09 录用日期 2015-12-11  
Manuscript received September 9, 2015; accepted December 11, 2015

国家自然科学基金(61471321), 教育部-中国移动科研基金(MCM20150503), 国家自然科学基金(61202400), 浙江省自然科学基金(LQ12F02014)资助

Supported by National Natural Science Foundation of China (61471321), Ministry of Education – China Mobile Research Fund (MCM20150503), National Natural Science Foundation of China (61202400), and Natural Science Foundation of Zhejiang Province (LQ12F02014)

本文责任编辑 胡清华  
Recommended by Associate Editor HU Qing-Hua

1. 浙江大学信息与电子工程学院 杭州 310027 2. 浙江大学 CAD&CG 国家重点实验室 杭州 310027

1. College of Information Science and Electronic Engineering,

在模式识别和计算机视觉领域中, 物体分类识别方法已经得到广泛的研究, 现存的算法也比较成熟. 但是在一个标记样本很少的数据库上, 这些传统的方法并不能得到较高的识别准确度, 这是因为几乎所有的监督学习算法<sup>[1-2]</sup>都需要大量的标记样本作为训练数据.

近年来, 基于一个大量标记的同构数据集上的物体识别模型, 设计一种基于迁移学习的分类器算法来提高新的只有少量标记样本的数据集上的识别准确度是一个越来越受关注的研究方向, 其中大样

Zhejiang University, Hangzhou 310027 2. The State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310027

本和小样本数据集分别被称作源域和目标域<sup>[3]</sup>, 其主要原因是获取大量标记样本是一个昂贵、极其耗费人力和时间的过程, 同时不可避免人为标记错误的情况<sup>[4]</sup>. 借助于源域来优化目标域的分类模型能很好地解决上述问题. 但是源域训练得到的分类器直接应用在目标域中并不能得到理想的效果, 这是因为不同域的特征分布有差异. 导致域间差异的因素有很多, 比如相机参数、光线、视角、分辨率、背景和姿态等<sup>[3, 5-6]</sup>, 如图 1 所示. 因此, 我们可以将源域作为目标域的先验知识, 利用域间的相关性, 通过迁移学习的方法来提高目标域的分类准确率<sup>[6-8]</sup>.



图 1 源域(上)与目标域(下)是存在差异的(与源域相比, 目标域中的背景更复杂, 分辨率更低, 视角更多样)

Fig.1 There exist differences between source domain (top) and target domain (bottom) (Compared with the source domain, the target domain contains more complex backgrounds, the lower resolution, and more camera angles.)

利用大样本数据学习小样本数据分类器问题的算法从采用的参数模型的角度主要可以分为三类: 基于支持向量机 (Support vector machine, SVM) 的算法、基于深度学习的算法和基于域间特征转换的算法. 基于 SVM 的算法, 比如文献 [9-13], 通过改变在源域上训练得到的支持向量机中的参数或者目标函数来适应目标域. 基于 SVM 的算法最大的缺点是不能够将算法扩展到目标域中标记样本未出现过的新类别的情况. 基于深度学习的算法是将在大量已知标记样本集上训练得到的神经网络应用到一个相对少量标记的数据域上. 文献 [14] 是基于 ILSVRC2012 数据库上训练好的 CNNs (Cellular neural networks), 利用目标域上的标记样本重新训练了全连接层. 文献 [15] 是基于在 ILSVRC13 数据库上训练得到的 OverFeat 网络模型, 利用目标域中的少量标记样本以及 OverFeat 学习出来的特征, 训练出适应目标域的分类器. 基于深度学习的算法的局限性是需要海量训练数据来保证分类准确度. 基于域间特征转换的算法也有很多, 比如文献 [3, 16-19] 等. 文献 [18] 提出了一种将源域和目标域中的高维特征映射到低维特征空间中, 实现域间迁移的学习算法. 基于域间特征转换的算法刚好可

以克服掉基于 SVM 的算法的最大缺点, 例如文献 [16] 提出的 ARC-t 算法. 但是 ARC-t (Asymmetric regularized cross-domain transformation) 算法对源域中类别的可分离性以及学习特征空间转换中训练点的选择缺乏充分的考虑.

综上, 为了充分利用源域的标记同构数据和保证目标域中不同类映射到源域后尽可能相互不重叠, 本文提出了一种新的基于域间特征转换的迁移学习方法. 首先, 针对源域的样本, 本文基于边际 Fisher 准则, 提出了特征向量点积作为距离度量标准, 尽可能最大化类间分离性和最小化类内紧凑性, 这个过程将直接影响后面的转换学习的效果. 其次, 本文设计了一种去除边缘奇异点的算法, 基于 K-medoids 思想从源域中选取合理的样本点, 学习与目标域标记样本间的特征转换, 实现目标域到源域的映射. 本文采用该去除奇异点算法的原因有两个: 1) 由于源域中每个类别中样本点的分散性, 必然在边界附近存在部分奇异点, 所以采用 K-medoids 思想能够避免将目标域相应类别映射到源域的边界附近; 2) k-means 算法同样可以达到以上效果, 但是该算法计算类的中心点时会受到奇异点的影响, 并不能将目标域映射到源域中样本密集分布区域. 由于线性的特征转换有时在特征空间上不能令不同类别可区分, 所以为实现特征的非线性转换, 本文利用了核化特征, 最后训练邻近算法 (k-nearest neighbor, kNN) 分类器并在通用数据库上计算物体分类识别准确度.

本文的结构如下: 第 1 节介绍本文所提算法涉及的相关工作; 第 2 节详细阐述本文提出的基于边际 Fisher 准则和迁移学习的分类器设计算法; 第 3 节是实验, 给出了本文算法的具体实现过程, 并把本文算法与近年来一些具有代表性的基于迁移学习的小样本分类器设计算法进行了性能对比; 第 4 节是论文总结及后续研究方向.

## 1 相关工作

在近年来, 利用大样本数据集来提高小样本集分类器性能的问题吸引了越来越多的关注. 国内外很多学者针对该问题做了许多研究工作, 也提出了很多思路<sup>[6]</sup>. 目前提出的方法从目标域是否进行特征映射的角度主要分为以下两类: 一类是在保持源域和目标域在特征分布空间不变的前提下, 在目标域上通过调节分类器的适应性参数来达到提高小样本数据分类准确率的的目的; 另一类是在保持目标域中的特征分布不变的前提下, 将目标域上的点在特征空间中映射到源域中, 直接利用在源域中训练好的分类器. 在第一类方法中, 文献 [9] 提出一种自适应 SVM 算法, 通过增加扰动函数和目标域中的标记样本来更新源域中训练好的分类器  $f^A(\mathbf{x})$ , 即

目标域上的分类器为  $f^T(\mathbf{x}) = f^A(\mathbf{x}) + \delta f(\mathbf{x})$ , 其中  $\delta f(\mathbf{x})$  是扰动函数. 文献 [11] 中将转换 SVMs 应用到了适应性学习中. 同时, 很多工作是基于第二类的思想进行研究的. 文献 [20] 提出的算法是基于测地流核的特征迁移学习算法实现的. 在文献 [21] 中, Long 等学习了一种能够直接使源域和目标域间的特征分布相匹配的域间不变核, 从而达到域间迁移学习的目的. 文献 [22] 通过一种自动域间适应的特征学习算法来适应不同图片数据域的特征差异性. 在文献 [3] 中, Saenko 等提出了一种基于信息论矩阵学习<sup>[23]</sup>的特征转换算法 (Symm 算法), 在特征空间上将目标域映射到源域, 通过源域上的分类器对目标域进行分类. 在文献 [16] 中, Kulis 等提出了一种叫做 ARC-t 的算法, 该算法的核心是一种基于域间特征核化非线性转换的模型并训练出 kNN 分类器.

与本文相关的工作还有文献 [24–25], 它们提出了边际 Fisher 分析算法 (Marginal Fisher analysis, MFA), 该算法是使用欧氏距离作为度量标准并基于边际 Fisher 准则来设计特征映射, 不仅对数据分布没有严格的要求, 而且能更好地表达出类内紧凑性和类外分离性, 有助于将不同类别更好地区分开. 为了提高域间适应性, 虽然基于信息论矩阵学习的算法和 ARC-t 算法能够将目标域在特征空间中映射到源域中, 但是存在两个明显的问题: 1) 源域内不同类如果在特征空间中相互交错, 不具有很好的类别

差异性, 那么目标域中的类别映射到源域也不能保证相互之间是可区分的. 2) 在特征转换训练过程中, 源域中的训练样本是随机选取并与目标域中的标记样本组成训练对, 如果该过程选取的点是在源域对应类别的边界上, 那么目标域样本也会被映射到源域的边界区域, 这样势必会影响类别间的分类. 综上所述, 首先我们需要将源域中的样本点降维到类间可分离的特征空间中, 使得在该空间中, 源域的不同类别相互可区分. 另外, 为了保证目标域的映射结果具有较高的分类准确度以及训练样本对的选择要避免边缘区域的样本点, 本文算法的设计正是基于上述考虑.

## 2 基于边际 Fisher 准则和迁移学习的小样本集分类器设计算法

为了利用大量已有的同构标记数据提高小样本集的分类器性能, 本文提出了基于边际 Fisher 准则和迁移学习的小样本数据的分类器设计算法. 该算法的思路是首先为了提高源域的内类紧凑性和类间区分性, 本文根据内积度量的边际 Fisher 准则优化源域特征分布, 其次利用筛选算法得到的训练样本对学习目标域到源域的非线性特征转换的迁移学习算法, 将目标域映射到源域, 提高域间适应性, 最后利用在源域上训练得到的 kNN 分类器对目标域的样本进行分类. 本文的分类器设计算法流程图见图 2.

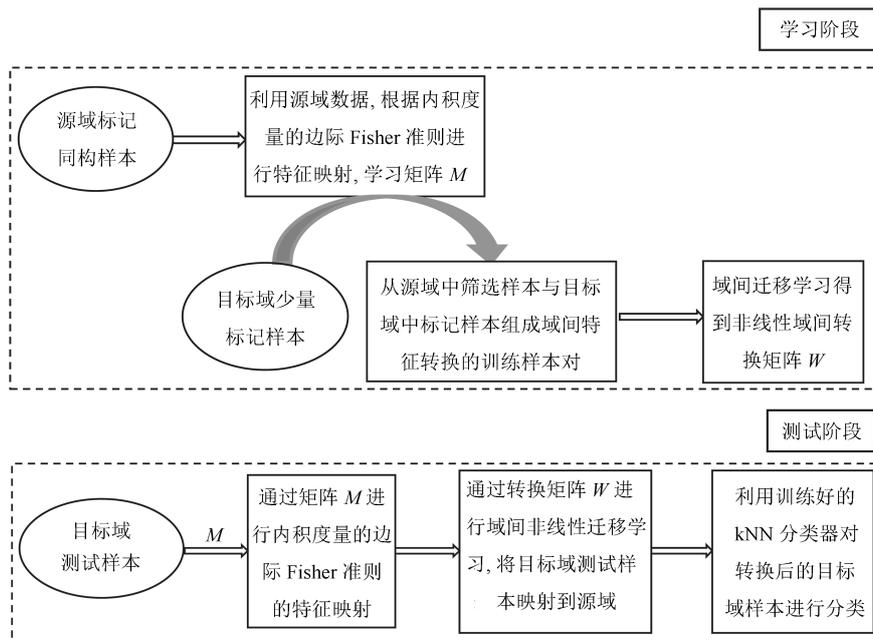


图 2 基于边际 Fisher 准则和迁移学习的小样本数据分类器设计算法流程图

Fig. 2 The flow diagram of classifier-designing algorithm on a small dataset based on margin Fisher criterion and transfer learning

首先, 我们考虑到了源域内特征分布, 利用内积度量的边际 Fisher 准则对源域进行特征映射, 提高源域中类内紧凑性和类间区分性. 其次, 在目标域到源域的特征转换的学习过程中, 为了筛选合理的训练样本对, 本文提出一种去除边界奇异点的算法来选择源域密集区域样本点, 与目标域中的标记样本点组成训练样本对. 并在核化空间上, 本文学习了目标域特征到源域特征的非线性转换, 将目标域映射到源域. 最后, 利用 kNN 分类器对映射后的目标域样本进行分类. 与其他小样本集分类器的算法相比, 本文算法考虑了域间转换的学习过程的细节. 以下是本文提出算法的具体描述.

## 2.1 内积度量的边际 Fisher 准则

为了提高类内紧凑性和类间可分性, 本文首先利用根据边际 Fisher 准则将源域的特征映射到新的特征空间中. 本文对原始 MFA<sup>[26]</sup> 方法进行了改进. 改进后的边际 Fisher 准则与原始 MFA 方法的区别在于: 本文方法的距离度量方式是内积, 而不是欧几里得距离. 这是为了优化算法的求解过程, 同时与后面的迁移学习算法中的距离度量方式保持一致, 以保证算法求解的一致性. 边际 Fisher 准则是基于图嵌入的框架, 设计出描述类内紧凑性的本征图和类间区分性的惩罚图, 如图 3 所示.

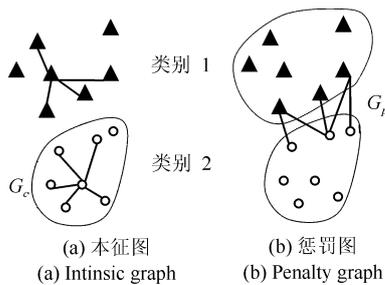


图 3 边际 Fisher 准则的图结构: 本征图和惩罚图

Fig. 3 The graph structure of margin Fisher criterion: intrinsic graph and penalty graph

在本征图  $G_c$  中, 同类点间的邻近关系是由每一个样本与  $k_1$  个与其同类并近邻的样本点的距离之和描述. 在惩罚图  $G_p$  中, 类间边界点邻近关系是由边界奇异点与  $k_2$  个与其异类并近邻的边界点的距离之和描述. 因此本征图  $G_c$  中类内紧凑性  $S_c$  可以表示为

$$S_c = \sum_i \sum_{i \in N_{k_1}(j) \text{ or } j \in N_{k_1}(i)} (M^T \mathbf{x}_i)^T M^T \mathbf{x}_j = \sum_i \sum_{i \in N_{k_1}(j) \text{ or } j \in N_{k_1}(i)} \mathbf{x}_i^T M M^T \mathbf{x}_j =$$

$$\text{tr} \left( M^T \left( \sum_i \sum_{i \in N_{k_1}(j) \text{ or } j \in N_{k_1}(i)} \mathbf{x}_i^T \mathbf{x}_j \right) M \right) \quad (1)$$

其中,  $M$  代表特征变换矩阵,  $N_{k_1}(i)$  表示  $k_1$  个与样本  $\mathbf{x}_i$  同类并与  $\mathbf{x}_i$  近邻的样本点的索引集.

惩罚图  $G_p$  中的类间分离性  $S_p$  可以表示为

$$S_p = \sum_i \sum_{(i,j) \in P_{k_2}(c_i) \text{ or } (i,j) \in N_{k_2}(c_j)} (M^T \mathbf{x}_i)^T M^T \mathbf{x}_j = \sum_i \sum_{(i,j) \in P_{k_2}(c_i) \text{ or } (i,j) \in N_{k_2}(c_j)} \mathbf{x}_i^T M M^T \mathbf{x}_j = \text{tr} \left( M^T \left( \sum_i \sum_{(i,j) \in P_{k_2}(c_i) \text{ or } (i,j) \in N_{k_2}(c_j)} \mathbf{x}_i^T \mathbf{x}_j \right) M \right) \quad (2)$$

其中,  $P_{k_2}(c_i)$  表示  $(i, j), i \in \pi_{c_i}, j \notin \pi_{c_i}$  中  $k_2$  个最邻近的边界样本对集. 根据边界 Fisher 准则, 最小化类内紧凑性和最大化类间分离性, 则特征变换矩阵  $M$  可以利用图嵌入结构得到:

$$M = \arg \min_M \frac{S_c}{S_p} \quad (3)$$

本文算法创新点之一就是本文通过改进传统的边界 Fisher 准则并将其与域间迁移学习相结合, 这样更加充分地利用了源域的标记样本信息. 本文的实验分别对集成了原始 MFA 的算法 (实验部分的 mmf-Euclid 算法) 和集成了改进后的边际 Fisher 准则算法 (实验部分的 mmf 算法) 进行了实验. 我们通过对实验结果可以验证, 集成了改进后的边际 Fisher 准则的算法比集成了原始 MFA 的算法更有效地提高分类器的准确率.

## 2.2 自适应训练样本筛选和域间迁移学习

针对域间特征分布的差异性以及相关性, 本文通过信息矩阵学习的方式, 学习出目标域到源域的映射矩阵. 源域  $A$  和目标源  $B$  的类别数均为  $c$ , 其中  $A$  有  $n_A$  个样本点, 设为  $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_A}]$ ,  $B$  中有  $n_B$  个样本点, 设为  $Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n_B}]$

### 2.2.1 训练样本对自适应筛选

训练样本对的选择会直接影响到训练结果的好坏. 在文献 [16] 提出的 ARC-t 算法, 对于样本对的选择是随机选取的. 但是事实上, 如果样本对恰好选择到源域中的奇异点, 那么目标域就会被映射到对应类别的边界区域, 类间差异性就会变得不明显, 如图 4 所示.

为了避免源域中奇异点对于学习特征转换的影响, 特征映射学习的训练样本对  $(\mathbf{x}_i, \mathbf{y}_j)$  不能随机选

取. 本文其中一个主要的创新点是考虑到域间特征转换学习的效果与训练样本对的质量是紧密相关的, 我们提出了一种通过去除边缘奇异点筛选出训练样本对的算法, 算法详细过程如下:

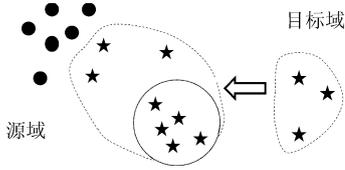


图 4 圆形虚线内是源域的同类样本的密集区, 源域的其他样本点被称为奇异点

Fig. 4 The concentrated region of samples belonging to the same category in the source domain is surrounded by the round dashed line, while the other samples in the source domain are called singular points

1) 计算每一个样本点  $\mathbf{x}_i$  与同类其余样本点间的距离, 记为  $S_{ij}$ , 在本文中, 距离统一都用点积作为度量标准, 即  $S_{ij} = \mathbf{x}_i^T \mathbf{y}_j, i \neq j$ ;

2) 计算  $S_i = \sum_{j=1}^{n^e} S_{ij}$ , 其中  $n^e$  是类别  $e$  中样本数,  $S_i$  表示样本点  $\mathbf{x}_i$  与同类其余样本点间的距离之和;

3) 针对每个  $i (1 \leq i \leq n^e)$  对  $S_i$  按降序进行排序, 倒数  $k_3$  个样本点会被认为是奇异点, 其余样本点组成样本候选集;

4) 从上述候选集中随机均匀选出  $k_A$  个样本点, 该集合记为  $T(A)$ , 这些点将与  $B$  域中的  $k_B$  个标记样本点 (记为  $T(B)$ ) 组成样本对. 随机选取是为了避免将  $B$  域中的样本过分地映射到  $A$  域的中心位置, 而是均匀地映射到  $A$  域的较密集区域.

至此, 域转换学习中所用到的训练样本对的选择结束. 算法最后得到的样本对将作为域间特征转换的训练样本对.

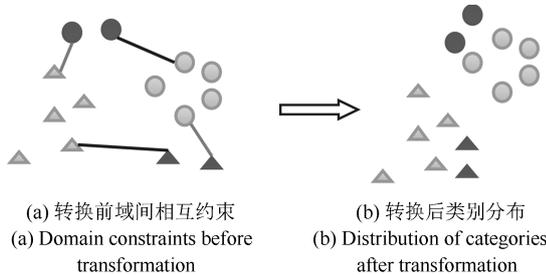


图 5 域间特征转换示意图 (灰色表示源域, 黑色表示目标域)

Fig. 5 The diagrammatic sketch of features transformation between the target domain and source domain (The gray points represent the target domain, while the balck ones represent the source domain.)

## 2.2.2 域间迁移学习

样本间相似性函数可表示为

$$\text{sim}_W(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T W \mathbf{y} \quad (4)$$

其中,  $\mathbf{x} \in A, \mathbf{y} \in B$ , 维数分别为  $d_A$  和  $d_B$ , 本文讨论的是维数相同的情况. 根据样本间的约束关系, 可计算出  $B$  到  $A$  的线性转换  $W$ . 为避免过拟合, 将  $W$  正则化, 记为  $\gamma(W)$ , 该优化问题可以表示为

$$\begin{aligned} \min_W \gamma(W) \\ \text{s. t. } f_i(X^T W Y) \geq 0, \quad 1 \leq i \leq c \end{aligned} \quad (5)$$

其中,  $f_i(X^T W Y)$  是域间的约束条件. 为了简化算法, 可以选择对  $W$  使用 LogDet 正则化. 正则化函数  $\gamma(W)$  可以写成  $W$  的奇异值之和的形式, 即如果  $\sigma_1, \sigma_2, \dots, \sigma_p$  是  $W$  的奇异值, 那么  $\gamma(W)$  可以写成  $\gamma(W) = \sum_{j=1}^p \gamma_j(\sigma_j)$  的形式,  $\gamma_j$  是一个标量函数. 但是这个方法有一个很大的局限性就是使用 LogDet 正则化需要  $W$  是正定矩阵, 为避免这个限制条件, 本文提出用样本之间特征向量的点乘来表示样本间的相似性函数<sup>[16]</sup>, 该问题的目标函数可以用内积的形式表示. 设内积  $K_A = X^T X, K_B = Y^T Y$ , 该问题的最优解  $W = X K_A^{-1/2} L K_B^{-1/2} Y^T$ , 其中  $L$  是一个  $n_A \times n_B$  的矩阵<sup>[12]</sup>. 那么, 该问题的约束函数可以表示为

$$\begin{aligned} f_i(X^T W Y) &= f_i(X^T X K_A^{-\frac{1}{2}} L K_B^{-\frac{1}{2}} Y^T Y) = \\ &= f_i(K_A^{\frac{1}{2}} L K_B^{\frac{1}{2}}) \end{aligned} \quad (6)$$

因此算法的计算过程中只涉及特征向量内积的计算问题.

考虑到类内相似性和类间差异性, 建立对应的约束条件. 假设  $(\mathbf{x}_i, l_i^A)$  是源域  $A$  中的标记样本, 其中  $l_i^A$  是  $\mathbf{x}_i$  的标签;  $(\mathbf{y}_j, l_j^B)$  是目标域  $B$  中的标记样本, 其中  $l_j^B$  是  $\mathbf{y}_j$  的标签. 从  $A$  和  $B$  中分别选出相应的样本对, 则将约束条件变成:

$$\begin{aligned} f_{ij}(\mathbf{x}_i W \mathbf{y}_j) &= \mathbf{x}_i^T W \mathbf{y}_j - l, \quad l_i^A = l_j^B \\ f_{ij}(\mathbf{x}_i W \mathbf{y}_j) &= u - \mathbf{x}_i^T W \mathbf{y}_j, \quad l_i^A \neq l_j^B \end{aligned} \quad (7)$$

其中,  $l$  和  $u$  分别是两个样本点相似性的上限和下限参数. 对于同一类别的样本对来说, 相似性应该更大; 对于不同类别的样本对来说, 相似性应该更小.

综上, 域间迁移学习的过程可以表述如下. 设  $\sigma_1, \sigma_2, \dots, \sigma_p$  是  $W$  的奇异值,  $\gamma_j$  是一个标量函数, 本文采用  $\gamma_j(\sigma_j) = \sigma_j^2/2$  的标量函数, 那么  $\gamma(W) = \frac{1}{2} \|W\|_F^2$  就是一个弗罗贝尼乌斯范数形

式. 则该迁移学习问题的求解可以用以下形式描述:

$$\begin{aligned} & \min_W \sum_{j=1}^p \gamma_j(\sigma_j) \\ & \text{s. t. } \mathbf{x}'_i \in T(A) \\ & \quad \mathbf{x}'_i{}^T W \mathbf{y}'_j - l \geq 0, \quad l_i^A = l_j^B \\ & \quad u - \mathbf{x}'_i{}^T W \mathbf{y}'_j \geq 0, \quad l_i^A \neq l_j^B \end{aligned} \quad (8)$$

其中,  $\mathbf{x}'_i = M \mathbf{x}_i$ ,  $\mathbf{y}'_j = M \mathbf{y}_j$ . 由于该问题是一个严格凸优化问题, 该凸优化问题的松弛系数记为  $\lambda$ , 则该问题可进一步表示为

$$\begin{aligned} & \min_W \sum_{j=1}^p \gamma_j(\sigma_j) + \lambda \sum_i \sum_j f_{ij}(\mathbf{x}'_i{}^T W \mathbf{y}'_j) \\ & \text{s. t. } \mathbf{x}'_i{}^T \in T(A), \quad \mathbf{x}'_j \in T(B) \end{aligned} \quad (9)$$

线性转换有时并不能很好地表达出域间的映射关系, 而非线性转换则能较好地表达, 所以算法可以引入特征的核化空间. 该算法的核化就是将算法中出现特征向量内积的地方用相应的核化函数表示. 当引入原始核函数  $k(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})^T \phi(\mathbf{y})$ , 核化之后  $f_{ij}(\phi(\mathbf{x}), \phi(\mathbf{y})) = \phi(\mathbf{x})^T W \phi(\mathbf{y})$ , 则  $W$  表示的是在希尔伯特 (Hilbert) 空间中的一个算子, 维度与  $\phi(\mathbf{x})$  和  $\phi(\mathbf{y})$  的维度相关. 因此, 如果引入的原始核函数为径向基函数 (Radial basis function, RBF) 核, 则  $W$  的维数就是无穷大. 尽管不能显式计算  $W$ , 但是仍然可以根据文献 [23] 中的核化算法将  $W$  递归展开成如下形式:

$$W = I + \sum_{i,j} \sigma_{ij} \phi(\mathbf{x}_i) \phi(\mathbf{y}_j)^T \quad (10)$$

其中,  $(\mathbf{x}_i, \mathbf{y}_j)$  为约束训练样本对,  $\sigma_{ij}$  为 Bregman 投影系数 (Bregman projection parameter). 此时我们可定义一个新的核函数  $\tilde{k}(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})^T W \phi(\mathbf{y})$ . 则核化之后的相似性函数可以用新引入的核化函数表示, 即  $f_{ij}(\phi(\mathbf{x}), \phi(\mathbf{y})) = \tilde{k}(\mathbf{x}, \mathbf{y})$ , 因此算法中出现特征向量内积的地方用相应的核函数表示, 算法的核化就完成了. 该凸优化问题有很多方法可以求解, 本文采用 Bregman 迭代算法进行求解, 可以很快得到收敛解, 具体求解过程可以参考文献 [27].

### 3 实验结果

本文提出的算法采用 Matlab 编程实现.

#### 3.1 数据库

实验所用数据库<sup>[3]</sup>包含了三个子数据域, 分别是 webcam, amazon 和 dsrl. 在这三个域中分别有 31 个物体类别, 涵盖了生活中常见的物体, 例如椅

子、背包、键盘等. 这三个数据域相互之间是有差异的. 首先, webcam 域中的图片是带有闪光灯的摄像头拍摄得到的, 具有姿态各异, 低分辨率 ( $640 \times 480$ ), 背景变化等特点, 每一类别中有 5 个不一样的对象且每个对象平均有 3 个不同的视角, webcam 域共有 795 张图片. 其次, amazon 域中的图片是从网上收集的, 具有画面单一、物体居中、白色背景等特点, 每一类别中平均有 90 张图片, 共计 2813 张图片. 最后, dsrl 域中的图片是由数码单镜头反光 (Single lens reflex, SLR) 相机在自然环境中拍摄得到的, 具有背景多样, 高分辨率  $4288 \times 2848$ , 自然光线的特点, 每一类别中有 5 个不同对象, 每个对象平均有 3 个不同的视角. dsrl 域共有 498 张图片.

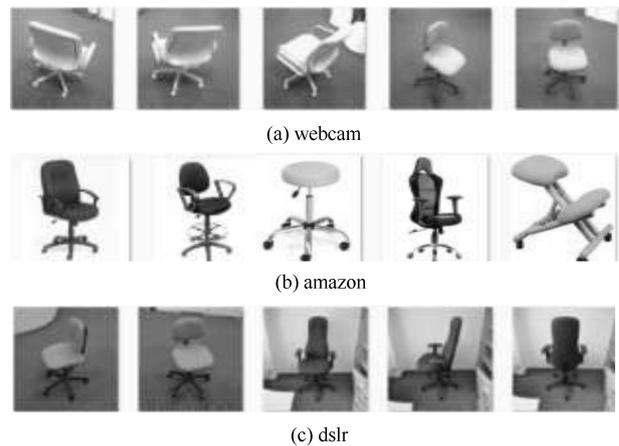


图 6 webcam, amazon 和 dsrl 中椅子类别中的样例图  
Fig. 6 The samples image of chair in three domain:  
webcam, amazon, dsrl

#### 3.2 本文算法的具体实现和参数

本文提出算法 (记为 mmf) 实验的具体实现方案如下. 如果每个算法采用的图像特征不是统一的, 那么很难对这些算法做比较. 为了能更好地与前人取得的成果做对比, 本实验中所用到的图像特征均采用文献 [3] 中提供的 SURF (Speeded up robust features) 特征, 将所有图片灰度化并提取其 SURF 特征, 最后被量化成 800 维的特征. 为保证各对比算法实验结果的可比性, 所有算法中的参数  $K_A$  和  $K_B$  均保持一致. 从  $A$  中选取  $K_A$  个样本点 (当  $A$  域为 amazon 时, 取  $K_A = 20$ , 当  $A$  域为 webcam 或 dsrl 时, 取  $K_A = 15$ ), 与  $B$  域中随机选择的 3 个样本点组成样本对, 取  $K_B = 3$ . 首先训练出映射矩阵  $M$ , 本文借鉴了对比方法中采用的核函数, 利用的是高斯 RBF 核函数, 其中  $\sigma = 1$ ,  $\lambda = 10^3$ . 通过迭代学习出域间的转换  $W$ , 本实验采用 kNN 分类器, 其中  $k = 1$ , 通过 20 次的随机实验, 最终取所有

实验的分类准确率的平均值作为本实验的平均分类准确率. 具体的实验参数设置如表 1 所示.

表 1 本文算法的实验参数设置

Table 1 The experiment parameters set of our method

$A \rightarrow B$	$k_1$	$k_2$	$k_A$	$k_B$	$\sigma$
$a \rightarrow w$	22	1700	20	3	1
$a \rightarrow d$	22	1700	20	3	1
$w \rightarrow a$	20	680	15	3	1
$w \rightarrow d$	20	680	15	3	1
$d \rightarrow a$	17	420	15	3	1
$d \rightarrow w$	17	420	15	3	1

### 3.3 实验结果及分析

本文方法与近年来具有代表性的一些相关研究成果进行了实验对比. 基于深度学习的算法需要源域是目标域的父集<sup>[15]</sup> 并且需要海量训练数据保证其准确率. 考虑到本文实验所用的数据库没有其对应的大样本库且其最小的数据域的样本仅有 498 个, 我们没有加入基于深度学习的算法对比. 对比实验中的基于 SVM 的算法和基于域间特征转换的算法描述如下:

1) kNN-ab 是在  $A$  域中的标记样本上训练出 kNN 分类器, 直接将  $B$  域中的测试数据用在该 kNN 分类器上计算分类的准确率.

2) kNN-bb 是将  $B$  域中的标记样本和  $A$  域中的标记样本作为训练样本训练出 kNN 分类器. 其余样本作为测试样本计算分类准确率.

3) symm 是 Saenko 等提出的算法<sup>[3]</sup>. 该算法主要通过信息论矩阵学习实现域间转换并训练出 kNN 分类器来提高目标域上的分类准确率.

4) ARC-t 是 Kulis 等提出的算法<sup>[16]</sup>. 该算法是利用域间的非线性转换的学习解决域间差异性问题的, 结合源域训练的 kNN 分类器对目标域样本进行分类.

5) gfk 是文献 [20] 中提出的算法, 是通过基于测地流核 (Geodesic flow kernel) 的迁移学习来训练目标域上的 kNN 分类器.

6) svm-s 是利用  $A$  域中的标记样本训练出的 SVM 分类器.

7) hfa 是文献 [28] 提出的, 该方法提出了一种新的边缘最大化的特征表示和 svm 损失函数相结合的算法, 最终通过 SVM 分类器达到提高目标域的分类准确度的目的.

表 2 给出了本文算法和其他算法的对比结果. 其中 mmf-Euclid 是集成了原始 MFA 的分类器算法, mmf 是集成了本文改进后的边际 Fisher 准则的分类器算法. 从实验结果可以看出, mmf 的分类准确度要比 mmf-Euclid 更高, 因此本文采用改进后的边际 Fisher 准则比原始 MFA 更有助于提高本文的分类器算法的准确率. 与 kNN-ab 和 kNN-bb 的实验结果相比较, 本文提出的算法通过将目标域的样本映射到源域的学习过程而不是简单地融合了两个域的标记样本数据, 有效提高了目标域上的分类准确率. 从基于 SVM 的算法和基于域间特征转换的算法的对比实验结果中, 我们可以看出本文的算法充分考虑了源域和目标域的特征分布及其差异性, 通过对源域和目标域的特征映射的学习过程使得目标域的分类准确率上有提高, 并且在平均的识别准确率上是最高的.

最后, 我们不仅验证了本文提出的算法能够获得高准确率, 而且在图 7 中还验证了该算法的性能随着目标域可用标记样本数的增大而提升. 源域中的训练样本数为 20, 目标域中可用标记的样本数从 1 到 5 变化. 从图 7 可以看到这 5 种方法都是递增的趋势, 本文的方法在整体上要比其他方法好.

综上, 与其他算法相比, 本文提出的小样本集的分类器设计算法能够更充分考虑到源域中不同类间的可分离性、同类间的紧凑性、域间特征分布的

表 2 在三个数据域上的分类准确率 (%) (加粗字体表示最佳性能, 缩写:  $a$ : amazom,  $w$ : webcam,  $d$ : dsrlr)

Table 2 Accuracy rates in the three domains (%) (The bold font represents the best performance, abbreviation:  $a$ : amazom,  $w$ : webcam,  $d$ : dsrlr.)

$A \rightarrow B$	kNN-ab	kNN-bb	symm <sup>[3]</sup>	ARCt <sup>[16]</sup>	gfk <sup>[20]</sup>	svm-s	hfa <sup>[28]</sup>	mmf-Euclid	mmf
$a \rightarrow w$	9.6 ± 1.0	51.0 ± 0.8	51.0 ± 1.4	55.7 ± 0.9	57.8 ± 1.0	34.5 ± 0.8	<b>61.5 ± 0.9</b>	49.3 ± 0.8	55.5 ± 0.7
$a \rightarrow d$	4.9 ± 1.1	47.9 ± 0.9	47.9 ± 1.4	50.2 ± 0.7	50.5 ± 0.8	35.3 ± 0.6	52.4 ± 1.0	50.6 ± 0.9	<b>57.6 ± 0.9</b>
$w \rightarrow a$	10.5 ± 0.6	40.1 ± 0.5	43.7 ± 0.7	43.4 ± 0.5	44.1 ± 0.4	34.9 ± 0.4	44.5 ± 0.7	42.9 ± 0.6	<b>44.6 ± 0.8</b>
$w \rightarrow d$	23.2 ± 0.8	54.1 ± 0.9	69.8 ± 1.0	71.3 ± 0.8	68.5 ± 0.5	65.8 ± 0.8	52.7 ± 1.1	70.1 ± 0.7	<b>71.5 ± 0.6</b>
$d \rightarrow a$	11.3 ± 0.5	35.6 ± 0.7	42.7 ± 0.5	42.5 ± 0.5	<b>45.7 ± 0.8</b>	33.8 ± 0.4	45.4 ± 0.9	41.4 ± 0.7	42.6 ± 0.6
$d \rightarrow w$	37.6 ± 0.8	54.0 ± 0.7	63.4 ± 0.9	65.3 ± 0.5	66.4 ± 0.5	68.1 ± 0.6	62.4 ± 0.8	68.5 ± 1.0	<b>69.5 ± 0.7</b>
平均准确率	12.2 ± 0.8	37.8 ± 0.8	55.6 ± 1.0	56.9 ± 0.7	57.7 ± 0.7	45.4 ± 0.6	53.2 ± 0.9	55.4 ± 0.8	<b>58.6 ± 0.8</b>

差异性以及域间特征转换学习中学习样本对的选择过程,并实现了目标到源域特征适应性更好的非线性转换.通过目标域中少量标记样本特征的转换学习,随着标记样本数的增加,学习的分类器性能就越好.

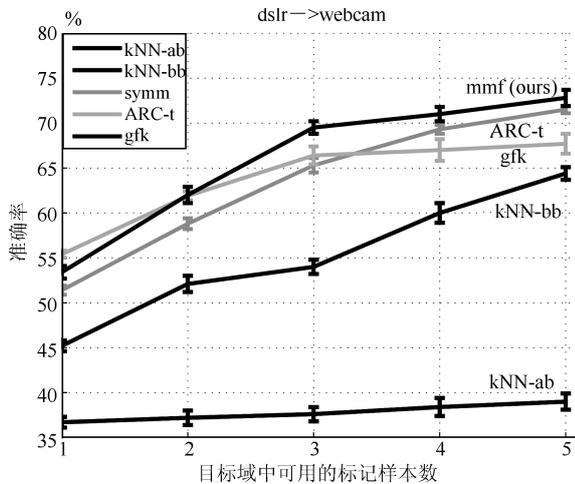


图7 在目标域上的分类准确率随目标域中可用标记样本数量的变化曲线(其中源域中选择了20个标记样本,源域为dslr,目标域为webcam)

Fig. 7 The accuracy rate curves in the target domain varying with the number of labeled samples in the target domain (Where 20 labeled samples in source domain is selected, the source domain here is dslr, while the target domain is webcam.)

## 4 结论

本文提出了一种新的基于边际 Fisher 准则和迁移学习的小样本集的分类器设计算法.通过对源域数据的特征分布的优化和目标域到源域的迁移学习,实现了利用大量标记同构样本集来提升小样本集上分类准确率.与最近针对小样本数据集上的分类器设计问题的算法进行对比,本文提出的算法不仅在小样本数据集上的分类器准确率有所提高,并且随着小样本集中的标记样本数量的增加,本文提出的分类器设计算法的性能提升得也更明显.

实验结果虽然证实了本文算法的有效性,但依然有待改进的地方.其中一个是可以将该算法在边际 Fisher 准则的基础上引入核化函数;另外一点是实验中用的是 800 维的 SURF 特征,不排除有更好的特征表示方法.因此,以上两点将是以后工作研究的重点.

## References

- Cristianini N, Shawe-Taylor J, Kandola J S. Spectral kernel methods for clustering. In: Proceedings of the 2001 Advances in Neural Information Processing Systems 14 (NIPS01). London: MIT, 2002. 649–655

- Bosch A, Zisserman A, Munoz X. Representing shape with a spatial pyramid kernel. In: Proceedings of the 6th ACM International Conference on Image & Video Retrieval. New York: ACM, 2007. 401–408
- Saenko K, Kulis B, Fritz M, Darrell T. Adapting visual category models to new domains. In: Proceedings of the 11th European Conference on Computer Vision, Lecture Notes in Computer Science. Heraklion, Crete, Greece: Springer, 2010. 213–226
- Patel V M, Gopalan R, Li R, Chellappa R. Visual domain adaptation: a survey of recent advances. *IEEE Signal Processing Magazine*, 2015, **32**(3): 53–69
- Hoffman J, Rodner E, Donahue J, Darrell T, Saenko K. Efficient learning of domain-invariant image representations. arXiv: 1301.3224, 2013.
- Shao L, Zhu F, Li X. Transfer learning for visual categorization: a survey. *IEEE Transactions on Neural Networks & Learning Systems*, 2015, **26**(5): 1019–1034
- Gu Xin, Wang Shi-Tong, Xu Min. A new cross-multidomain classification algorithm and its fast version for large datasets. *Acta Automatica Sinica*, 2014, **40**(3): 531–547 (顾鑫, 王士同, 许敏. 基于多源的跨领域数据分类快速新算法. *自动化学报*, 2014, **40**(3): 531–547)
- Zhang Qian, Li Ming, Wang Xue-Song, Cheng Yu-Hu, Zhu Mei-Qiang. Instance-based transfer learning for multi-source domains. *Acta Automatica Sinica*, 2014, **40**(6): 1176–1183 (张倩, 李明, 王雪松, 程玉虎, 朱美强. 一种面向多源领域的实例迁移学习. *自动化学报*, 2014, **40**(6): 1176–1183)
- Yang J, Yan R, Hauptmann A G. Cross-domain video concept detection using adaptive SVMS. In: Proceedings of the 15th ACM International Conference on Multimedia. New York: ACM, 2007. 188–197
- Wang Xue-Song, Pan Jie, Cheng Yu-Hu, Cao Ge. Self-adaptive transfer for decision trees based on similarity metric. *Acta Automatica Sinica*, 2013, **39**(12): 2186–2192 (王雪松, 潘杰, 程玉虎, 曹戈. 基于相似度衡量的决策树自适应迁移. *自动化学报*, 2013, **39**(12): 2186–2192)
- Bergamo A, Torresani L. Exploiting weakly-labeled web images to improve object classification: a domain adaptation approach. In: Proceedings of the Advances in Neural Information Processing Systems 23. Vancouver, British Columbia, Canada: Curran Associates, Inc., 2010. 181–189
- Li X. Regularized Adaptation: Theory, Algorithms and Applications [Ph.D. dissertation], University of Washington, USA, 2007
- Dong Ai-Mei, Wang Shi-Tong. A shared latent subspace transfer learning algorithm using SVM. *Acta Automatica Sinica*, 2014, **40**(10): 2276–2287 (董爱美, 王士同. 共享隐空间迁移 SVM. *自动化学报*, 2014, **40**(10): 2276–2287)
- Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, OH: IEEE, 2014. 1717–1724
- Razavian A S, Azizpour H, Sullivan J, Carlsson S. CNN features off-the-shelf: an astounding baseline for recognition. *Computer Vision & Pattern Recognition Workshops*, 2014. 512–519
- Kulis B, Saenko K, Darrell T. What you saw is not what you get: domain adaptation using asymmetric kernel transforms. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, RI: IEEE, 2011. 1785–1792

- 17 Farhadi A, Tabrizi M K. Learning to recognize activities from the wrong view point. In: Proceedings of the 10th European Conference on Computer Vision — ECCV 2008. Marseille, France: Springer, 2008. 154–166
- 18 Chopra S, Balakrishnan S, Gopalan R. DLID: deep learning for domain adaptation by interpolating between domains. In: Proceedings of the ICML Workshop on Representation Learning. Atlanta, Georgia, USA, 2013.
- 19 Wang H, Nie F P, Huang H, Ding C. Dyadic transfer learning for cross-domain image classification. In: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona: IEEE, 2011. 551–556
- 20 Gong B, Shi Y, Sha F, Grauman, K. Geodesic flow kernel for unsupervised domain adaptation. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI: IEEE, 2012. 2066–2073
- 21 Long M, Wang J, Sun J, Yu P S. Domain invariant transfer kernel learning. *IEEE Transactions on Knowledge & Data Engineering*, 2015, **27**(6): 1519–1532
- 22 Shao L, Liu L, Li X L. Feature learning for image classification via multiobjective genetic programming. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(7): 1359–1371
- 23 Davis J V, Kulis B, Jain P, Sra S, Dhillon I S. Information-theoretic metric learning. In: Proceedings of the 24th International Conference on Machine Learning. Corvallis, Oregon, USA: ACM, 2007. 209–216
- 24 Yan S C, Xu D, Zhang B Y, Zhang H J, Yang Q, Lin S. Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(1): 40–51
- 25 Li Y L, Liu P, Du H S, Li Z, Liu J H, Yu D Y, Li M Q. Marginal fisher analysis-based feature extraction for identification of drug and explosive concealed by body packing. *Analytical Methods*, 2013, **5**(22): 6331–6337
- 26 Wang Z Q, Sun X. Optimal kernel marginal fisher analysis for face recognition. *Journal of Computers*, 2012, **7**(9): 2298–2305
- 27 Al Censor Y, Zenios S A. Parallel optimization: theory, algorithms and applications. *Scalable Computing: Practice & Experience*. New York: Oxford University Press, 1997.
- 28 Duan L X, Xu D, Tsang I W. Learning with augmented features for heterogeneous domain adaptation. In: Proceedings of the 29th International Conference on Machine Learning. Edinburgh, Scotland, UK: Omnipress, 2012. 711–718



**舒 醒** 浙江大学信息与电子工程学院硕士研究生. 主要研究方向为计算机视觉, 图像分类.

E-mail: 21331093@zju.edu.cn

(**SHU Xing** Master student at the College of Information Science and Electronic Engineering, Zhejiang University. Her research interest covers

computer vision and image classification.)



**于慧敏** 浙江大学信息与电子工程学院教授. 主要研究方向为医学图像处理, 计算机视觉. 本文通信作者.

E-mail: yhm2005@zju.edu.cn

(**YU Hui-Min** Professor at the College of Information Science and Electronic Engineering, Zhejiang University. His research interest covers medical image processing and computer vision. Corresponding author of this paper.)



**郑伟伟** 浙江大学信息与电子工程学院博士研究生. 主要研究方向为图像匹配, 跟踪技术.

E-mail: 3090102748@zju.edu.cn

(**ZHENG Wei-Wei** Ph.D. candidate at the College of Information Science and Electronic Engineering, Zhejiang University. His research interest

covers object matching and target tracking.)



**谢 奕** 浙江大学信息与电子工程学院博士研究生. 主要研究方向为图像匹配, 跟踪技术.

E-mail: Yixie@zju.edu.cn

(**XIE Yi** Ph.D. candidate at the College of Information Science and Electronic Engineering, Zhejiang University. His research interest covers object matching and target tracking.)



**胡浩基** 浙江大学信息与电子工程学院副教授. 主要研究方向为图像处理, 计算机视觉. E-mail: haoji\_hu@zju.edu.cn

(**HU Hao-Ji** Assistant professor at the College of Information Science and Electronic Engineering, Zhejiang University. His research interest covers image processing and computer vision.)



**唐慧明** 浙江大学信息与电子工程学院副教授. 主要研究方向为图像识别, 计算机视觉.

E-mail: tanghm@isee.zju.edu.cn

(**TANG Hui-Ming** Assistant professor at the College of Information Science and Electronic Engineering, Zhejiang University. His research interest

covers image recognition and computer vision.)