

## 基于二次表示的空间目标图像分类

蒋飞云<sup>\*①</sup> 孙锐<sup>①②</sup> 张旭东<sup>①</sup> 李超<sup>①</sup>

<sup>①</sup>(合肥工业大学计算机与信息学院 合肥 230009)

<sup>②</sup>(奇瑞汽车博士后工作站 芜湖 241009)

**摘要:** 针对空间目标图像的特点, 该文提出一种基于局部不变特征的空间目标图像分类方法。该方法首先提取每幅图像的局部不变特征, 利用混合高斯模型(GMM)建立全局的视觉模式, 然后依据最大后验概率匹配局部特征和视觉模式来构造整个训练集图像的共现矩阵, 采用概率潜在语义分析(PLSA)模型得到图像的潜在类别表示来实现图像的二次表示, 最后利用 SVM 算法实现分类。实验结果验证了该方案的有效性。

**关键词:** 空间目标分类; 局部不变特征; 视觉模式; 二次表示

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2013)05-1247-05

DOI: 10.3724/SP.J.1146.2012.01289

## Space Target Image Categorization Based on the Second Representation

Jiang Fei-yun<sup>①</sup> Sun Rui<sup>①②</sup> Zhang Xu-dong<sup>①</sup> Li Chao<sup>①</sup>

<sup>①</sup>(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

<sup>②</sup>(The Chery Car Postdoctoral Workstation, Wuhu 241009, China)

**Abstract:** According to the characteristics of space target image, an novel method of space target image categorization based on local invariant features is proposed. The method extracts firstly local invariant features of each image and uses Gaussian Mixture Model (GMM) to establish global visual modes. Then co-occurrence matrix of the entire training set is constructed by matching local invariant features and visual models with maximum a posteriori probability and Probability Latent Semantic Analysis (PLSA) model is used to obtain latent class vector of images to achieve second representation. Finally, the SVM algorithm is used to implement image categorization. The experimental result demonstrates the effectiveness of the proposed method.

**Key words:** Space target categorization; Local invariant features; Visual mode; Second representation

### 1 引言

随着越来越多的平台被送入太空, 空间目标识别技术变得越来越重要, 它是人类研究空间、控制空间的基础, 对国家安全具有重要的意义。空间目标主要是指卫星, 随着天基光学系统对空间目标具备逐步成像的能力, 这为从光学图像角度识别空间目标提供了可靠条件。天基光学系统目标成像, 除了与目标本身的外形、材料成分有关外, 还具有以下几个特点<sup>[1]</sup>: (1)目标运行时姿态呈现一定的规律性变化; (2)目标所在区域的光照条件不稳定; (3)由于传感器相机视角的变化引起图像视角以及仿射变化; (4)目标在由远到近和运动状态的实际条件下, 造成目标在图像中表现出尺度、旋转、遮挡等变化, 另外空间目标分类是一种类内的分类, 各类

图像间存在着相似性和同义性, 这些都给空间目标识别带来很大难度。

在已有的方法中, 文献[2]根据空间目标的不变矩、灰度以及轮廓特征, 提取出目标的特征向量, 并利用一种基于Voronoi距离的核聚类算法, 通过距离度量实现目标分类识别。文献[3]的方法则是从空间目标图像序列中提取出每幅图像的小波矩函数, 然后利用DS证据理论(Dempster-Shafer evidence theory)结合这些序列图像的特征向量, 最后利用神经网络得到识别结果。以上方法虽然都实现了一定的分类效果, 但是空间目标成像有其自身的特点, 利用上述特征描述往往需要特定的目标区域, 而当图像存在遮挡、视角、光照模糊、仿射等变化时, 要想从图像上准确地分割目标往往是很困难的。Ren等人<sup>[4]</sup>利用压缩感知(Compressive Sensing, CS)的原理, 该方法首先利用训练图像得到一个稀疏的表示字典, 其次通过测试矩阵得到一个关于测试图像的线性判别函数进行识别, 但是稀疏表示缺乏有

2012-10-10 收到, 2012-12-28 改回

安徽省自然科学基金(11040606M149)资助课题

\*通信作者: 蒋飞云 317033310@qq.com

效的判别信息。Cao等人<sup>[5]</sup>采用基于克利福德代数(Clifford Algebra, CA)的仿生模式识别来实现空间目标识别, 该方法是通过克利福德代数将空间目标图像扩展到高维的向量空间进行属性识别, 这种方法向量维数高, 忽略了图像特征的作用。

与已有的方法不同, 本文研究基于局部特征的空间目标图像分类方法, 结合特征统计建模和分解得到图像的最终表示实现最后的分类。

## 2 基于二次表示的空间目标分类方法

本文针对空间目标图像, 使用高斯混合模型(GMM)<sup>[6]</sup>进行全局的图像局部特征建模, 通过计算局部特征相对于视觉模式(GMM 中的单高斯模式)的后验概率, 依据最大后验概率准则进行二者的匹配, 统计视觉模式在图像上出现的次数, 构造整个训练集图像的共现矩阵, 对共现矩阵使用概率潜在语义分析(PLSA)<sup>[7]</sup>得到每个图像的二次表示来处理空间目标图像的相似性和同义性问题, 最后使用支持向量机的方法进行分类, 图 1 给出了算法的主要流程。

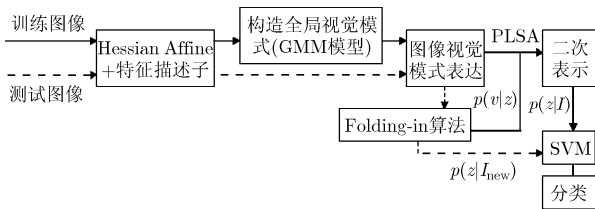


图 1 本文算法流程框图

本文的方法主要有以下几个特点: (1)采用 GMM 模型代替 k-means 进行聚类, 局部特征与视觉模式的匹配不是按照简单的欧式距离, 而是按照概率大小方式<sup>[8]</sup>, 每个局部特征针对不同的视觉模式能够得到更多的判别信息。(2)共现矩阵以概率匹配通过计数方式得到的, 视觉模式在图像中出现次数潜在反应了图像目标的类别, 这种计数实际上反映了一种统计信息, 文献[9]表明这种统计信息在物体分类识别中是一种很重要的因素。(3)PLSA 是一种层次结构 Bayes 模型, 通过建立目标和图像之间的概率映射关系, 寻求同一类别的目标在特征空间中的统一概率分布, 选取最大概率值对应的类别作为图像的代表, 由于空间目标图像的分类是一种类内的分类, 不同类别的空间目标图像往往有一些较为明显的相似之处(颜色、外形等), 利用 PLSA 从潜在反应类别的共现矩阵中提取类别信息可以很好地解决空间目标图像存在的相似性和同义性的问题<sup>[10]</sup>。

## 2.1 局部不变特征提取

SIFT(Scale-Invariant Feature Transform), MSER(Maximally Stable Extremal Regions), HOG(Histograms of Oriented Gradients)等是一些常用的局部特征, 但 SIFT 不具备仿射不变性, 旋转不变性存在偏差<sup>[11]</sup>, 并且维数很高; MSER 特征所检测的特征点较少; HOG 特征较多运用在行人等直立物体的检测上, 不具备旋转和尺度不变性。本文针对以上不足, 结合空间目标成像特点, 利用 Hessian-Affine 检测图像中的仿射区域, 对椭圆形的仿射区域进行归一化为圆形区域, 按照旋转不变的梯度方法(如图 2 所示)计算特征区域内的梯度信息; 对特征区域按照灰度大小顺序划分(如图 3 所示), 有序的空间灰度分布能够处理非线性的亮度变化, 对光照模糊具有很好的抵抗性<sup>[12,13]</sup>; 最后合并所有灰度区间内的梯度朝向直方图形成特征向量, 同样具备尺度、旋转、光照、仿射等不变性, 而且维数低。

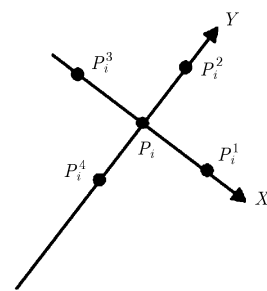


图 2 特征区域坐标轴

图 2 所示一个关于模板点  $P_i$  的局部  $X - Y$  坐标系,  $P_i$  是特征区域的一个像素点。无论图像如何旋转,  $P_i$  点的梯度朝向相对于该坐标系永远不变, 因此梯度方向从理论上说完全是旋转不变的。梯度信息的计算公式如式(1), 式(2)所示:

$$D_x(P_i) = I(P_i^1) - I(P_i^3) \tag{1}$$

$$D_y(P_i) = I(P_i^2) - I(P_i^4) \tag{2}$$

这里  $P_i^j (j=1,2,3,4)$  是  $P_i$  在两坐标轴上的 4 个近邻点,  $I(P_i^j)$  代表点  $P_i^j$  的灰度。



图 3 局部区域的灰度空间划分

本文中将梯度方向划分为 8 个单元, 特征区域按照灰度大小顺序划分为 6 个单元, 统计每个灰度区间内的梯度和朝向, 形成  $8 \times 6 = 48$  维的特征描述子, 如图 4 所示。

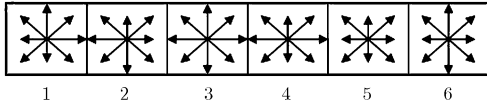


图 4 特征描述子构造

### 2.2 图像二次表示

本文选用 GMM 估计局部特征的统计分布，将 GMM 的一个单独模式看成视觉模式，这样全局的 GMM 描述了各类图像特征可能存在的视觉模式。给定图像  $I$  以及通过 GMM 生成的  $r$  个视觉模式  $v = \{v_1, v_2, \dots, v_r\}$ ，每个模式  $v_i$  对应一个高斯模式  $N(\mu_i, \Sigma_i)$ ，根据贝叶斯公式估计局部特征  $x_l$  属于视觉模式  $v_i$  的后验概率：

$$p(v_i | x_l) = \frac{N\left(x_l | \mu_i, \Sigma_i\right)}{\sum_{k=1}^r N\left(x_l | \mu_k, \Sigma_k\right)} \quad (3)$$

根据式(3)可以将局部特征  $x_l$  匹配到具有最大后验概率值的视觉模式  $v_i$ ，进而得出图像  $I_t$  上全部与  $v_i$  匹配的局部特征的个数  $q$ ，则  $v_i$  在图像  $I_t$  上出现的次数就记为  $q_{it}$ ，相应地可以得出所有视觉模式在图像  $I_t$  上出现的次数。则图像  $I_t$  就可以表示成： $w = [q_{1t}, q_{2t}, q_{3t}, \dots, q_{rt}]^T$ ，从而构造整个训练图像集的共现矩阵，如图 5 所示。

$$\begin{pmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{r1} & q_{r2} & \dots & q_{rn} \end{pmatrix}$$

图 5 共现矩阵

空间目标图像分类是一种类内的分类，不同类别的图像之间存在一些较为明显的相似之处(颜色、外形等)，上述所述的图像表示仅仅借助了底层特征表示，对于图像间存在的相似性和同义性判别性较差，为了处理在类内存在的相似性和同义性问题，因此在保留图像底层特征的基础上采用 PLSA 进一步推导其潜在类别表示。对于 PLSA 模型可以用图像的语法描述：一幅图像可以看成是主题变量  $z$  的混合生成模型， $z$  是一个  $k$  维的向量  $\{z_1, z_2, \dots, z_k\}$ ， $k$  表示图像目标类别数。利用 PLSA 对共现矩阵(图 5)通过最大对数似然函数式(4)估计进行参数迭代。

$$L = \sum_{i=1}^r \sum_{j=1}^n d_{ij} \log p(I_j, v_i) \quad (4)$$

得到  $p(z_u | I_t)$  和  $p(v_i | z_u)$  两个参数，图像则可以看成为条件概率  $p(z_u | I_t)$  的混合， $u = 1, 2, \dots, k$ ，我

们使用条件概率组合作为图像的二次表示，即

$$a(I) = \{p(z_1 | I_t), p(z_2 | I_t), \dots, p(z_k | I_t)\} \quad (5)$$

二次表示具有以下几个特点：(1)图像表示更具判别性：局部特征与视觉模式按照最大概率匹配，共现矩阵使用统计模型计数，潜在反应了图像目标的类别，通过使用PLSA分解抽取类别信息，得到的  $a(I)$  建立了所有目标类别与图像的映射关系，与图像具有相同类别的条件概率  $p(z_u | I_j)$  的值较大，反之就较小<sup>[14]</sup>，可处理类内图像分类存在的相似性和同义性，提高了图像表示的判别性，同时实现了降维。(2)由于PLSA的条件独立假设，因此  $p(v_i | z_u)$  这个分布独立于具体的训练图像，对训练集外的图像也可以适用，因此对于一幅未知的图像  $I_{new}$ ，可以通过视觉模式表示  $w$  和训练得到的参数  $p(v_i | z_u)$ ，使用Folding-in算法就可得到新图像的向量表示  $a(I_{new})$ ；最后采用SVM算法进行分类识别。

## 3 实验结果与分析

### 3.1 实验结果

实验将 3 ds 文件格式的卫星模型直接导入 3 ds max 软件进行仿真，得到了 10 类空间目标仿真模型的图像(如图 6 所示)，每类图像都存在模糊、噪声、视角变化等不利因素，各类目标含 400 幅图像，每类图像随机被分成两个部分，其中 300 幅用来作为训练，其余的用作测试图像，流程按照图 1 所示；局部特征的检测如图 7 所示。视觉模式是通过训练图像得到的，一旦得到就是固定的，其个数  $s$  及其参数都是固定不变的。

实验采用的分类器是一对一(one against one)规则的SVM分类器，核函数采用径向基函数。实验中需要考虑的因素有两个：(1)高斯混合模型的成分个数  $s$ ；(2)PLSA分解中  $k$  的取值。

模型成分个数的确定对分类有很重要的影响，本文将最小描述长度的准则(Minimum Description Length, MDL)<sup>[15]</sup>和GMM结合。在建立全局的视觉模式时，通过统计每类图像局部特征个数的均值(如图8所示)，我们将视觉模式的个数  $s$  搜索范围确定在  $[100, 250]$  的区间上，为了精确地得到最终的全局视觉模式个数  $s$ ，我们选择1为间隔依次迭代选择最佳的高斯成分个数，间隔越小计算精度就越高。在实验室中，根据上述条件限制，计算出Rissanen最小对应的  $s$  值为152，得到152个不同均值和方差的视觉模式。

由于PLSA模型中  $k$  的取值与实际图像的表达密切相关，鉴于  $k$  的取值与实际具体类别数目相关，因此本文通过实验的方法来寻找一个最佳最合理的  $k$  值来进行图像的二次表示，理想的情况应是

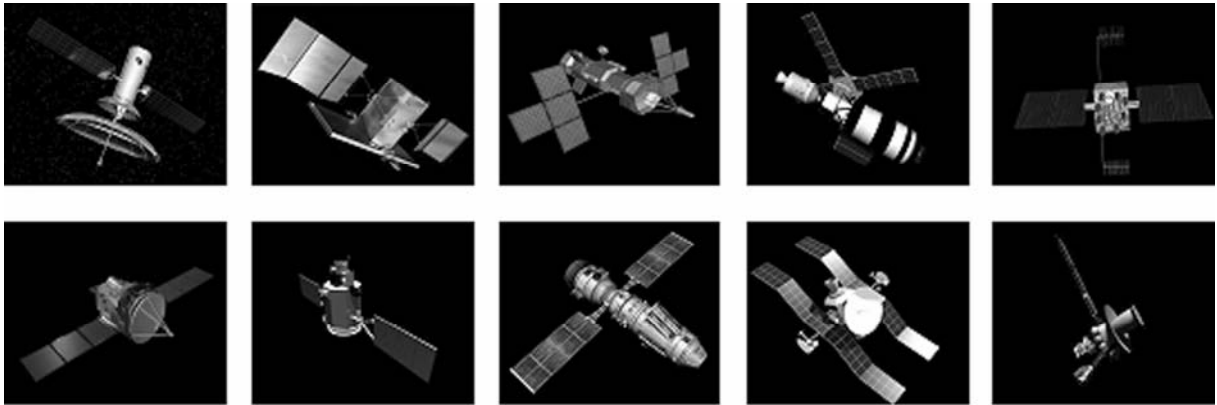


图6 10类空间目标代表图

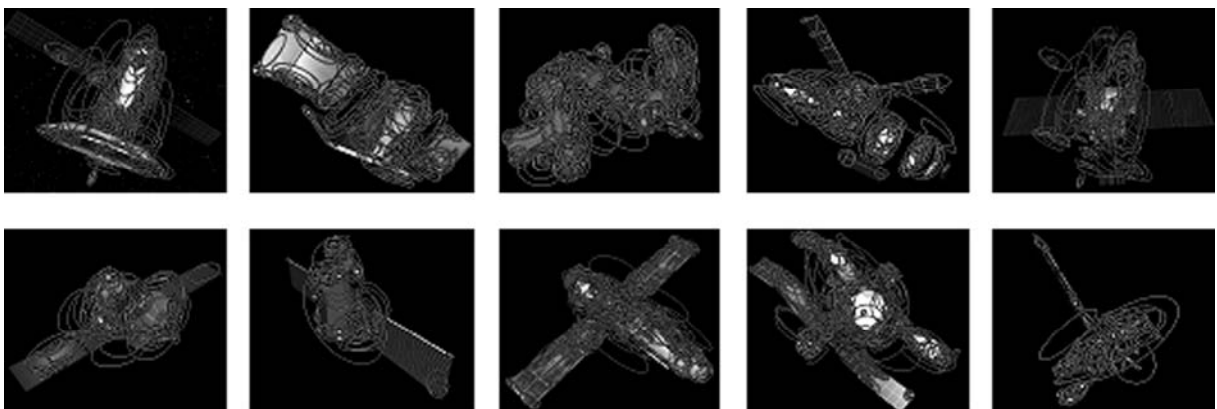


图7 10类空间目标图像的特征检测

$k=10$ ,但是考虑到实际会有偏差,因此通过做多组实验来确定最终的 $k$ 值。

我们将 $k$ 的取值设定在 $[8,16]$ 这个区间上,通过实验得出分类正确率与 $k$ 的曲线图。由图9可知,通过将 $k$ 设置成12时,分类的正确率会达到93.5%左右,因此本文选择 $k=12$ 。

### 3.2 比较分析

为了证明本文算法的有效性,我们将本文算法与其他的空间目标识别算法进行了比较,实验中比较采用的方法是文献[2],文献[3]的方法;为进一步验证本文所提出的二次表示方法有效性,实验中采用了与基于GMM建模图像表示的算法进行比较,该方法即本文中未经PLSA分解的方法,最后分类同样采用SVM方法。分类的评价标准采用所有类别的平均分类准确率CR,每类图像的分类准确率的计算公式为

$$CR = \frac{\text{正确图像分类数量}}{\text{测试图像数量}} \times 100\% \quad (6)$$

最终的分类效果比较如表1所示。

通过比较算法可以得出,本文的算法取得了令人满意的效果。文献[2,3]的方法需要从图像上分割出完整的目标,但是由于视角、模糊、遮挡等因素,

表1 分类算法的比较

分类方法	文献[2]方法	文献[3]方法	基GMM建模表示方法	本文方法
CR(%)	74.40	76.78	80.27	93.50

要从图像上完全分割目标几乎是不可能的;基于GMM的建模方法实际上是一种生成模式的表示方法,这种方法借助底层特征来表示,维数过高而且缺乏类别信息对图像存在的相似性无法处理,判别能力不够。相比于上述几种方法,本文方法将底层特征的表示提升到类别的二次表示,嵌入类别信息,使图像表示更具判别性,同时实现了降维,对于各类空间目标图像的相似性和同义性问题能够很好处理,因此一定程度上提高了分类效果。

为了验证本文局部特征的分类效率,在实验中与SIFT,MSER特征进行了分类比较,其中SIFT特征采用128维,MSER特征192维,算法的其余部分同本文方法一样。实验中得出了3类特征在不同 $k$ 值情况下的分类正确率(如图10所示),可以看出本文所用的特征的分类效果优于使用SIFT特征和MSER特征的方法。

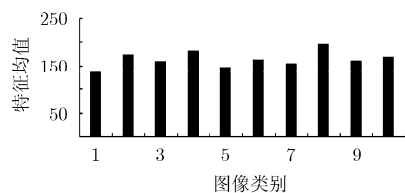


图8 各类图像所检测的特征数目均值

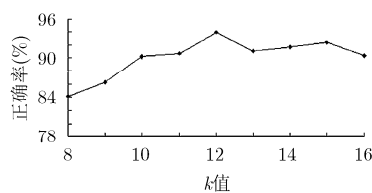


图9 k的取值与分类正确率的曲线图

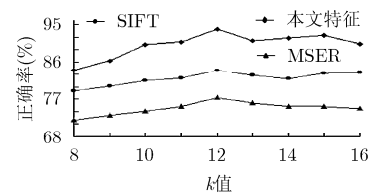


图10 3类特征的分类比较

## 4 结束语

本文针对空间目标图像提出了一种新的算法，该方法对局部特征进行全局建模，利用最大后验概率进行局部特征与视觉模式的匹配生成共现矩阵，然后通过PLSA分解得到基于类别的二次表示。该方法有效结合对象区域的底层特征，通过对图像的二次表示来提高最终的分类效率和计算效率，论文最后将本文的算法与已有的方法进行了比较。国内外对空间目标识别技术的研究较少，由于保密等因素关于空间目标真实图像的数据非常少，所进行的研究都采用仿真数据，因此本文下一步工作继续优化算法以提高适用性。

## 参考文献

- [1] 李元祥, 许鹏, 敬忠良, 等. 空间目标序列图像识别技术[J]. 哈尔滨工业大学学报, 2009, 41(11): 115-119.  
Li Yuan-xiang, Xu Peng, Jing Zhon-liang, *et al.*. Automatic space target recognition based on sequential images[J]. *Journal of Harbin Institute of Technology*, 2009, 41(11): 115-119.
- [2] 王晓雪, 杨永胜, 敬忠良. 基于改进核聚类算法的空间目标识别方法[J]. 中国空间科学技术, 2012, 32(2): 35-42.  
Wang Xiao-xue, Yang Yong-sheng, and Jing Zhong-liang. Space object identification method based on improved kernel clustering algorithm[J]. *China Space Science and Technology*, 2012, 32(2): 35-42.
- [3] Song W G, Zhou P H, Lu G L, *et al.*. Outer space target recognition based on image sequence[J]. *Key Engineering Materials*, 2010, 426(3): 161-165.
- [4] Ren Y, Zhang Y, Li Y, *et al.*. A space target recognition method based on compressive sensing[C]. 2011 Sixth International Conference on Image and Graphics, Xi'an, 2011: 582-586.
- [5] Cao W, Feng H, Hu L, *et al.*. Space target recognition based on biomimetic pattern recognition[C]. 2009 First International Workshop on Database Technology and Applications, Wuhan, 2009: 64-67.
- [6] Paalanen P, Kamarainen J K, Ilonen J, *et al.*. Feature representation and discrimination based on gaussian mixture model probability densities[J]. *Pattern Recognition*, 2006, 39(7): 1346-1358.
- [7] Thomas Hofmann. Unsupervised learning by probabilistic latent semantic analysis[J]. *Machine Learning*, 2001, 42(1): 177-196.
- [8] Matsukawa T and Kurita T. Image representation for generic object recognition using higher-order local autocorrelation features on posterior probability images[J]. *Pattern Recognition*, 2012, 45(2): 707-719.
- [9] Sivic J, Russell B C, Efros A A, *et al.*. Discovering objects and their location in images[C]. IEEE International Conference on Computer Vision, Beijing, 2005, 1: 370-377.
- [10] Quelhas P, Monay F, Odobez J M, *et al.*. A thousand words in a scene[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(9): 1575-1589.
- [11] Fan B, Wu F C, and Hu Z Y. Aggregating gradient distributions into intensity orders: a novel local image descriptor[C]. IEEE Conference on Computer Vision and Pattern Recognition, Colorado, 2011: 2377-2384.
- [12] Tang F, Lim S H, Chang N L, *et al.*. A novel feature descriptor invariant to complex brightness changes[C]. IEEE Conference on Computer Vision and Pattern Recognition, Miami, 2009: 2631-2638.
- [13] Wang Z H, Fan B, and Wu F C. Local intensity order pattern for feature description[C]. IEEE International Conference on Computer Vision, Barcelona, 2011: 603-610.
- [14] Felzenszwalb P, Girshick R, McAllester D, *et al.*. Object detection with discriminatively trained part based models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 32(9): 1627-1645.
- [15] Rissanen J. A universal prior for integers and estimation by minimum description length[J]. *Annals of Statistics*, 1983, 11(2): 417-431.

蒋飞龙：男，1987年生，硕士生，研究方向为模式识别。

孙锐：男，1976年生，副教授，硕士生导师，研究方向为计算机视觉、多媒体信息处理。

张旭东：男，1966年生，博士，副教授，硕士生导师，研究方向为机器视觉、传感器技术、智能信息处理。

李超：男，1987年生，硕士生，研究方向为视频指纹分析。