

# 基于仿射不变离散哈希的遥感图像多目标分类\*

孔 颖, 孙权森, 徐 晖, 刘亚洲, 纪则轩

(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

通讯作者: 孙权森, E-mail: sunquansen@njust.edu.cn



**摘 要:** 遥感图像的多目标分类是一个具有挑战性的课题。首先, 由于数据的复杂性以及算法对存储的高需求, 传统分类方法很难兼顾到分类的精度和速度; 其次, 遥感成像过程中产生的仿射变换, 使得目标的快速解译难以实现。为此, 提出一种基于仿射不变离散哈希(AIDH)的遥感图像多目标分类方法。该方法采用具有低存储、高效率优势的监督离散哈希框架, 结合仿射不变优化因子, 构造仿射不变离散哈希, 通过将具有相同语义信息的仿射变换样本约束到相似的二值码空间实现分类精度的提高。实验结果表明, 在NWPU VHR-10和RSDO-dataset两个数据集下, 相比于经典的哈希方法和分类方法, 所提方法在具备了高效性的同时, 其精度也得到了保证。

**关键词:** 遥感; 监督哈希; 仿射不变性; 多目标分类; 平均分类精度

中图法分类号: TP391

中文引用格式: 孔颖, 孙权森, 徐晖, 刘亚洲, 纪则轩. 基于仿射不变离散哈希的遥感图像多目标分类. 软件学报, 2019, 30(4): 914-926. <http://www.jos.org.cn/1000-9825/5661.htm>

英文引用格式: Kong J, Sun QS, Xu H, Liu YZ, Ji ZX. Multi-object classification of remote sensing image based on affine-invariant supervised discrete hashing. Ruan Jian Xue Bao/Journal of Software, 2019, 30(4): 914-926 (in Chinese). <http://www.jos.org.cn/1000-9825/5661.htm>

## Multi-object Classification of Remote Sensing Image Based on Affine-invariant Supervised Discrete Hashing

KONG Jie, SUN Quan-Sen, XU Hui, LIU Ya-Zhou, JI Ze-Xuan

(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

**Abstract:** The multi-object classification of remote sensing images has been a challenging task. Firstly, due to the complexity of the data and the high requirement of storage, the traditional classification methods are difficult to achieve both the accuracy and speed of the classification. Secondly, the affine transformation caused by the remote sensing imaging process, the real-time performance of the object interpretation is difficult to be realized. To solve the problem, a multi-object classification of remote sensing image is proposed based on affine-invariant discrete hashing (AIDH). This method uses supervised discrete hashing with the advantage of low storage and high efficiency, jointed with affine-invariant factor, to construct affine-invariant discrete hashing. By constraining the affine transform samples with the same semantic information to the similar binary code space, the method achieves the enhancement on classification precision. Experiments show that under the two datasets of NWPU VHR-10 and RSDO-dataset, the method presented in this paper is more efficient than classical hash method and classification method, and it is also guaranteed in accuracy.

**Key words:** remote sensing; supervised hashing; affine-invariant; multi-object classification; mean average precision (MAP)

## 1 引 言

随着不同用途卫星的发射以及遥感技术的飞速发展, 遥感图像的分辨率越来越高, 涵盖的光谱波段越来越

\* 基金项目: 国家自然科学基金(61673220)

Foundation item: National Natural Science Foundation of China (61673220)

本文由“多媒体数据的知识关联与理解专题”特约编辑蒋树强研究员、刘青山教授、孙立峰教授、李波教授推荐。

收稿时间: 2018-04-13; 修改时间: 2018-06-13; 采用时间: 2018-09-30

广,有效完成多目标分类工作从而正确解译卫星数据成为了一个具有挑战性的课题<sup>[1-6]</sup>。目前应用于自然图像分类的经典方法有  $K$  近邻分类法( $K$ -NN)、随机森林法(RF)、支持向量机(SVM)、稀疏表示分类器(SRC)<sup>[7-10]</sup>等。 $K$ -NN 法以测试样本邻域信息为参照确定其所属类别,能够较快地完成分类工作,然而分类精度受数据集的影响较大。RF 法通过训练集重构来放大分类模型间的差异,从而提高组合分类模型的外推预测能力,实现分类精度的提升,具有较强的分类鲁棒性,缺陷是算法训练时间过长。SVM 以其在小样本学习上的优势,降低了学习机对样本信息的依赖性,然而,当样本数增多以及为解决多目标分类而创建多个二类 SVM 分类器时,过高的机器内存要求和不断增加的运算时间,都影响了 SVM 方法的效率。SRC 方法以稀疏表示思想为基础,通过寻找测试样本在不同类别训练样本上的最优稀疏表示系数确定其类别号,具有很高的分类精度,然而生成过完备词典对电脑存储的高要求以及过长的运算时间是其主要的缺陷<sup>[11]</sup>。而应用于遥感图像的分类方法,除了以上几种分类方法外<sup>[12-15]</sup>,早期的工作是围绕目标物光谱分析为基础的,生成具有代表性的目标特征曲线或归一化指数进行分类判别。归一化指数的方法,是指特殊波段间运算所产生的能够代表目标明显特征的指数,如归一化植被指数(NDVI)、归一化水体指数(NDWI)、归一化雪盖指数(NDSI)分别用于提取植物、水体、雪覆盖信息<sup>[16]</sup>。特征曲线分析方法,是以同类物质样本在光谱空间的概率相似性为基础的,主要方法有最大似然法、最小距离法等<sup>[17]</sup>。这些早期遥感图像的分类方法具有很强的先验知识的支撑,具有很高的分类精度,然而适用范围却局限于多源数据、多光谱信息。近年来,基于人工神经网络的遥感图像分类方法通过模拟人类神经系统的结构与功能,将样本信息以神经元形式传递、加权分析和反馈的方式实现遥感图像的多目标分类。人工神经网络,主要包括前向网络、反馈网络和自组织网络,其结构设计很复杂,具有很强的特征捕捉能力,在多目标分类工作中取得了不错的效果,然而过长的学习时间以及网络结构难以收敛等因素都制约了其应用价值<sup>[18,19]</sup>。结合这些经典分类方法的利弊,我们希望找到一种新的方法,既不过分依赖数据集的构成,又能降低算法的存储需求,并能保证较高的精度和速度。

哈希方法以其低存储、高效率等优势在图像检索领域率先得到了广泛应用,近年来逐渐推广到图像分类、识别等问题中<sup>[20]</sup>。哈希方法的主要思想是,在不改变原数据语义相似性的基础上,通过将高维特征空间映射到低维的汉明空间,以低字节二值哈希码代替原有的高维实数特征以表征原有数据信息。早期哈希方法如局部敏感哈希(LSH)是数据无关的哈希技术,对投影向量及二值码初始化采取特定分布中随机采样的策略,虽然在模拟数据中有足够的理论支撑,但在实际场景下性能得不到保证<sup>[21]</sup>。近 10 年的哈希方法研究重心逐渐转化到利用数据自身以及一些监督信息设计高效的哈希函数上,其主要的分支划分为无监督哈希和监督哈希。无监督哈希方法主要有谱哈希(SH)<sup>[22]</sup>、可扩展图哈希(SGH)<sup>[23]</sup>、流形哈希(IMH)<sup>[24]</sup>等,通过挖掘数据集的分布信息,以谱聚类方法、流形结构等模型求解。而通过将语义信息等先验知识嵌入到哈希函数学习中,监督哈希方法近年来发展迅速,其主要代表方法有核监督哈希(KSH)<sup>[25]</sup>、最小损失哈希(MLH)<sup>[26]</sup>、监督离散哈希(SDH)<sup>[27]</sup>、列采样监督离散哈希(COSDISH)<sup>[28]</sup>、旋转不变离散哈希(RIDISH)<sup>[29]</sup>。以标签化类别信息加入到样本训练和测试中,通过引入核学习、度量学习等方法用于生成二值码,监督哈希方法在多目标分类应用中更为适用<sup>[30]</sup>。KSH<sup>[25]</sup>(CVPR 2012)通过将核映射思想引入到哈希技术中,以内积代替汉明距离运算,在保证样本相似性的基础上提高了分类精度。SDH<sup>[27]</sup>(CVPR 2015)采用适当的嵌入学习算法,将哈希码直接映射到标签信息上,并采用离散循环坐标梯度下降方法(DCC)在离散约束下按位求解,简化了运算过程,大大提高了算法效率。COSDISH<sup>[28]</sup>(AAAI 2016)以迭代采样语义相似矩阵的多列完成所有数据的训练学习,代替了传统的松弛化约束求解的方式,提高了分类方法的速度和稳定性。

目前为止,哈希方法在自然图像分类问题上已取得很大的成功,然而要引入到遥感领域还需要结合遥感卫星成像前后的具体星情。遥感图像的目标识别,除了受样本有限、纹理丰富等遥感图像数据因素的影响外,还会因在轨相机的姿态、传感器拍摄时不同方位角和俯仰角而发生变化。随着遥感目标数量的不断增大,不断增加的几何校正预处理过程已经不能满足在轨实时目标检索和匹配工作的需求。对常见卫星遥感图像,其成像模型可用中心投影成像来近似,可以认为观测数据和实际数据间近似满足仿射变换关系<sup>[31]</sup>。为此,在将哈希引入遥感图像多目标分类时,需要考察样本的仿射变换信息,使这些信息的哈希特征表达具有相似性并能更好地表达

原样本的类别信息.针对上述问题,本文提出一种仿射不变离散哈希的框架,在 SDH 目标函数中引入仿射不变因子约束,通过求解汉明空间的离散优化问题实现优质二值码的生成.而仿射不变因子约束项迫使仿射变化扩大的同型号样本集具有相同的二值码,从而保证了哈希算法的仿射不变性.本文的框架图如图 1 所示.

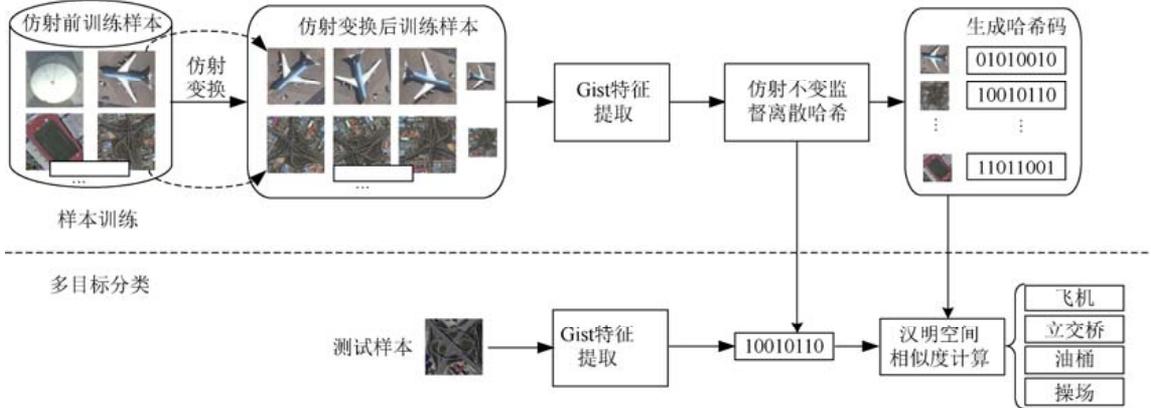


Fig.1 The frame of the proposed method  
图 1 本文算法框架图

本文的主要贡献有以下两个方面.

- (1) 提出一种仿射不变离散哈希方法,通过在监督离散哈希二值码的生成过程中加入仿射不变项模型优化,有效地解决了因卫星成像形成的仿射变换畸变对分类精度的影响.
- (2) 通过在不同数据集下的实验对比,其结果表明,不论是横向对比于经典的哈希方法,还是纵向对比于经典的分类方法,本文方法均兼顾了多目标分类的精度和速度.

## 2 仿射不变离散哈希(AIDH)

### 2.1 监督离散哈希(SDH)

假设有  $n$  个样本,  $X = \{x_i\}_{i=1}^n \in R^{d \times n}$ , 每一列代表一个  $d$  维的样本.为了正确表达样本语义信息,哈希学习需要生成一组二值码  $B = \{b_i\}_{i=1}^n \in \{-1, 1\}^{L \times n}$ , 其中,  $b_i$  是对应于样本  $x_i$  的字节长度为  $L$  的二值化编码.

不同于无监督哈希最小化样本间距离相似性的求解方式,监督哈希利用了训练样本的标签信息,使得真实类别(标签)与预测类别间的差异尽可能地小.监督哈希的正则化目标函数可表示为

$$\left. \begin{aligned} \min_{B, W, H} \sum_{i=1}^n \|y_i - W^T b_i\|^2 + \lambda_1 \|W\|^2 \\ \text{s.t. } b_i = \text{sgn}(H(x_i)), i = 1, \dots, n \end{aligned} \right\} \quad (1)$$

其中,  $W = \{w_k\}_{k=1}^C \in R^{L \times C}$  是哈希码在  $C$  个分类上的投影矩阵;  $Y = \{y_i\}_{i=1}^n \in R^{C \times n}$  是样本的标签矩阵,若  $y_{ki}=1$ ,表示样本  $x_i$  属于类别  $k$ ,反之,若  $y_{ki}=0$ ,则表示样本  $x_i$  不属于类别  $k$ ;  $H(x_i)$  为哈希函数;  $\lambda_1$  是正则化参数.

为了获得高质量的哈希码,使得学习到的二值编码能够保持原有样本空间的相似性,并减少因随机映射生成哈希函数的误差,考虑构建哈希函数  $H(X)$  将样本的连续特征映射到二值化空间上,对二值编码  $B$  进行离散约束.于是,式(1)改写为

$$\left. \begin{aligned} \min_{B, W, H} \|Y - W^T B\|^2 + \lambda_1 \|W\|^2 + \lambda_2 \|B - H(X)\|^2 \\ \text{s.t. } B \in \{-1, 1\}^{L \times n}, i = 1, \dots, n \end{aligned} \right\} \quad (2)$$

式(2)的第 3 项代表哈希函数  $H(X)$  与二值码  $B$  之间的拟合误差,  $\lambda_1$  为惩罚因子.哈希函数通常是采用映射学

习算法获取,为保持样本的非线性结构,采用如下形式:

$$H(X) = P^T \phi(X) \quad (3)$$

其中,  $\phi(X) \in R^{m \times n}$  为径向基核映射产生的矩阵,每一列  $\phi(x) = [\exp(-\|x - a_1\|^2/\sigma), \dots, \exp(-\|x - a_m\|^2/\sigma)]^T$ , 而  $\{a_j\}_{j=1}^m$  是训练样本  $X$  中随机抽取的  $m$  个样本点,  $\sigma$  则是核函数的步长. 矩阵  $P \in R^{m \times L}$  是降维矩阵,用于将核函数空间投影到哈希码空间. 通过式(3)的非线性哈希函数构造过程,可以明显减少式(2)中原始距离和汉明距离的重构误差<sup>[27]</sup>.

## 2.2 仿射不变离散哈希

若直接将哈希用于遥感图像的目标分类中,会受限于拍摄时的高度和角度等因素,形成不同程度的仿射变换. 为此,我们希望在 SDH 的基础上做出改进,生成一种具有仿射不变性的哈希编码.

设图像仿射变换前后坐标分别为  $G(x, y)$  和  $G'(x', y') \in R^2$ , 则有:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

其中,非奇异阵  $A = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} = \begin{bmatrix} s \cdot \cos \alpha & -s \cdot \sin \alpha \\ s \cdot \sin \alpha & s \cdot \cos \alpha \end{bmatrix} \neq 0$  由物体的旋转变换  $T_\alpha$  和缩放变换  $T_\beta$  组成,而  $t = [t_1, t_2]^T$  代表物体的平移变换. 考虑到样本图像的稠密性,目标平移产生的影响几乎没有,  $t$  项不作考虑<sup>[32]</sup>.

依据 MSA 特征提取的方法可知,图像在仿射变换前后的对应像素点亮度函数具有相同分布,这种同分布的关系是独立于仿射变换的;那么图像的仿射变换不会改变图像所代表物体的语义信息<sup>[33]</sup>. 由于仿射变换前后的图像对应着相同标签模式类的信息,那么可以将仿射后扩大的样本加入到训练样本集,并加以仿射约束信息优化原有目标函数,生成更优质的哈希码<sup>[2]</sup>.

定义  $R$  个旋转角度  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_R\}$  及其旋转变换  $T_\alpha = \{T_{\alpha_1}, T_{\alpha_2}, \dots, T_{\alpha_R}\}$ , 其中,  $T_{\alpha_i}$  表示经  $\alpha_i$  角度旋转变换;定义  $S$  个缩放因子  $\beta = \{\beta_1, \beta_2, \dots, \beta_S\}$  及其缩放变换  $T_\beta = \{T_{\beta_1}, T_{\beta_2}, \dots, T_{\beta_S}\}$ , 其中,  $T_{\beta_i}$  表示  $\beta_i$  倍缩放变换. 那么仿射变换可表达为旋转变换和缩放变换的笛卡尔积:  $T_{\alpha, \beta} = T_\alpha \times T_\beta$ . 将  $T_{\alpha, \beta}$  应用于所有训练样本  $X$ , 结合原始训练集,即构成了新的用于仿射不变离散哈希模型的样本集  $T_{\alpha, \beta} = \{X, T_{\alpha, \beta} X\}$ .

由于仿射变换前后不改变物体本身的语义信息,那么所得到的哈希码应该是相似的,生成的哈希码对数据集具有仿射不变性<sup>[5, 34]</sup>. 类似于旋转不变 HOG 特征的处理方案,添加仿射不变约束项如下:

$$Q(B, T_{\alpha, \beta} B) = \lambda_3 \sum_{i=1}^n \|b_i - \overline{T_{\alpha, \beta} b_i}\|^2 = \lambda_3 \|B - \overline{T_{\alpha, \beta} B}\|^2 \quad (5)$$

其中,  $\overline{T_{\alpha, \beta} b_i} = \frac{1}{R \times S} \sum_{j=1}^R \sum_{k=1}^S T_{\alpha_j} T_{\beta_k} b_i$  表示样本  $x_i$  所有仿射变换张成的样本子空间所生成二值码的均值.

这样,式(2)的原目标函数就可以改写为

$$\left. \begin{aligned} & \min_{B, W, H} \|Y - W^T B\|^2 + \lambda_1 \|W\|^2 + \lambda_2 \|B - H(X)\|^2 + \lambda_3 \|B - \overline{T_{\alpha, \beta} B}\|^2 \\ & \text{s.t.: } B \in \{-1, 1\}^{L \times n}, i = 1, \dots, n \end{aligned} \right\} \quad (6)$$

可以看到,通过添加此约束项可以迫使每个样本生成哈希码更接近于其张成的仿射样本子空间的均值,从而实现整体样本集  $X_{\alpha, \beta}$  与其二值哈希编码  $B$  尽可能地相似. 式(6)即为仿射不变离散哈希所需优化的目标方程<sup>[35]</sup>.

## 2.3 优化求解方案

仿射不变离散哈希的未知变量分别为  $B$ 、 $W$ 、 $H$ , 可以通过交替优化的方式把式(6)的问题分解为 3 个子问题,分别进行求解.

(1) *P-step*: 固定  $B$ 、 $W$ , 优化  $P$ .

由于  $P$  优化的过程只与  $H(X)$  有关,因而,只需优化式(6)中的第 3 项. 求解结果如下:

$$P = (\phi(X)\phi(X)^T)^{-1}\phi(X)B^T \quad (7)$$

(2) *W-step*: 固定  $P$ 、 $B$ , 优化  $W$ .

原方程简化为求解只有前两项的监督哈希问题, 通过最小二乘法可求解:

$$W = (BB^T + \lambda_1 I)^{-1} B Y^T \quad (8)$$

(3) *B-step*: 固定  $W$ 、 $P$ , 优化  $B$ . 原问题转化为求解式(9)所示函数:

$$\left. \begin{aligned} \min_B & \left\{ \|Y - W^T B\|^2 + \lambda_2 \|B - H(X)\|^2 + \lambda_3 \|B - \overline{T_{\alpha, \beta} B}\|^2 \right\} \\ \text{s.t. } & B \in \{-1, 1\}^{L \times n} \end{aligned} \right\} \quad (9)$$

将式(9)展开, 再化简, 等价于求解:

$$\left. \begin{aligned} \min_B & \left\{ \|W^T B\|^2 - 2Tr(B^T A) \right\} \\ \text{s.t. } & B \in \{-1, 1\}^{L \times n} \end{aligned} \right\} \quad (10)$$

其中,  $A = WY + \lambda_2 H(X) + \lambda_3 \overline{T_{\alpha, \beta} B}$ , 而  $Tr(\cdot)$  表示矩阵的迹.

受 SDH 优化方式的启发, 采用 DCC 来按位求解哈希码, 每一位对应  $B$  的一行数据.

令  $b_i$  为矩阵  $B$  的第  $i$  行,  $q_i$  为矩阵  $Q$  的第  $i$  行,  $w_i$  为矩阵  $W$  的第  $i$  行;  $\tilde{B}$  为矩阵  $B$  去除  $b_i$  后的矩阵,  $\tilde{A}$  为矩阵  $A$  去除  $a_i$  后的矩阵,  $\tilde{W}$  为矩阵  $W$  去除  $w_i$  后的矩阵, 那么:

$$\left. \begin{aligned} \|W^T B\|^2 &= const + 2w_i^T \tilde{W}^T \tilde{B} b_i \\ Tr(B^T A) &= const + a_i^T b_i \end{aligned} \right\} \quad (11)$$

式(11)中  $const$  为常数项, 不影响优化过程. 将式(11)代入式(10)中约掉常数项和倍数项, 可得:

$$\left. \begin{aligned} \min_b & (w_i^T \tilde{W}^T \tilde{B} - a_i^T) b_i \\ \text{s.t. } & b_i \in \{-1, 1\}^n \end{aligned} \right\} \quad (12)$$

这个问题可以获得如下的闭解:

$$b_i = \text{sgn}(a_i - \tilde{B}^T \tilde{W} w_i) \quad (13)$$

因而, 本文方法的基本步骤可见下面的算法<sup>[35]</sup>.

**算法.** 仿射不变离散哈希(AIDH).

输入: 训练样本  $\{x_i, y_i\}_{i=1}^n$ , 字节数  $L$ , 锚点数  $m$ , 最大迭代次数  $it$ , 参数  $\lambda_1, \lambda_2, \lambda_3$ ;

输出: 二值码  $B = \{b_i\}_{i=1}^n \in \{-1, 1\}^{L \times n}$ .

(1) 在训练样本集中随机选取  $m$  个样本, 并利用径向基函数求得式(3)中的  $\phi(X)$ .

(2) 随机方式初始化所有样本的二值码  $b_i \in \{-1, 1\}^L$ .

(3) 循环下述 3 个步骤直至收敛.

*W-step*: 利用式(8)计算  $W$ .

*P-step*: 利用式(7)计算  $P$ , 代入式(3)中, 求解  $H(X)$ .

*B-step*: 利用式(13)按位迭代学习  $b_i$ , 得到哈希码  $B = \{b_i\}_{i=1}^n \in \{-1, 1\}^{L \times n}$ .

### 3 实验部分

#### 3.1 数据集介绍及算法运行环境

为了验证 AIDH 算法的有效性, 本文采用两个数据集 NWPU VHR-10 和 RSOD-Dataset. NWPU VHR-10 是西北工业大学团队标注的航天遥感目标检测数据集, 而 RSOD-Dataset 是武汉大学团队标注的一个遥感图像公共数据集<sup>[36]</sup>. 两个数据集通过提供了源图像以及各自图像中所含带目标坐标信息的标签文件(ground truth), 给出了可提取的多类目标样本. 这两个数据集的优势就在于源图像的存在给出了旋转样本边缘的信息, 避免了样

本自旋转产生的边缘黑带。

NWPU VHR-10 提供了 650 幅图像,生成了 3 896 个目标,分为 10 个数据类,分别为 757 架飞机、302 艘舰船、655 个油桶、390 座棒球场、524 座网球场、159 座篮球场、163 座足球场、224 个港口、124 座桥梁、598 个车辆。这些数据来自谷歌地球(分辨率 0.5m~2m)和法伊英根数据(分辨率 0.08m)<sup>[4]</sup>。RSOD-Dataset 提供了 936 幅图像,生成了 7 400 个目标,分为 4 个数据类,分别为 5 374 架飞机、1 698 个油桶、176 架立交桥、149 个操场。这些数据取自谷歌地图和天地图等,图 2 给出了该数据集的各类目标信息,而各类目标的分辨率见表 1。

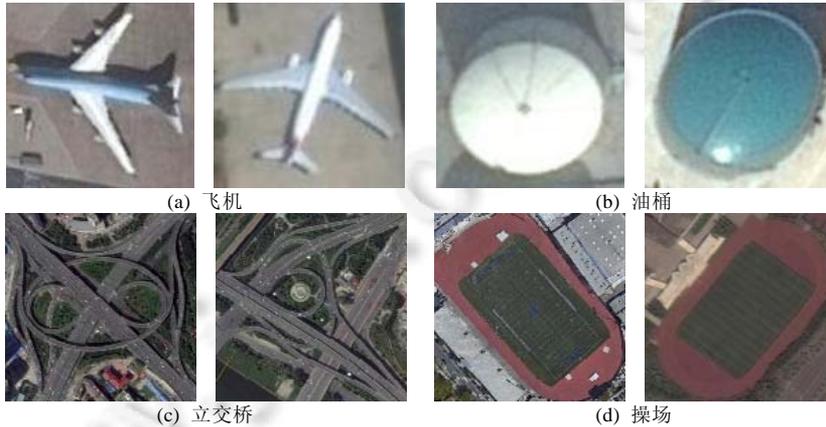


Fig.2 Four-class object samples of RSOD-Dataset

图 2 RSOD-Dataset 的 4 类目标示例

Table 1 The resolution of different objects of RSOD-Dataset

表 1 RSOD-Dataset 各类目标分辨率

| 目标类<br>分辨率(m) | 飞机    | 油桶    | 立交桥    | 操场    |
|---------------|-------|-------|--------|-------|
|               | 0.5~2 | 0.3~1 | 1.25~3 | 0.4~1 |

在数据的特征选择上,使用的是国际上公认的 Gist 特征<sup>[37]</sup>。该特征是 512 维的,由基于视觉词袋模型的网格化 Gabor 小波系数组成。在样本集设置上(以 RSOD-Dataset 为例),按比例随机抽取 2 000 个来自 4 类的目标作为测试样本集;然后,对剩余的 5 400 个样本集作仿射变换,并对扩大后的样本集打乱次序后作为训练样本集,将训练和测试样本集的特征及其类别标签合成到一个 .mat 文件中用于算法验证。本文的实验中,旋转系数为  $R=11$ ,  $\alpha=\{30^\circ, 60^\circ, \dots, 330^\circ\}$ , 缩放系数都是  $S=2, \beta=\{0.5, 0.75\}$ 。此外,为了验证算法的有效性,我们生成了 10 组 .mat 文件,取平均值作为算法的精度结果。本文的所有实验是在软件 MATLAB 2014a 上仿真实现的,电脑的处理器的处理器是英特尔酷睿 i7-8550U, CPU 主频 1.99GHz, 内存为 8GB RAM。

### 3.2 实验对比及定量分析

本文是以哈希方法为基础的遥感多目标分类,实验分为两部分,一是比较本文方法和经典的哈希方法之间的精度,二是比较本文方法与传统的多目标分类方法之间的优越性。

#### 3.2.1 哈希方法之间的性能对比

为了验证仿射不变离散哈希方法在哈希函数改进上的优越性,本文选取了谱哈希(SH)<sup>[22]</sup>、可扩展图哈希(SGH)<sup>[23]</sup>、核监督哈希(KSH)<sup>[25]</sup>、列采样监督离散哈希(COSDISH)<sup>[28]</sup>、监督离散哈希(SDH)<sup>[27]</sup>、旋转不变离散哈希(RIDISH)<sup>[29]</sup>作为对比方法以验证算法的精度。此外,为了验证本文方法相对于旋转不变离散哈希的优势,特增加一步实验,旋转&尺度不变离散哈希(RI&SIDISH)代表的是旋转不变离散哈希方法在旋转不变约束项中加入缩放变换的结果。

定量指标上,选取 4 个指标来评价算法的精度。表 2~表 5 显示出这些指标在各种哈希算法下的评价结果。

(1) 平均分类精度(MAP):准确分类的目标占总数的比例.

(2) 前 1 000 个样本的分类精度(precision of top 1000 returned samples):对所有测试样本寻找其匹配度较高的前 1 000 个训练样本计算其分类精度,再对这些测试样本分类精度取平均值.

(3) 汉明距离在半径为 2 时的分类精度(precision of Hamming distance with radius 2):以汉明距离小于等于 2 为测度,对所有测试样本寻找与其接近的训练样本并计算分类精度,再对这些测试样本的分类精度取平均值.

(4) 各分类精度(AP):各类目标的分类精度.

表 2~表 4 从宏观的角度给出了 7 种哈希方法在多目标分类中的精度对比情况.在哈希字节的选取上,给出了 8 字节、16 字节、32 字节和 64 字节的数据.可以看到,随着哈希字节数的增长,各种哈希方法的增长趋势是一致的,直至 32 字节,而到了 64 字节时,很多哈希方法的精度反而会有所降低.这种情况的发生往往取决于数据集的分布以及过拟合的发生.鉴于此,在表 5 中给出了 32 字节时各目标分类的结果.

**Table 2** MAP of different hashing method in 8,16,32,64 bits

**表 2** 不同哈希方法在 8、16、32、64 字节下的平均分类精度

| 方法 \ 数据集    | NWPU VHR-10 |         |         |         | RSOD-Dataset |         |         |         |
|-------------|-------------|---------|---------|---------|--------------|---------|---------|---------|
|             | 8-bits      | 16-bits | 32-bits | 64-bits | 8-bits       | 16-bits | 32-bits | 64-bits |
| 谱哈希         | 0.197 2     | 0.198 5 | 0.226 8 | 0.214 1 | 0.498 6      | 0.511 3 | 0.538 6 | 0.496 2 |
| 可扩展图哈希      | 0.266 8     | 0.279 7 | 0.293 6 | 0.298 2 | 0.559 8      | 0.582 0 | 0.600 0 | 0.626 2 |
| 核监督哈希       | 0.677 3     | 0.709 9 | 0.735 7 | 0.759 5 | 0.727 4      | 0.774 5 | 0.800 1 | 0.816 8 |
| 列采样监督离散哈希   | 0.729 8     | 0.784 5 | 0.823 7 | 0.832 0 | 0.826 0      | 0.888 4 | 0.915 0 | 0.927 8 |
| 监督离散哈希      | 0.708 9     | 0.847 5 | 0.910 1 | 0.914 1 | 0.760 1      | 0.897 6 | 0.938 9 | 0.958 1 |
| 旋转不变离散哈希    | 0.835 3     | 0.886 3 | 0.920 5 | 0.928 4 | 0.892 0      | 0.927 4 | 0.952 9 | 0.964 2 |
| 旋转&尺度不变离散哈希 | 0.789 5     | 0.843 4 | 0.870 4 | 0.883 2 | 0.871 3      | 0.899 4 | 0.935 4 | 0.949 5 |
| 仿射不变离散哈希    | 0.850 5     | 0.902 0 | 0.926 3 | 0.933 8 | 0.911 1      | 0.936 4 | 0.959 0 | 0.972 9 |

**Table 3** The precision of top 1000 returned samples of different hashing method in 8,16,32,64 bits

**表 3** 不同哈希方法在 8、16、32、64 字节下的前 1 000 个样本的分类精度

| 方法 \ 数据集    | NWPU VHR-10 |         |         |         | RSOD-Dataset |         |         |         |
|-------------|-------------|---------|---------|---------|--------------|---------|---------|---------|
|             | 8-bits      | 16-bits | 32-bits | 64-bits | 8-bits       | 16-bits | 32-bits | 64-bits |
| 谱哈希         | 0.259 3     | 0.284 2 | 0.341 6 | 0.328 4 | 0.514 3      | 0.516 3 | 0.536 2 | 0.502 6 |
| 可扩展图哈希      | 0.392 0     | 0.400 6 | 0.410 4 | 0.423 9 | 0.659 9      | 0.681 3 | 0.721 0 | 0.755 7 |
| 核监督哈希       | 0.696 2     | 0.748 7 | 0.789 0 | 0.792 0 | 0.727 8      | 0.780 5 | 0.805 0 | 0.819 7 |
| 列采样监督离散哈希   | 0.683 0     | 0.730 0 | 0.764 1 | 0.774 1 | 0.766 9      | 0.819 7 | 0.869 9 | 0.899 4 |
| 监督离散哈希      | 0.744 3     | 0.868 6 | 0.915 3 | 0.918 5 | 0.763 0      | 0.900 6 | 0.944 2 | 0.949 4 |
| 旋转不变离散哈希    | 0.852 5     | 0.901 0 | 0.925 3 | 0.930 6 | 0.894 4      | 0.926 9 | 0.950 1 | 0.959 5 |
| 旋转&尺度不变离散哈希 | 0.832 9     | 0.871 1 | 0.891 8 | 0.899 1 | 0.868 7      | 0.918 0 | 0.931 8 | 0.941 9 |
| 仿射不变离散哈希    | 0.875 9     | 0.913 8 | 0.930 0 | 0.937 4 | 0.917 9      | 0.934 5 | 0.952 7 | 0.987 8 |

**Table 4** The precision of Hamming distance with radius 2 of different hashing method in 8,16,32,64 bits

**表 4** 不同哈希方法在 8、16、32、64 字节下的汉明距离在半径为 2 时的分类精度

| 方法 \ 数据集    | NWPU VHR-10 |         |         |         | RSOD-Dataset |         |         |         |
|-------------|-------------|---------|---------|---------|--------------|---------|---------|---------|
|             | 8-bits      | 16-bits | 32-bits | 64-bits | 8-bits       | 16-bits | 32-bits | 64-bits |
| 谱哈希         | 0.214 4     | 0.381 8 | 0.407 3 | 0.005 5 | 0.514 9      | 0.740 2 | 0.788 3 | 0.019 7 |
| 可扩展图哈希      | 0.287 1     | 0.435 9 | 0.420 0 | 0.032 0 | 0.572 1      | 0.822 3 | 0.812 9 | 0.039 3 |
| 核监督哈希       | 0.588 6     | 0.728 8 | 0.783 3 | 0.680 7 | 0.722 9      | 0.782 0 | 0.784 4 | 0.819 8 |
| 列采样监督离散哈希   | 0.617 8     | 0.599 0 | 0.414 0 | 0.367 0 | 0.751 0      | 0.665 1 | 0.462 7 | 0.359 0 |
| 监督离散哈希      | 0.431 5     | 0.837 4 | 0.881 7 | 0.821 7 | 0.661 9      | 0.904 6 | 0.933 8 | 0.902 0 |
| 旋转不变离散哈希    | 0.513 4     | 0.890 8 | 0.918 7 | 0.871 7 | 0.759 0      | 0.921 5 | 0.960 6 | 0.970 4 |
| 旋转&尺度不变离散哈希 | 0.524 9     | 0.848 7 | 0.882 7 | 0.854 2 | 0.780 7      | 0.877 3 | 0.941 0 | 0.962 5 |
| 仿射不变离散哈希    | 0.670 1     | 0.901 3 | 0.920 7 | 0.925 0 | 0.748 7      | 0.929 4 | 0.967 3 | 0.968 1 |

两个数据集均能有效验证 7 种哈希方法的分类精度.从宏观对比上看,武汉大学 RSOD-Dataset 数据集的分类精度高于西北工业大学 NWPU VHR-10 数据集的分类精度,究其原因,是武汉大学的数据集虽然样本数很

多,但是数据类型较少(4 个),样本的差异性较大,数据集较为简单,使得分类的精度偏高.而西北工业大学的数据集,包含的数据类型较多(10 个),样本集的多样性增加了准确分类的难度,更能层次化地反映出不同算法在多目标的分类效果.

从监督和无监督哈希方法之间对比上来看,两种无监督哈希方法 SH、SGH 的精度较低,甚至连 30% 都不到,这符合了机器学习中无监督方法在没有先验信息时,即使挖掘更多的数据自身的特征也难以有效提升分类精度的结论.而 6 种监督哈希方法的平均分类精度都达到了至少 65% 以上.而且, KSH 方法在 8 字节处的分类精度就已达到了 67.73% 的精度,说明了哈希方法以较少的字节依然能够很好地区分出不同目标类的语义信息. COSDISH 方法在 8 字节处就取得了高于 KSH、SDH 的分类精度,然而,随着哈希码字节数的增多,分类精度的增长却不如其他方法明显,该方法的优点在于,对样本数目不敏感,两个数据集下分类精度随字节增长的幅度相似,代表着这种算法的稳定性很高.

观察表 2 后 4 行的数据可以发现, RIDISH 相对于 SDH 有更大的提升在于 8 字节处就能获得 8 成以上的分类精度,这主要得益于该方法考虑了样本的旋转多样性.而观察 RIDISH 和 RI&SIDISH 的精度可以发现,在加入原数据的缩放样本后,数据的精度反而有所下降(约为 4%).这个现象表明,数据的尺度性会影响到 RIDISH 的分类精度.而本文方法在考察了样本的仿射变换信息后,精度达到了 93.74%,高于 RI&SIDISH 方法的 89.91%,即说明本文方法在不同尺度目标的语义表达的有效性,达到了改进哈希方法对仿射不变性的要求.同时,本文方法在 MAP 等精度上也略高于 RIDISH 方法,表明本文方法继承并发扬了旋转不变哈希在哈希码约束项上改造的优势.

图 3 和图 4 直观地给出了两个数据集下 7 种哈希方法在不同字节数时的 3 种分类精度指标.

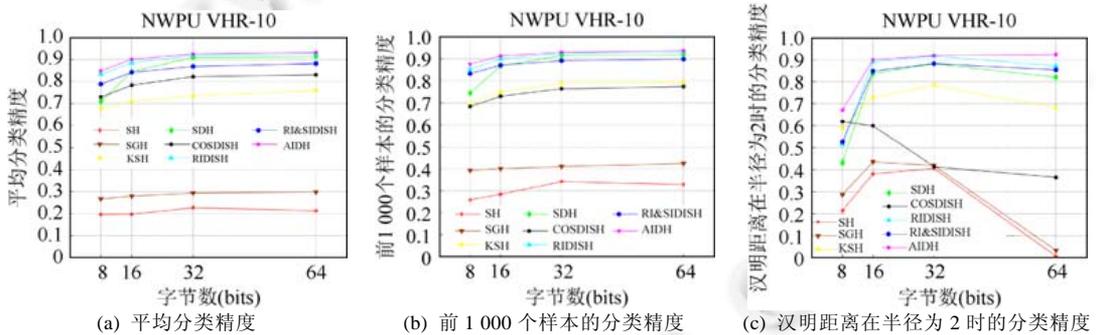


Fig.3 The diagram of three precision indexes varied with bits in NWPU VHR-10 dataset

图 3 NWPU VHR-10 数据集的 3 个精度指标随字节变化示意图

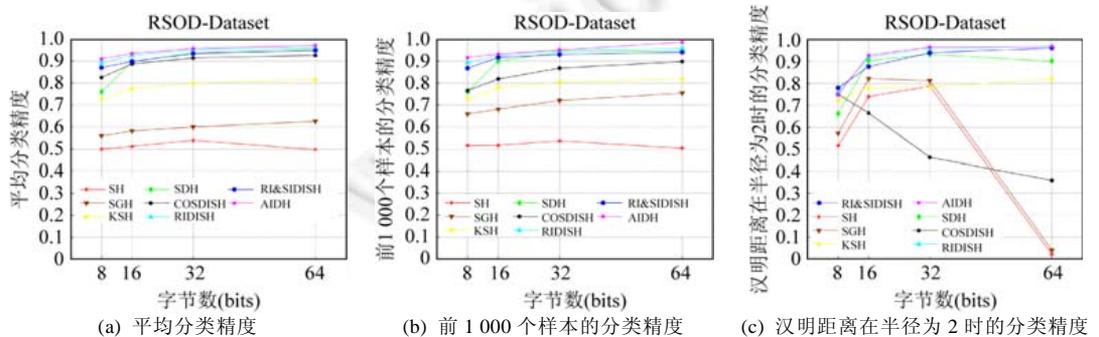


Fig.4 The diagram of three precision indexes varied with bits in RSOD-Dataset

图 4 RSOD-Dataset 数据集的 3 个精度指标随字节变化示意图

不难看出,SH、SGH 方法的分类精度偏低,32 字节之后精度不再提升甚至会有所下降,尤其是汉明半径在 2

以内的分类精度在 64 字节处极低,表明汉明半径以 2 为测度时加入进了过多的误判样本.而 KSH、COSDISH、SDH 则高于两种无监督哈希方法,这一方面得益于类标签信息的引入,另一方面则是核学习方法的优势.然而,这两种方法在 8 字节处的精度还是偏低于后 3 种方法,毕竟字节数过少难以准确地表达出样本的特征.而 RISDH、RI&SIDISH 和本文方法在各字节下均保持着较高的精度,而且分类效果相对稳定,这主要得益于样本集扩充和算法的有效性,以样本点的成倍增加弥补了特征维数低的不足.

表 5 给出了两个数据集中不同哈希方法在 32 字节下,各目标的分类精度以及总体分类精度(MAP).从纵向对比可以看到,NWPU VHR-10 数据集在飞机、油桶、车辆等小目标的分类精度上均较高,而篮球场、桥梁、足球场等大目标的精度则偏低.这种现象的出现原因有两个:(1) 小目标的分辨率都较高,辨识度高,降低了分类的难度;反之,大目标覆盖范围广,分辨率较低,又包含了很多有用信息,无形中提高了分类的难度.(2) 从数据集的构成上看,小目标的样本数较多,而大目标的样本数较少.这一点无疑更加重要,从“网球场”这一类来看,同样是大目标数据,在样本数足够多的情况下依然取得了不错的效果.同样地,对于 RSOD-Dataset 而言,立交桥和操场的分类精度要低于飞机和油桶.而从横向对比不同方法在各目标分类精度的结果来看,本文方法虽然在个别目标类的分类精度与 RIDISH 不相上下(表格中加粗显示),但是,整体分类精度的提高,尤其是篮球场和桥梁两个“短板”精度上的提高,反映出仿射不变离散哈希在多目标分类上的优势.

Table 5 AP and MAP of different hashing methods in 32 bits

表 5 不同哈希方法在 32 字节下各目标的分类精度和总体分类精度

| 数据集          | 方法  | 谱哈希     | 可扩展图哈希  | 核监督哈希   | 采样监督离散哈希 | 监督离散哈希  | 旋转不变离散哈希       | 旋转&尺度不变离散哈希 | 仿射不变离散哈希       |
|--------------|-----|---------|---------|---------|----------|---------|----------------|-------------|----------------|
|              |     |         |         |         |          |         |                |             |                |
| NWPU VHR-10  | 飞机  | 0.267 9 | 0.337 3 | 0.931 1 | 0.896 5  | 0.976 0 | <b>0.980 4</b> | 0.969 9     | 0.975 3        |
|              | 轮船  | 0.149 2 | 0.194 7 | 0.493 3 | 0.712 9  | 0.841 0 | 0.866 1        | 0.826 0     | <b>0.869 4</b> |
|              | 油桶  | 0.382 9 | 0.503 2 | 0.975 5 | 0.970 2  | 0.991 6 | <b>0.993 8</b> | 0.985 3     | 0.990 4        |
|              | 棒球场 | 0.204 7 | 0.242 6 | 0.702 4 | 0.833 7  | 0.918 9 | 0.910 2        | 0.911 4     | <b>0.947 6</b> |
|              | 网球场 | 0.226 4 | 0.309 1 | 0.845 2 | 0.905 4  | 0.958 5 | <b>0.976 7</b> | 0.950 4     | 0.971 9        |
|              | 篮球场 | 0.064 1 | 0.097 5 | 0.185 3 | 0.435 1  | 0.677 6 | 0.592 7        | 0.527 0     | <b>0.606 9</b> |
|              | 足球场 | 0.078 5 | 0.184 3 | 0.472 0 | 0.677 8  | 0.777 5 | <b>0.874 7</b> | 0.774 0     | 0.863 5        |
|              | 港口  | 0.102 1 | 0.140 6 | 0.214 0 | 0.672 8  | 0.834 4 | <b>0.872 4</b> | 0.799 5     | 0.860 8        |
|              | 桥梁  | 0.053 3 | 0.087 4 | 0.172 4 | 0.364 9  | 0.463 1 | 0.656 0        | 0.603 2     | <b>0.716 7</b> |
|              | 车辆  | 0.222 6 | 0.251 2 | 0.819 4 | 0.848 1  | 0.950 7 | 0.921 6        | 0.916 6     | <b>0.941 6</b> |
|              | MAP | 0.226 8 | 0.293 6 | 0.735 7 | 0.823 7  | 0.910 1 | 0.920 5        | 0.896 4     | 0.926 3        |
| RSOD-Dataset | 飞机  | 0.565 4 | 0.714 8 | 0.898 7 | 0.884 2  | 0.998 3 | <b>0.998 7</b> | 0.997 2     | 0.997 5        |
|              | 油桶  | 0.593 2 | 0.630 0 | 0.870 5 | 0.971 9  | 0.979 6 | 0.986 5        | 0.982 6     | <b>0.987 4</b> |
|              | 立交桥 | 0.381 9 | 0.329 5 | 0.589 2 | 0.838 3  | 0.887 5 | <b>0.860 5</b> | 0.846 7     | 0.844 1        |
|              | 操场  | 0.444 5 | 0.419 8 | 0.468 1 | 0.918 8  | 0.642 3 | 0.741 3        | 0.656 3     | <b>0.823 7</b> |
|              |     | MAP     | 0.538 6 | 0.600 0 | 0.800 1  | 0.915 0 | 0.938 9        | 0.952 9     | 0.935 4        |

最后,我们给出各参数对本文方法的影响.AIDH 主要有 3 个参数( $R$ 、 $S$ 、 $m$ ),分别是旋转系数、缩放系数以及锚点数.下面通过变化其中一个参数、固定另外两个参数来分析各参数对算法精度的影响.旋转系数对算法的影响是以旋转间隔角的变化来体现的,通过变换旋转间隔角以分析其对分类精度的影响.缩放系数方面,可以通过缩放尺度变化以分析其对分类精度的影响.而锚点数,则通过增加哈希模型中改变锚点的个数来分析其对分类精度的影响.下面以 NWPU VHR-10 为例,图 5 给出了本文方法在哈希码为 8 字节、32 字节时 MAP 随参数变化的影响情况,其中旋转间隔角以 15、30、45、90、180、360 完成实验;缩放比例按照 0.75、0.5、0.25、(0.5,0.75)、(0.25,0.75)、(0.25,0.5)作对比(不作旋转);锚点数则给出了 500、1 000、1 500、2 000、4 000 个锚点作为对比参数.

图 5 所示实验结果表明,当算法的 3 个参数在 32 字节时对 MAP 的提升幅度小于 8 字节时,其中不同字节下锚点数的增长对算法的精度都有稳定性提升,这是来源于 SDH 方法引入嵌入学习对算法精度提升的效果.在 8 字节时,随着旋转间隔角的减小,算法的精度不断提升,直至 15°后( $R=12$ )趋于平稳;而在缩放系数方面,在 0.5 和 0.75 这两个尺度下,达到算法精度的极值点;当缩放系数小于 0.5 时,算法精度则有所降低.为此,本文选取(0.5, 0.75)两个尺度相结合作为缩放系数;而从锚点数的影响上来看,当锚点数小于 2 000 时,随着锚点数的增加,分类

精度稳步提升,而大于 2 000 后,分类精度则基本不发生变化,因此实验中选取 2 000 作为锚点数.从定量上来看,锚点数可以对分类精度带来 10%~15%的稳步提升,而旋转系数和缩放系数在 8 字节下对算法精度影响更大,有 5%~10%的提升.分析这种现象的原因在于,本文算法能够在低字节下更充分地挖掘样本信息并做出更有效的哈希码表达,而随着字节数的增大,哈希码承载的维度越高,表达信息就越多,算法的优势也就不再明显,从而 MAP 的提升有限.

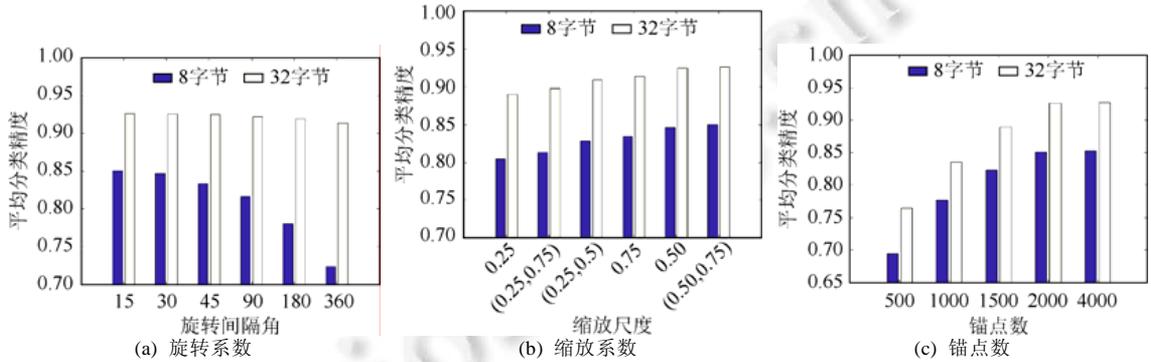


Fig.5 MAP of the proposed method affected by the parameter in 8, 32 bits

图 5 在 8 字节、32 字节下本文方法参数变化对 MAP 的影响

### 3.2.2 不同分类器方法的性能对比

为了验证仿射不变哈希方法在分类方法上的优越性,本文选取了  $K$ -NN<sup>[12]</sup>、RF<sup>[13]</sup>、SVM<sup>[14]</sup>、SRC<sup>[15]</sup>作为对比方法以验证算法的精度,以 MAP 和运算用时作为评价标准.表 6 给出几种分类方法的对比情况,其中,稀疏表示分类器采用 L1 范数的 L1-SRC 测试,本文方法给出了 32 字节的数据.

Table 6 MAP and time consuming comparison of different classification methods

表 6 不同分类方法的平均分类精度和用时对比

| 方法                  | 数据集 | NWPU VHR-10 |        | RSOD-Dataset |        |
|---------------------|-----|-------------|--------|--------------|--------|
|                     |     | 平均分类精度      | 时间(s)  | 平均分类精度       | 时间(s)  |
| $K$ 近邻分类法(512-Gist) |     | 0.761 2     | 23.45  | 0.854 3      | 40.61  |
| 随机森林(512-Gist)      |     | 0.794 5     | 84.52  | 0.845 6      | 185.48 |
| 支持向量机(512-Gist)     |     | 0.849 3     | 59.74  | 0.926 4      | 112.82 |
| L1 范数稀疏表示(512-Gist) |     | 0.927 0     | 136.88 | 0.964 5      | 295.29 |
| 仿射不变离散哈希(32 bits)   |     | 0.926 3     | 8.35   | 0.959 0      | 15.13  |

表 6 给出了 5 种方法在两个数据集下的 MAP 和所用时间对比情况.首先看 MAP 数据,稀疏表示分类器方法的分类精度与本文方法不相上下,而其余的经典分类方法的精度都低于这两种方法;而在 RSOD-Dataset 数据集上,这种精度优势则有所缩小. $K$  近邻分类方法和 RF 方法精度相近, $K$ -NN 在样本数较多的 RSOD-Dataset 数据集上精度略高于 RF 的算法精度.SVM 方法以其小样本学习的优势具有较好的分类精度,在几种方法中排在中间的位置,换到类别数较少的 RSOD-Dataset 数据集上,也很接近于 SRC 和 AIDH.

再比较一下几种方法运行耗时情况.比较两个数据集后发现,RSOD-Dataset 数据集比 NWPU VHR-10 数据集的用时要多出一倍左右,这是因为 RSOD-Dataset 的样本数比 NWPU VHR-10 的样本数要多出一倍;而各种方法随着样本数的增加,所消耗的时间也随之成倍增加.再比较各分类算法, $K$ -NN 方法虽然精度较低,但是运行时间却是很快的,在 NWPU VHR-10 数据集下只需要 20 多秒即完成了运算,并且两个数据集下都很稳定.综合考虑,在样本数适当提高的情况下, $K$ -NN 方法在效率上是可观的.SVM 和 RF 都是比较耗时的方法,在小数据集上也花去了 1 分钟左右的时间.而在精度上有优势的 L1-SRC 方法,由于算法随样本数增长时运算量过大同时也更占用内存,因而运行时间也增加很多,在大数据集下花掉了将近 5 分钟的时间.本文方法只在式(3)的锚点计算以及式(5)的约束平均项计算上消耗了一点时间,哈希码迭代的次数也仅在 6 次之内完成,而且哈希码的存储单位

只在 32 字位上,因此,算法的运行时间很短,大约是 10s 左右。

此外,再比较几种方法的机器存储类型.几种对比方法的特征都是 512 维的特征,在 RSOD-Dataset 数据集下,它们占用的存储空间达到了 15MB.即使由于仿射变换扩大了数据集构成,32 位的 AIDH 方法的样本存储量也只有 767KB.以 32 位的二值码代替 512 维的浮点型数据,哈希方法的优势是大幅降低算法用时的主要原因。

综上所述,相比于其他几种分类方法,本文方法以其数据存储和编码优化的优势,在较短的运算时间内获得了更高的分类精度。

## 4 结 论

本文首先回顾了经典分类方法在遥感图像处理中的优势与不足,将兼具低存储和高效性的哈希方法引入到遥感多目标分类领域中.为解决哈希方法在仿射变换数据监测的有效性,构建仿射不变离散哈希模型.通过在监督离散哈希方法中,加入仿射不变项约束,迫使同型号目标分享相同的二值码,增加同类目标哈希码的相似性,实现哈希方法的仿射不变性.为了验证算法的有效性,本文采用 NWPU VHR-10 和 RSOD-Dataset 数据集作为样本库进行了实验.实验中,以 6 种不同的哈希方法以及 4 种不同的分类方法作对比,并以平均分类精度、运算用时等作为定量指标评价.实验结果表明,相比于各对比方法,本文方法改善了几个样本类别数较少的目标类的精度,从而取得了更高的总体分类精度.此外,哈希方法低存储的优势使得本文方法相较于经典分类方法在速度上得到了提升。

## References:

- [1] Zhang NJ, Zhang J, Zhang X, Lang HT. Task distribution balancing for parallel two-parameter CFAR ship detection. *Journal of Remote Sensing*, 2016,2:344–351 (in Chinese with English abstract).
- [2] Zhen JX, Fu J, Fu X. Aircraft target recognition in remote sensing images based on distribution of the feature points and invariant moments. *Journal of Image and Graphics*, 2014,4:592–602 (in Chinese with English abstract).
- [3] Chen Z, Ma HC, Zhang L. Cloverleaf interchange boundary extraction from airborne LiDAR data based on advanced neighborhood structure and contour analysis. *Journal of Remote Sensing*, 2013,17(5):1146–1157 (in Chinese with English abstract).
- [4] Cheng G, Han JW, Zhou PC, Guo L. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS Journal of Photogrammetry & Remote Sensing*, 2014,98(1):119–132.
- [5] Cheng G, Zhou PC, Yao XW, Yao C, Zhang YB, Han JW. Object detection in VHR optical remote sensing images via learning rotation-invariant HOG feature. In: *Proc. of the Int'l Workshop on EORSA*. 2016. 433–436.
- [6] Chen XM. The study of disaster target automatic classification based on high-resolution remote sensing images [Ph.D. Thesis]. Beijing: China University of Geosciences, 2016 (in Chinese with English abstract).
- [7] Hou YT, Peng JY, Hao LW, Wang R. Research of classification method for natural images based on adaptive feature-weighted  $K$ -nearest neighbors. *Application Research of Computers*, 2014,31(3):957–960 (in Chinese with English abstract).
- [8] Bosch A, Zisserman A, Munoz X. Image classification using random forests and ferns. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. 2007. 14–21.
- [9] Wu W, Nie JY, Gao GL. Improved SVM multiple classifiers for image annotation. *Computer Engineering & Science*, 2015,37(7): 1338–1343 (in Chinese with English abstract).
- [10] Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y. Robust face recognition via sparse representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2009,31(2):210–227.
- [11] Zhao LJ, Tang P. Scalability analysis of typical remote sensing data classification methods: A case of remote sensing image scene. *Journal of Remote Sensing*, 2016,20(2):157–171 (in Chinese with English abstract).
- [12] Wang XM, Zhang HL. Hyperspectral remote sensing image classification using geodesic-based KNN. *Journal of Shanxi Coalmining Administrators College*, 2013,26(4):135–137 (in Chinese with English abstract).
- [13] Pal M. Random forest classifier for remote sensing classification. *Int'l Journal of Remote Sensing*, 2005,26(1):217–222.
- [14] Zhao CH, Liu W, Xu Y, Wen JH. A spectral-spatial SVM-based multi-layer learning algorithm for hyperspectral image classification. *Remote Sensing Letters*, 2018,9(3):218–227.

- [15] Wu SL, Chen HD, Bai Y, Zhu GK. A remote sensing image classification method based on sparse representation. *Multimedia Tools and Applications*, 2016,75(19):12137–12154.
- [16] Bulley H, Bishop MP, Shroder JF. Integration of classification tree analyses and spatial metrics to assess changes in supraglacial lakes in the Karakoram Himalaya. *Int'l Journal of Remote Sensing*, 2012,34(2):387–411.
- [17] Sisodia PS, Tiwari V, Kumar A. Analysis of supervised maximum likelihood classification for remote sensing image. In: *Recent Advances & Innovations in Engineering*. 2014. 1–4.
- [18] Cheng G, Zhou PC, Han JW. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Trans. on Geoscience and Remote Sensing*, 2016,54(12):7405–7415.
- [19] Meher Sk. Knowledge-encoded granular neural networks for hyperspectral remote sensing image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015,8(6):2439–2446.
- [20] Wang JD, Shen HT, Song JK, Ji JQ. Hashing for similarity search: A survey. *Computer Science*, 2014, 1–29.
- [21] Gionis A, Indyk P, Motwani R. Similarity search in high dimensions via hashing. In: *Proc. of the 25th VLDB Conf.* 1999,8(2): 518–529.
- [22] Weiss Y, Torralba A, Fergus R. Spectral hashing. In: *Proc. of the Int'l Conf. on Neural Information Processing Systems (NIPS)*. 2008,282(3):1753–1760.
- [23] Jiang QY, Li WJ. Scalable graph hashing with feature transformation. In: *Proc. of the Int'l Conf. on Artificial Intelligence (ICAI)*. 2015,9(3):2248–2254.
- [24] Shen FM, Shen CH, Shi QF, Hengel AVD, Tang ZM. Inductive hashing on manifolds. In: *Proc. of the Computer Vision and Pattern Recognition (CVPR)*. 2013. 1562–1569.
- [25] Liu W, Wang J, Ji RR, *et al.* Supervised hashing with kernels. In: *Proc. of the Computer Vision and Pattern Recognition (CVPR)*. 2012. 2074–2081.
- [26] Norouzi M, Blei DM. Minimal loss hashing for compact binary codes. In: *Proc. of the Int'l Conf. on Machine Learning (ICML)*. 2011. 353–360.
- [27] Shen FM, Shen CH, Liu W, Shen HT. Supervised discrete hashing. In: *Proc. of the Computer Vision and Pattern Recognition (CVPR)*. 2015. 37–45.
- [28] Kang WC, Li WJ, Zhou ZH. Column sampling based discrete supervised hashing. In: *American Association for Artificial Intelligence (AAAI)*. 2016. [http://cs.nju.edu.cn/lwj/paper/AAAI16\\_COSDISH.pdf](http://cs.nju.edu.cn/lwj/paper/AAAI16_COSDISH.pdf)
- [29] Xu H, Liu YZ, Sun QS. Object classification of remote sensing images based on rotation-invariant discrete hashing. In: *Proc. of the Pacific RIM Conf. on Multimedia*. 2017. 264–274.
- [30] Shen XB. Study of multi-view embedding learning techniques with applications [Ph.D. Thesis]. Nanjing: Nanjing University of Science and Technology, 2017 (in Chinese with English abstract).
- [31] Tang T. Affine invariant feature and its application to target recognition in remote sensing images [Ph.D. Thesis]. Changsha: National University of Defense Technology, 2006 (in Chinese with English abstract).
- [32] Wang FG, Feng XC, Zhang XB. Stationary wavelet transform for affine invariant image object recognition. *Computer Engineering and Applications*, 2007,43(21):239–241 (in Chinese with English abstract).
- [33] Zhang JY, Chen Q, Bai XJ, Sun QS, Xia DS. Affine invariant feature extraction algorithm based on generalized canonical correlation analysis. *Journal of Electronics & Information Technology*, 2009,31(10):2465–2469 (in Chinese with English abstract).
- [34] Gao F, Wen XJ. A new method for affine invariants extraction based on affine geometry. *Journal of Image and Graphics*, 2011, 16(3):389–397 (in Chinese with English abstract).
- [35] Kong J, Sun QS, Ji ZX, Liu YZ. A novel fast object detection method in remote sensing image based on affine-invariant supervised discrete hashing. *Journal of Nanjing University Natural Science*, 2019,55(1):49–60 (in Chinese with English abstract).
- [36] Long Y, Gong YP, Xiao ZF, Liu Q. Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Trans. on Geosciences and Remote Sensing*, 2017,55(5):2486–2498.
- [37] Liu J, Guo J, He ZL. Scene classification based on gist and PHOG feature. *Computer Engineering*, 2015,41(4):232–235 (in Chinese with English abstract).

## 附中文参考文献:

- [1] 张临杰,张杰,张晰,郎海涛.任务分配均衡的双参数 CFAR 舰船检测并行算法.遥感学报,2016,2:344-351.
- [2] 曾接贤,付俊,符祥.特征点和不变矩结合的遥感图像飞机目标识别.中国图像图形学报,2014,4:592-602.
- [3] 陈卓,马洪超,张良.改进邻域结构与轮廓分析的 LiDAR 点云立交桥提取.遥感学报,2013,17(5):1146-1157.
- [6] 程希萌.基于高分遥感影像的灾害目标自动分类技术[博士学位论文].北京:中国地质大学,2016.
- [7] 侯玉婷,彭进业,郝露微,王瑞.基于 KNN 的特征自适应加权自然图像分类研究.计算机应用研究,2014,31(3):957-960.
- [9] 吴伟,聂建云,高光米.一种基于改进的支持向量机多分类器图像标方法.计算机工程与科学,2015,37(7):1338-1343.
- [11] 赵理君,唐姆.典型遥感数据分类方法的适用性分析——以遥感图像场景分类为例.遥感学报,2016,20(2):157-171.
- [12] 王小美,张红利.基于测地距离的 KNN 高光谱遥感图像分类.山西煤炭管理干部学院学报,2013,26(4):135-137.
- [30] 沈肖波.多视图嵌入学习方法及其应用研究[博士学位论文].南京:南京理工大学,2017.
- [31] 唐涛.图像仿射不变特征及其在遥感图像目标识别中的应用[博士学位论文].长沙:国防科技大学,2006.
- [32] 王凤国,冯象初,张小波.平稳小波变换在仿射不变性目标识别中的应用.计算机工程与应用,2007,43(21):239-241.
- [33] 张洁玉,陈强,白小晶,孙权森,夏德深.基于广义典型相关分析的仿射不变特征提取方法.电子与信息学报,2009,31(10):2465-2469.
- [34] 高峰,文项坚.利用仿射几何的仿射不变特征提取方法.中国图像图形学报,2011,16(3):389-397.
- [35] 孔颖,孙权森,纪则轩,刘亚洲.基于仿射不变离散哈希的遥感图像快速目标检测新方法.南京大学学报(自然科学),2019,55(1):49-60.
- [37] 刘静,郭建,贺遵亮.基于 Gist 和 PHOG 特征的场景分类.计算机工程,2015,41(4):232-235.



孔颖(1985—),男,安徽舒城人,硕士,主要研究领域为遥感图像处理,机器学习与应用.



刘亚洲(1979—),男,博士,副教授,主要研究领域为计算机视觉,模式识别.



孙权森(1963—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为模式识别,机器学习,遥感信息处理.



纪则轩(1984—),男,博士,副教授,CCF 专业会员,主要研究领域为模式识别,图像处理,医学影像学.



徐晖(1992—),女,硕士,主要研究领域为模式识别,图像处理.