

6-8-2020

Research of Remote Sensing Image Classification Technology Based on Multi-feature Combining and BoW Model

Li Ke

1. The Information Engineering University, Zhengzhou 450052, China;;

You Xiong

1. The Information Engineering University, Zhengzhou 450052, China;;

Du Lin

1. The Information Engineering University, Zhengzhou 450052, China;;2. Troop 72515, PLA, Jinan 250014, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Research of Remote Sensing Image Classification Technology Based on Multi-feature Combining and BoW Model

Abstract

Abstract: *A new algorithm of image classification of multi feature combination and BoW model was proposed. SIFT, GIST, Census and Gabor color, and many other types of features were extracted from the images, and then through the experimental analysis to determine the best feature combination. According to the general K-means algorithm which did not consider the weight of each features, different feature component was put forward by using automatic weighted k-means algorithm, respectively SIFT, GIST, Gabor feature construct weights based on image features of vocabulary, using the soft coding algorithm for image coding, and using the SVM algorithm to complete the image classification. Experiments show that this method can effectively improve the classification accuracy of images.*

Keywords

multi-feature combining, k-means, bag of words, SVM

Recommended Citation

Li Ke, You Xiong, Du Lin. Research of Remote Sensing Image Classification Technology Based on Multi-feature Combining and BoW Model[J]. Journal of System Simulation, 2016, 28(6): 1386-1393.

基于多特征组合与优化 BoW 模型的影像分类技术研究

李科¹, 游雄¹, 杜琳^{1,2}

(1. 信息工程大学, 郑州 450052; 2. 72515 部队, 济南 250014)

摘要: 提出一种基于多特征组合与优化 BoW 模型的影像地物分类新方法。提取影像的 SIFT、GIST、颜色、Census 和 Gabor 等多种类型特征, 通过实验分析确定最佳特征组合。针对一般 K-Means 算法没有考虑各个特征值的权重, 提出利用自动加权 k-Means 算法计算不同特征分量的权值, 分别对 SIFT、GIST、Gabor 特征构建了基于权重的影像特征词汇表, 采用基于 Soft 的词汇编码算法进行影像编码, 使用 SVM 算法完成影像分类。通过实验表明方法能有效提高遥感影像分类准确性, 并且具有较好的稳定性和鲁棒性。

关键词: 特征组合; K-means; 视觉词袋; 支持向量机

中图分类号: P237.3

文献标识码: A

文章编号: 1004-731X (2016) 06-1386-08

Research of Remote Sensing Image Classification Technology Based on Multi-feature Combining and BoW Model

Li Ke¹, You Xiong¹, Du Lin^{1,2}

(1. The Information Engineering University, Zhengzhou 450052, China; 2. Troop 72515, PLA, Jinan 250014, China)

Abstract: A new algorithm of image classification of multi feature combination and BoW model was proposed. SIFT, GIST, Census and Gabor color, and many other types of features were extracted from the images, and then through the experimental analysis to determine the best feature combination. According to the general K-means algorithm which did not consider the weight of each features, different feature component was put forward by using automatic weighted k-means algorithm, respectively SIFT, GIST, Gabor feature construct weights based on image features of vocabulary, using the soft coding algorithm for image coding, and using the SVM algorithm to complete the image classification. Experiments show that this method can effectively improve the classification accuracy of images.

Keywords: multi-feature combining; k-means; bag of words; SVM

引言

遥感影像数据是虚拟地理环境仿真重要的数据源, 随着各种航空、航天、地面传感器广泛的应用, 各行各业采集了海量的、高分辨率的影像数据,

这些数据为地理环境仿真提供了坚实的数据基础。影像分类技术可以从海量的遥感影像数据中自动、智能的解译出有用的空间环境信息, 是当前地理环境仿真应用研究的热点、难点问题。国内外许多学者对影像特征描述和分类技术进行了广泛深入的研究, 其中特征词汇模型(Bag of Words, BoW)是一种有效的影像表达模型。该模型首先获取影像的特征向量, 通过聚类算法建立视觉词汇表, 然后将一副图像解析成大量的视觉单词, 每个视觉单词都是特征量化后的局部特征描述子, 最后用得到的影



收稿日期: 2015-01-05 修回日期: 2015-03-02;
基金项目: 国家自然科学基金(41201390), 国家“863”
计划(2013AA12A202);
作者简介: 李科(1977-), 男, 河北沧州, 博士, 副教授, 研究方向为地理空间数据工程; 游雄(1962-),
男, 福建罗源, 博士, 教授, 研究方向为虚拟地理环境; 杜琳(1982-), 女, 河南开封, 硕士, 工程师, 研究方向为地理信息工程。

<http://www.china-simulation.com>

• 1386 •

像视觉单词直方图来训练分类器。该模型将影像信息映射成视觉关键词的集合, 这样既保存了图像的局部特征信息又压缩了图像特征的维数^[1]。在特征词汇模型中, 对分类精度影响较大的因素有: 影像特征种类、词汇表构建算法以及词汇编码算法。许多学者针对这些关键因素进行了深入研究和优化, 在影像特征方面, 文献[1]将形状和颜色两种特征进行了组合应用, 在特征组合的基础上, 引入了稠密的特征空间, 使得影像的特征表述更完善; 文献[2]对 SIFT 特征(Scale Invariant Feature Transform, SIFT) 和 HOG 特征 (Histogram of Oriented Gradients, HOG)进行组合使用, 减少特征提取对影像描述的损失; 文献[3]整合了像素级和对对象级两种特征。在词汇编码算法方面, 文献[4]对 BoW 模型中的词汇编码方式进行了优化, 采用了 LDA 模型进行影像词汇编码, 通过语义模型来表达影像, 减少影像描述的损失; 文献[5] BoW 模型中的词汇编码方式采用了稀疏编码方式, 论文通过大量实验表明稀疏性具有一定的去噪能力。从分类结果上, 这些算法的改进对影像分类精度均有一定程度提升。从这些文献可以看出, 目前研究主要集中在

特征选取以及词汇编码方面, 对词汇表构建优化算法研究较少。在特征选取方面, 一般选取少量特征, 这样无法从纹理、颜色、形状、空间位置等方面综合描述影像, 导致影像分类算法适应性不好等问题。另外, 当前对 BoW 模型优化算法中, 对词汇表构建算法改进较少, 主要以聚类算法为主, 例如 k-Means 算法, 该算法简单高效, 应用广泛, 但该算法没有考虑影像特征的权值, 因此对分类结果有较大影响。

针对已有方法的缺点, 本文提出一种基于多特征组合与优化 BoW 模型的影像分类新方法。针对影像多种不同类型的特征, 通过大量实验分析确定最佳特征组合, 然后提出利用自动加权 k-Means 算法计算不同特征的权值, 根据特征数值的权重构建特征词汇表。然后利用基于 Soft 算法进行影像词汇编码, 达到影像综合表达的效果。最后将颜色特征、Census 特征, 多个不同种类的词汇特征相结合, 对每幅图像形成一个多维特征矩阵, 代入 SVM 模型进行分类训练, 找出其最可能属于的主题或者类别, 从而完成影像分类。如图 1 所示。

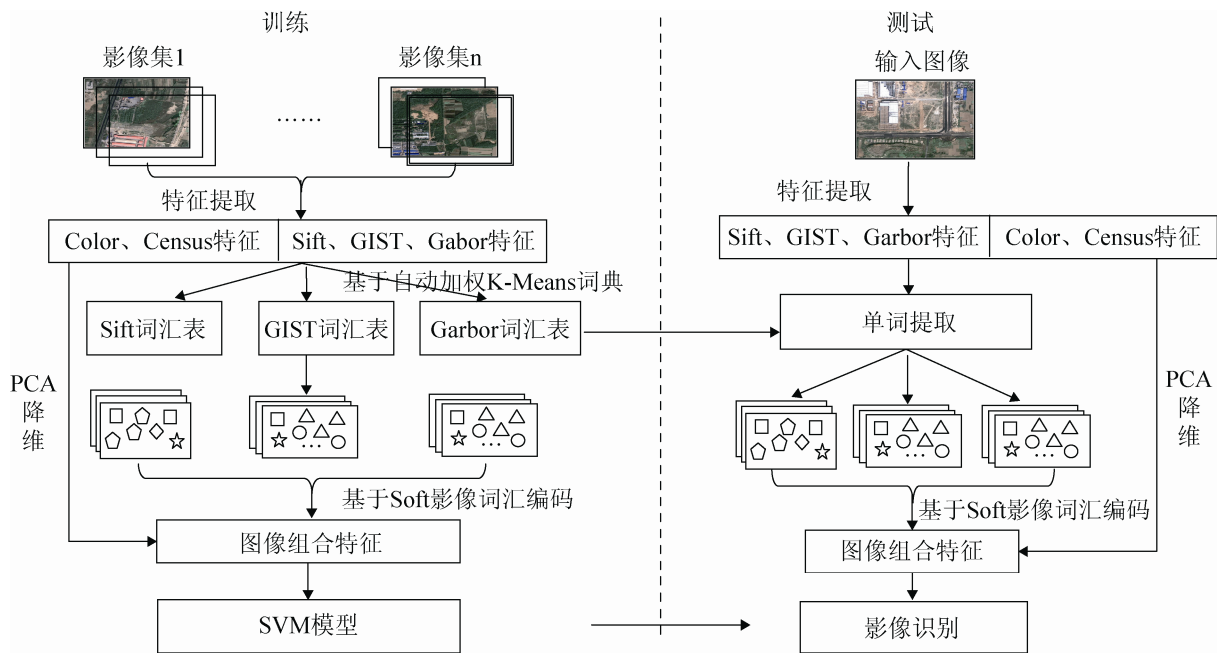


图 1 算法流程图

该算法分为 4 个阶段。(1) 多特征提取, 为弥补单一特征无法充分表达影像的缺点, 综合比较各种颜色、纹理、尺度、形状特征, 结合遥感影像的特点提取每幅影像的 SIFT、GIST、Gabor、Color、Census 特征集;(2) 基于权值的视觉词典生成, 采用自动加权 K-Means 聚类算法分别对训练影像数据集中 SIFT、GIST、Gabor 特征聚类, 生成视觉词典来提高分类结果的准确性和鲁棒性, 此外自动计算各种特征的影响因子, 并得出各种特征的权重系数。(3) 基于 Soft 的影像词汇表达, 将每幅图像的特征向视觉词典的每个单词通道进行投影, 即计算每幅图像的特征在每个词汇上的特征分布。(4) 影像分类, 将颜色特征、Census 特征, 词汇特征相结合, 对每幅图像形成一个多维特征矩阵, 代入 SVM 模型进行分类训练, 找出其最可能属于的主题或者类别, 从而完成影像分类。

1 影像特征分析与提取

相关文献^[2,5-6]的实验结果表明将若干种特征进行组合, 分类效果要优于单特征, 但并非使用特征的种类越多分类效果就越好。目前用于影像分类的特征主要包括: 纹理、颜色、形状、空间位置等特征描述^[7]。它们对影像的表达各有所侧重, 本文运用 5 种具有代表性的图像特征进行组合应用:

(1) GIST 特征^[8], 它模拟人的视觉提取图像中粗略但简明扼要的结构信息, 也就是将一组多方向多尺度 Gabor 小波函数基与图像进行卷积计算, 因此能够检测出图像中具有相应的方向频率信息的显著特征。假设灰度图像 $f(x, y)$ 用 m 尺度 n 方向的 Gabor 滤波器组进行滤波, 即分别同 t_c 个通道的滤波器进行卷积, 其中 $t_c = m \times n$, 再级联卷积的结果, 称为图像 GIST(IG)特征, 即

$$G^I(x, y) = \underset{t}{cat}(f(x, y) * g_{mm}(x, y)) \quad (1)$$

式中: cat 为级联运算符; $*$ 为卷积运算符; G^I 的维数为 $t_c \times r \times c$ 。由于颜色信息是影像十分重要特征, 因此, 本文将颜色空间引入到 GIST 特征中, 利用 3 个尺度, 各 8 个方向的 Gabor 滤波器组提取

影像的 GIST 特征(Gabor 滤波器 3 个尺度对应的频率分别为 0.3, 0.162 2, 0.087 7, 8 个方向分别为 $0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8$)。如图 2(b)所示, 是 24 个 Gabor 滤波器分别对图像 R, G, B 三个通道进行滤波后级联的结果。

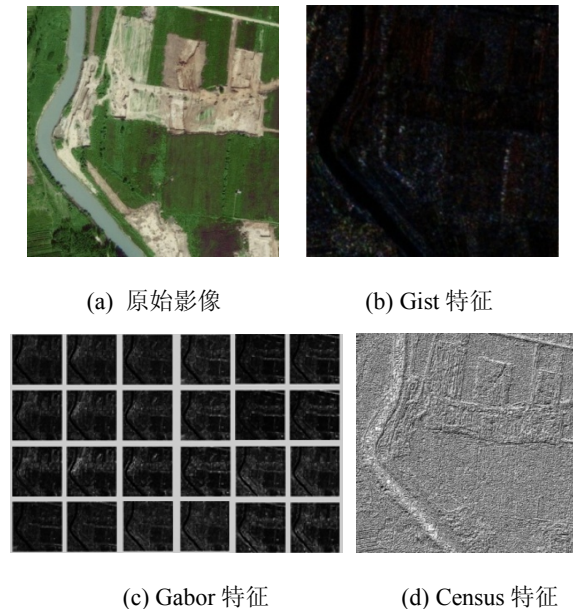


图 2 图像特征提取

(2) 尺度不变特征变换(SIFT)^[9]。它对平移、旋转、尺度缩放、亮度变化、遮挡和噪声等具有好的不变性, 对观察视角的变化、仿射变换也具有一定的稳定性, 属于局部结构特征, 因此在影像分类中得到了广泛应用。

(3) 颜色直方图, 它描述的是不同颜色特征在整幅图像中所占的比例, 反映了图像颜色的统计分布, 属于全局特征。由人类视觉模型可知, 人眼能感知的色彩在 Lab 模型中进行计算具有简单快速等特点。因此, 本文将原始影像从 RGB 颜色空间转化到 Lab 颜色空间, 在 Lab 颜色模型上进行颜色直方图统计, 提取影像的颜色特征分布。

(4) Gabor 特征, 是对原始影像通过二维 Gabor 函数计算得到的特征, 适用于检测和描述图像纹理特征^[4,10]。多尺度多方向 Gabor 滤波器组是在 Gabor 滤波器 $g(x, y)$ 基础上通过尺度和旋转变换扩展而成的多通道滤波方案, 是自相似 Gabor 小波, 即

$$\begin{aligned} g_{mn}(x, y) &= a^{-m} g(x', y'), \quad a > 1 \\ x' &= a^{-m} (x \cos \theta + y \sin \theta) \\ y' &= a^{-m} (-x \sin \theta + y \cos \theta) \\ \theta &= n\pi / (n+1) \end{aligned} \quad (2)$$

式中: θ 为滤波器的方向; a^{-m} 为母小波膨胀的尺度因子。 n 为滤波器组的方向数, m 为滤波器组的尺度数, 改变 m, n 的数值, 可以得到不同的 Gabor 小波滤波器组。特征可视化结果如图 2(c)所示。

(5) Census 变换直方图主成份分析^[4,11], 它是一种基于局部图像像素灰度值比较而提出的纹理特征。其核心思想是将二值矩阵看作一个二值纹理模式(纹理基元), 通过比较一个像素点和它周围的八邻域像素点的灰度值来获得变换值。其计算公式为:

$$\begin{aligned} f(x) &= \bigotimes_{p \in W(p)} \xi(I(x), I(x')) \\ \xi(I(x), I(x')) &= \begin{cases} 1 & I(x) < I(x') \\ 0 & \text{其它} \end{cases} \end{aligned} \quad (3)$$

式中: $f(x)$ 为像素 x 对应的特征值; $w(x)$ 为中心像素 x 对应的 Census 变换窗口; $I(x)$ 为像素 x 的灰度值。如图 2(d)所示, 该特征保留了图像的全局特征和局部特征。另外, Census 特征是一个 256 维的向量, 根据实验分析, 该特征向量具有较高的稀疏性, 因此对 Census 特征进行主成分分析, 本文使用 40 个主要分量值对特征向量进行降维。

综上所述, SIFT 与 GIST 特征主要关注影像的局部结构特征, colorhist 属于图像的全局特征, 主要侧重图像颜色分布特征的描述, 而 Census 与 Gabor 特征侧重于图像的纹理。本文对这些特征进行优化改进, 在后面的实验中, 将其组合应用于遥感图像场景分类。

2 视觉词汇表构建

在视觉词汇表构建时, 使用最广泛是无监督聚类算法, 例如 K-Means 算法。然而, 在影像特征描述符中的每个特征维度的取值范围不同, 并且对于影像分类来说, 每个维度具有不同的重要性。但

是, 传统的 K-Means 算法不能很好的区分不同特征的重要性。因此, 为了使构建的视觉词汇表更有效, 本文采用了加权 K-Means 算法, 该算法应用了明科斯基距离和特征加权^[12]。

假设 M 维的 2 个特征 $F_1 = (F_1^1, \dots, F_1^M)$ 和 $F_2 = (F_2^1, \dots, F_2^M)$, 其明科斯基距离计算公式为:

$$d_\beta(F_1, F_2) = \left(\sum_{m=1}^M |F_1^m - F_2^m|^\beta \right)^{1/\beta} \quad (4)$$

在 K-Means 算法中, 通过迭代计算将所有特征向量到中心的距离之和最小化, 公式如下:

$$W(S, C, w) = \sum_{k=1}^K \sum_{i=1}^N s_{ik} w^\beta d_\beta(x_i, c_k) \quad (5)$$

式中: N 为特征个数; K 为聚类中心数量; $S = \{s_1, s_2, \dots, s_K\}$ 是一个包含非重叠的子集; s_K, s_{iK} 是一个 $\{0, 1\}$ 变量。假设当 $F_i \in S_K$ 时, $s_{iK} = 1$; 其它, $s_{iK} = 0$ 。 $C = \{c_1, c_2, \dots, c_K\}$ 是子集的中心, w 是权重系数。通过 3 组变量(子集, 中心和权重), 将聚类最小化, 其中权重的初始值平均分布, 通过迭代不断优化。简要计算过程如下:

(1) 假定重心 C_K 和权值 w , 利用公式(4), 在明科斯基距离空间, 根据最小距离准则, 不断更新特征子集的划分。

(2) 假定子集 S_K 和权重 w , 不断更新每个聚类 S_K 的中心 C_K , 作为明科斯基中心。 C_{KV} 的值由 c 的最小值定义, 其公式如下:

$$d(c) = \sum_{i \in S} |y'_i - c|^\beta \quad (6)$$

其中 $y'_{iv} = (w_v, v_{iv})$, v 是特征的维度索引。

(3) 假定子集 S_K 和中心 C_K , 不断更新权重, 主要依据是第一个步骤中最优条件最小化公式(4)的约束, 并且权值的总和为 1:

$$W_v = \frac{1}{\sum_{u=1}^M \left| \frac{D_{v\beta}}{D_{u\beta}} \right|^{1/\beta-1}} \quad (7)$$

其中 $D_{V\beta} = \sum_k \sum_{i=1}^N s_{ik} |y_i - c_{kv}|^\beta$ 。

3 影像特征词汇编码

在视觉词汇表构建完成后,需计算一幅图像不同区域特征在每个词汇上的特征分布,即计算图像不同区域的特征词汇编码,代表性算法主要有 VQ(Vector Quantization)和局部稀疏编码等算法^[11-12]。VQ 算法是利用一个最近邻视觉单词来代替图像的某些区域,因此可能造成过多图像特征信息的丢失。局部稀疏编码是利用少量不同的视觉单词来线性拟合图像,减少了图像特征信息丢失,降低图像特征的重构误差,因此其算法比 VQ 算法更优,但是局部稀疏编码可能出现负数,这个特性为后续的图像分类带来一些问题,有时会导致分类精度较低。

针对 VQ 和局部稀疏编码算法的缺点,本文提出了基于 Soft 的词汇编码算法,其基本思路是:将一幅图像每个子区域的特征向量向视觉词汇表中的每个单词通道进行投影,即计算每个图像区域的特征向量在每个词汇上的特征分布。假设视觉词汇表是 D 维的特征向量的集合,即 $P = \{p_1, p_2, p_3, \dots, p_N\} \in R_{D \times N}$, N 表示词汇个数。影像中任意子区域的特征向量为 $p(x)$,则特征向量 $p(x)$ 在视觉词汇表上的特征分布表示 $\theta(x) = \{\theta_1(x), \theta_2(x), \dots, \theta_D(x)\}^T$, $\theta(x)$ 是 D 维向量,其中

$$\theta_i(x) = ce^{-\frac{\|p(x)-p_i\|}{\delta}} \quad (8)$$

式中: δ 为各视觉词汇之间距离的最大值; c 为归一化因子,使得 $\sum_{i=1}^D \theta_i = 1$ 。

基于 Soft 的词汇编码算法将图像任意子区域的特征向量转化为视觉词汇表分布直方图,有效避免了 VQ 算法仅利用一个最近邻视觉单词来代替图像子区域特征向量的缺点,也避免了局部稀疏编码具有负数的缺点。因此,本文根据上述公式(5)统计每幅图像的视觉单词词频信息,得到图像的视觉词汇分布直方图向量集合作为图像分类的基本特征向量。

4 实验与分析

4.1 实验设置

为验证提出方法的有效性,本文设计了 4 组实验来验证相关方法。实验中的遥感影像数据均来自于百度地图,共有 350 幅遥感影像,大小为 600 像素×600 像素,主要包括土路、居民地、道路、荒地、耕地、树林、水系等 7 类。

在选择训练样本和测试样本方面,采用了随即挑选了 n 幅影像作为训练样本,剩余的影像作为测试样本。在本文实验中 n 分别取 50, 100, 150, 200, 250, 300, 每次随即选择训练影像和测试影像,重复 5 次,文中实验结果是 5 次实验的平均值。

4.2 实验流程

首先利用超像素分割算法将一幅遥感影像中的颜色、纹理等相似的像素划分为同一区域,如图 3 所示;接着对每个图像块分别提取纹理、颜色、形状、空间位置等特征,并通过大量实验确定了最佳特征组合;然后利用自动加权 k-Means 聚类算法分别将影像的 SIFT 特征、GIST 特征、Gabor 特征汇聚成 3 个最优视觉词汇表,每个词汇表的个数均为 300 个单词;然后将每幅图像的特征利用 Soft 模型向视觉词汇表的每个单词通道进行投影,即计算每幅图像的特征在每个词汇上的特征分布;最后对每幅影像形成的多维特征矩阵,代入 SVM 模型进行分类训练,找出其最可能属于的主题或者类别,从而完成影像分类。其中支持向量机(SVM)采用 LibSVM 工具包的 libsvm 3.12 工具箱,并使用 RBF 函数作为核函数。



(a) 原始影像 (b) 影像分割结果

图 3 影像分割

在特征提取方面, SIFT 特征、GIST 特征、Gabor 特征均采用网格密集特征, 其中 GIST 特征、Gabor 特征提取了 3 个尺度, 每个尺度 8 个方向(Gabor 滤波器 3 个尺度对应的频率分别为 0.3, 0.162 2, 0.087 7, 8 个方向分别为 0, $\pi/8$, $\pi/4$, $3\pi/8$, $\pi/2$, $5\pi/8$, $3\pi/4$, $7\pi/8$)。由于 SIFT, GIST, Census 特征维度较高, SIFT, GIST, Census 特征维度分别为 128, 72, 256, 导致分类器过拟合从而影响分类精度, 因此本文利用主成份分析, 对上述的特征进行了降维处理, 每种特征的维数如表 1 所示。

表 1 经过 PCA 降维后的特征维数

特征	Sift	GIST	ColorHist	Gabor	Census
维数	40	30	78	24	40

4.3 分类结果

(1) 特征组合分类

本文将 5 个特征进行 10 种不同组合, 在相同分类器 libsvm 下得到分类结果, 如表 2 所示。从表中可以看出, 分类精度并非与特征组合个数成正比

表 2 特征组合分类

个数	特征									
	colorhist + SIFT	Colorhist + GIST	colorhist +census	colorhist +Gabor	GIST+ census	colorhist+ census+GIST	colorhist+ census+SIFT	sift+GIST+ census	sift+GIST +colorhist	SIFT+census colorhist+ GIST
50	0.665	0.648	0.659	0.641	0.619	0.707	0.715	0.615	0.668	0.677
100	0.676	0.667	0.673	0.665	0.637	0.732	0.738	0.636	0.682	0.691
150	0.691	0.692	0.698	0.687	0.651	0.751	0.763	0.649	0.701	0.715
200	0.713	0.719	0.721	0.698	0.668	0.778	0.797	0.664	0.719	0.738
250	0.731	0.728	0.749	0.717	0.673	0.791	0.811	0.683	0.737	0.751
300	0.747	0.745	0.762	0.738	0.685	0.809	0.827	0.691	0.749	0.762

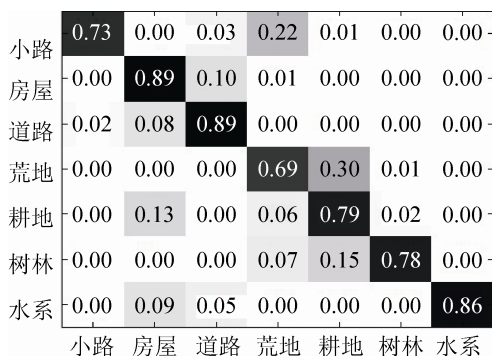


图 4 特征组合得到的分类混淆矩阵

比, 如 GIST+ census 两种特征组合, 分类精度与 GIST 单个特征分类基本相同, SIFT, census, colorhist, GIST 四种特种组合的分类精度比 colorhist, census, GIST 三种分类精度低。将结构、纹理及颜色各取一种特征进行组合, 如 colorhist, census, sift 或 colorhist, Census, GIST 三种特征组合, 分类精度要优于使用其它 5 种特征的组合, 该组合分类精度最高。另外, 从表中可以看出, colorhist, census 两种特征组合, 其分类精度比其它两种特征组合效果好, 这说明 census 和 colorhist 特征具有一定程度的互补性。

图 4 给出了在使用 colorhist, census, SIFT 三个特征优化组合, 训练样本为 260 个情况下所得到的混淆矩阵。矩阵的横轴与纵轴分别代表地物类别, 其对角线上元素的值代表每类地物的分类准确率。从图 4 中可以看出, 小路有很大一部分错分到了荒地中, 这是由于荒地与小路在颜色、纹理等特征上具有一定的相似性, 这是错分的主要因素。居民地与道路容易错分, 另外, 荒地、耕地、树林也容易错分, 这 3 类地物在局部特征、纹理特征十分相似。

(2) 词汇表构建

在上述实验的基础上, 提取了影像的 colorhist、census、SIFT 三种特征, 然后对 SIFT 特征分别利用自动加权 K-Means 和一般 K-Means 算法构建视觉词汇表, 每个词汇表的个数均为 300 个单词, 然后利用 SVM 算法进行影像分类, 结果如表 3 所示。从结果可以看出, 自动加权 K-Means 算法较好的利用不同特征分量的权重, 使视觉单词

构建更为合理，从而提高了影像分类精度。

表 3 基于不同视觉词汇表分类结果

算法名称	个数					
	50	100	150	200	250	300
自动加权 K-means	0.728	0.771	0.796	0.837	0.853	0.863
一般 K-means	0.707	0.743	0.773	0.806	0.809	0.821

(3) 特征词汇编码

在实验 1, 2 的基础上, 本文分别利用了 VQ、局部稀疏编码以及本文提出的 Soft 算法 3 种方法对每幅影像根据视觉词汇本进行编码, 结果如图 5 所示。从该图可以看出, Soft 编码算法比 VQ 算法分类精度明显具有优势, 与局部稀疏算法相比, 分类精度基本相同, 但由于 Soft 编码具有非负数和算法简单高效的优点, 因此在影像分类中更具有优势。

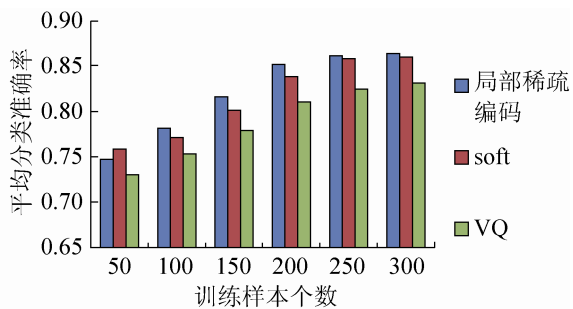


图 5 基于不同特征词汇编码分类结果

(4) 对比实验

为验证提出方法的有效性, 本文设计了与一般 BoW 模型对比实验。在采用一般 BoW 模型实验中, 影像特征分别为 SIFT 特征和 colorhist、census、SIFT 组合特征两种, 采用本文提出的优化算法实验中, 影像特征为 colorhist、census、SIFT, 实验结果如图 6 所示。相对于一般 BoW 模型, 经过优化的 BoW 算法(主要是通过自动加权 K-means 视觉词汇表构建算法和基于 Soft 的词汇编码算法进行优化), 在单一特征和多特征组合方式下, 分别提高了 10.6%和 4.97%。同时也可以看出, 同样采用传统 BoW 模型, 通过多个影像特征组合比单一影像特征的分类精度也有明显提升。

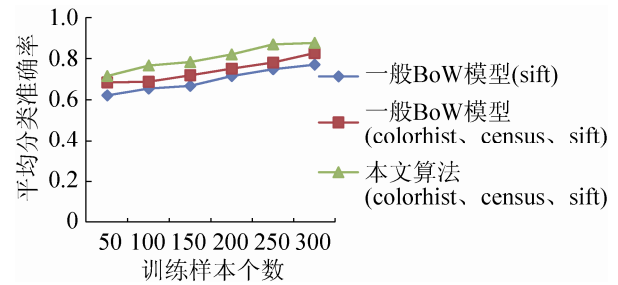


图 6 对比实验分类结果

如图 7 所示, 是根据本文提出的多特征组合与优化 BoW 模型的影像分类新算法的实验结果, 其中图 7 中两幅影像像素分类的正确率分别是 85.537%和 93.19%, 实验中 50 幅测试样本像素分类的平均正确率为 89.836%, 可以看出本文提出的算法对本次测试数据集中的 7 类地物有较好的区分度。其中红色区域是错分的地物, 从这两幅图中可以看出, 错误发生在多种地物混合的区域, 如居民地中有小片的树木, 荒地 with 耕地交叉的地区。

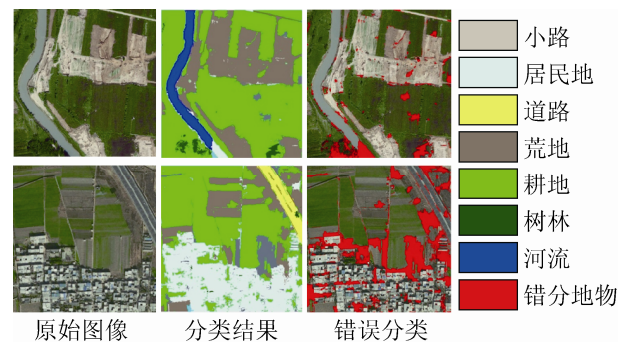


图 7 影像分类结果

5 结论

为了提高遥感影像地物分类精度, 本文提出了一种基于多特征组合与优化 BoW 模型的影像分类新算法。首先提取影像的纹理、颜色、形状、空间位置等特征, 通过大量实验确定了最佳特征组合, 然后利用加权 K-means 聚类算法生成不同视觉词汇本, 采用基于 Soft 的词汇编码算法进行影像编码, 得到训练影像数据集的不同特征组合, 最后利用 SVM 分类器对影像进行分类。实验结果表明, 本文方法能有效提升遥感影像不同地物分类的准确性, 具有较好的稳定性和鲁棒性。

参考文献:

- [1] Harzallah H, Jurie F, Schmid C. Combining Efficient Object Localization and Image Classification [C]// Proc of the 12th IEEE International Conference on Computer Vision, Kyoto, Japan. USA: IEEE, 2009: 237-244.
- [2] Jiang Y, Wang R S. Scene Classification Based on Multi-feature Extended pLSA Model [J]. Signal Processing (S1003-0530), 2010, 26(4): 539-544.
- [3] 李刚, 万幼川. 基于改进的像素级和对象级的遥感影像合成分类 [J]. 测绘学报, 2012, 41(6): 891-903.
- [4] 李宇, 王宏琦, 孙显. 利用多尺度语义模型的复杂图像目标自动提取方法 [J]. 武汉大学学报 (信息科学版), 2012, 37(9): 1064-1067.
- [5] 罗会兰, 郭敏杰, 孔繁胜. 集成多特征与稀疏编码的图像分类方法 [J]. 模式识别与人工智能, 2014, 27(4): 345-355.
- [6] Bronstein A M, Bronstein M M, Guibas L J, et al. Shape google: Geometric words and expressions for invariant shape retrieval [J]. ACM Transactions on Graphics (TOG), (S1067-7055), 2011, 30(1): 1-8
- [7] 徐佩, 陈丽君, 杨文, 等. 利用特征选择的遥感图像场景分类 [J]. 哈尔滨工业大学学报, 2011, 43(9): 117-121.
- [8] Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope [J]. International Journal of Computer Vision (S0920-5691), 2001, 42(3): 145-175.
- [9] LOWE D G. Distinctive image features from scale-invariant key-points [J]. International Journal of Computer Vision (S0920-5691), 2004, 60(2): 91-110.
- [10] 寇程, 柯长青. 小波变换与 FCM 聚类的 QuickBird 影像特征空间识别分析 [J]. 地球信息科学学报, 2010, 12(3): 444-450.
- [11] 程少光, 何毕, 布树辉, 等. 基于超像素空间金字塔模型的场景识别研究 [J]. 计算机应用研究, 2014, 38(7): 152-159.
- [12] De Amorim R C, Mirkin B. Minkowski metric, feature weighting and anomalous cluster initializing in K-Means clustering [J]. Pattern Recognition (S0031-3203), 2012, 45(3): 1061-1075.
- [11] 顾立忠, 苏建波. 仿人机器人的头眼协调运动控制研究 [J]. 机器人, 2008, 30(2): 165-170.
- [12] 毛晓波, 陈铁军. 头眼协调运动控制仿生模型 [J]. 生物医学工程学杂志, 2011, 28(5): 895-900.
- [13] Edwad G. Freedman. Interactions between eye and head control signals can account for movment kinematics [J]. Biological Cybernetics (S0340-1200), 2001, 84(6): 453-462.
- [14] Xiaolin Zhang. A cooperative interaction control methodology of a pair independent control system [C]// Control Automation Robotics & Vision. Singapore: IEEE, 2010: 42-49.
- [15] Wang X, Jin J. A quantatitive analysis for decomposing visual signal of the gaze displacement [C]// Proceeding of the pan-sydney area workshop on visual information processing. Darlinghurst, Australia: Australian Computer Society, Inc, 2001: 153-159.
- [16] Law J, Shaw P, Lee M. A biologically constrained architecture for developmental learning of eye-head gaze control on a humanoid robot [J]. Autonomous Robots (S0929-5593), 2013, 35(1): 77-92.
- [17] Robinson D A. Models of the saccadic eye movement control system [J]. Kybernetik (S0023-5946), 1973, 14(2):71-83.
- [18] Coubard O A, Kapoula Z. Saccades during symmetrical vergence [J]. Graef Arch Clin Exp Ophthalmol (S0721-832X), 2008, 246(4): 521-536.
- [19] Erkelens C J. A dual visual-local feedback model of the vergence eye movement system [J]. Journal of Vision (S1534-7362), 2011, 11(10): 1-14.

(上接第 1385 页)