

# 基于多维金字塔表达和 AdaBoost 的高分辨率 SAR 图像 城区场景分类算法

殷慧<sup>1</sup> 曹永锋<sup>1</sup> 孙洪<sup>1</sup>

**摘要** 提出了多维金字塔表达算法, 并使用基于多维金字塔表达的 AdaBoost 实现了高分辨率合成孔径雷达 (Synthetic aperture radar, SAR) 图像的城区场景分类. 多维金字塔表达算法首先在局部特征的多维计算金字塔表达矢量, 再将所有的金字塔表达矢量连接起来构成多维金字塔表达矢量. 多维金字塔表达算法克服了金字塔表达算法在处理高维局部特征时, 遇到的输出金字塔表达矢量的区分力受计算效率制约的问题. 本文分别在一个 TerraSAR-X 图像库和一张大幅 TerraSAR-X 图像上比较基于金字塔表达的 AdaBoost 和基于多维金字塔表达的 AdaBoost 的分类性能. 实验结果表明, 与前者相比, 后者显著提高了计算效率同时保证了分类精度.

**关键词** 多维金字塔表达, AdaBoost, 高分辨率合成孔径雷达图像, 分类

**DOI** 10.3724/SP.J.1004.2010.01099

## Urban Scene Classification Based on Multi-dimensional Pyramid Representation and AdaBoost Using High Resolution SAR Images

YIN Hui<sup>1</sup> CAO Yong-Feng<sup>1</sup> SUN Hong<sup>1</sup>

**Abstract** This paper presents a new image representation algorithm called multi-dimensional pyramid representation (MPR), and combines MPR and AdaBoost (MPR-AdaBoost) for urban scene classification using high resolution SAR images. MPR calculates a pyramid representation (PR) vector in each dimension of local feature and combines the PR vectors together to get an MPR vector. The computational complexity of PR is high when the local feature is high-dimensional, which leads PR vector to lose discriminative information in real applications. MPR overcomes these limitations. Its computational complexity is low and the MPR vector has discriminative information, even when the local feature is high-dimensional. Using a TerraSAR-X data set and a TerraSAR-X image, AdaBoost based on PR (PR-AdaBoost) and MPR-AdaBoost are compared. The experimental results have shown that MPR-AdaBoost gives comparable results and reduces the computational cost.

**Key words** Multi-dimensional pyramid representation (MPR), AdaBoost, high resolution SAR (Synthetic aperture radar) image, classification

合成孔径雷达 (Synthetic aperture radar, SAR) 以其高分辨率、全天候、全天时、大面积的数据获取能力而成为世界各国普遍重视的对地观测技术. SAR 图像获取及其信息应用的能力, 对我国的自然保护、国土资源利用、自然灾害监测等起着十分重要的作用.

近 20 年来, SAR 高分辨率成像技术极大地促进了空间遥感的发展. SAR 图像数据源的日益丰富, 一方面为该领域的研究和应用提供了可能, 另一方面也为从海量数据中发现和探索不同应用层面的有用信息提出了巨大的挑战. 面对数量巨大的图像数据, 传统的依靠人工对图像进行分类与标注的管理方式因为需要耗费大量的人力资源而变得不可行. 同时底层特征所建模的类别语义和用户高层语义之间常常存在着巨大的差别 (就是所谓的语义鸿沟). 语义鸿沟使得人工专家的参与和解译, 在应用分析过程中几乎必不可少, 如果一个系统可以自动地学习图像的高层语义并分类, 而不是仅仅完成独立像素/区域的分类, 那么 SAR 的应用领域将大大受益. 本文研究的重点就是利用计算机自动地完成高分辨率 SAR 图像的城区场景分类.

现在, 领先的机载 SAR 系统获取的图像的分辨率已经能够达到分米级水平, 星载系统也达到米级水平<sup>[1]</sup>, 随着 SAR 图像分辨率的不断提高, SAR 图

收稿日期 2009-11-12 录用日期 2010-03-29  
Manuscript received November 12, 2009; accepted March 29, 2010

国家高技术研究发展计划 (863 计划) (2007AA12Z155), 国家自然科学基金 (40901207), 测绘遥感信息工程国家重点实验室专项科研经费, 武汉大学博士研究生科研团队创新研究项目 (20082120201000043) 资助

Supported by National High Technology Research and Development Program of China (863 Program) (2007AA12Z155), National Natural Science Foundation of China (40901207), Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS) Special Research Funding, and Doctor's Independent Research Project of Wuhan University (20082120201000043)

1. 武汉大学电子信息学院信号处理实验室 武汉 430079  
1. Signal Processing Laboratory, School of Electrical Information, Wuhan University, Wuhan 430079

像逼近并达到光学图像的水平是 SAR 系统发展的一个主要趋势. 在高分辨率条件下, SAR 图像中的空间信息更加丰富, 地物目标的几何结构和纹理信息更加明显, “更像”光学图像, 然而这同时也带来了新问题: 从 SAR 的物理成像原理推导有用模型和方法的策略变得难以实施; 而且, 以往针对中低分辨率的一系列模型及方法在高分辨率下效果也不理想.

因此, 计算高分辨率 SAR 图像场景特征需要加强考虑场景图像的纹理、统计量分布、几何形状和空间上下文等计算机视觉特征. 随着 SAR 图像分辨率的不断提高, 细节和结构的复杂性与随机性愈加明显, 原来具有“特征一致性”的简单类别可能会变成“特征不一致”的复杂类别, 从而使基于底层全局特征建模的统计分类器的性能下降. 在计算机视觉和模式识别领域, 流行的趋势是先使用检测算子找到图像的关键区域, 然后在关键区域上计算局部特征, 并通过描述模型对图像进行描述, 最后进行分类识别<sup>[2-4]</sup>. 受这种处理方法的启发, 我们在高分辨率 SAR 图像上使用基于底层局部特征的统计分类器.

将图像的各部分分别用一个局部特征矢量表达, 图像的所有特征矢量构成一个特征矢量集合, 称为“特征袋”, 我们用这个特征矢量集合来表达该图像. 这个方法不仅在光学图像处理中是目前流行的方法, 而且在高分辨率遥感图像中也同样适用<sup>[5]</sup>. 但是, 很多传统的机器学习方法不能直接利用局部特征表达得到的特征矢量集合, 因为这些机器学习方法默认输入都是特征矢量. 为了解决这个问题, 人们通常使用 Bag-of-features (BOF)<sup>[2]</sup>. BOF 事先通过对特征矢量集合聚类(如使用  $k$ -means) 建立特征原型库, 然后用特征原型在图像中出现的概率构成的矢量表达图像. 如何自动确定最佳的聚类中心数是 BOF 有待解决的问题.

金字塔表达 (Pyramid representation, PR) 算法<sup>[4]</sup>, 通过直接划分特征空间将特征矢量集合聚类, 回避了确定最佳聚类中心数的难题, 并且金字塔表达对特征空间由粗到细的多层划分(每次划分对应于金字塔的一层), 较细致地刻画了特征矢量在特征空间的分布. 但是金字塔表达存在两个缺点: 1) 金字塔表达矢量长度(即维数)与局部特征维数呈指数关系, 与金字塔的层数呈线性关系. 在使用高维局部特征情况下, 其表达效率不可接受; 2) 由于高维局部特征矢量往往只是聚集在特征空间的某个小角落<sup>[6]</sup>, 有限层(计算效率的限制)的金字塔并不能将这个“角落”精细划分, 导致不同类别输出的金字塔表达矢量在后续分类应用中缺乏区分力.

为了克服以上问题, 我们提出了一种新的表达算法, 叫做多维金字塔表达. 多维金字塔表达将特征矢量的各维独立看待, 并在每一维计算金字塔表达

矢量, 最后将各维的金字塔表达矢量连接起来. 与金字塔表达最大的不同是, 当输入是  $n$  ( $n > 1$ ) 维局部特征矢量集合时, 金字塔表达计算的是一个  $n$  维金字塔表达矢量, 而多维金字塔的表达是计算  $n$  个一维金字塔表达矢量(当  $n = 1$  时, 两种算法等价). 多维金字塔表达默认特征矢量各维间是独立的. 因此, 有必要事先对原始局部特征矢量进行处理(例如利用主成分分析 (Principle component analysis, PCA) 或者流型方法), 使得处理后的矢量不仅各维间独立而且保留原始局部特征各维间的有用关系. 多维金字塔表达矢量的长度(即维数)与特征的维数呈线性关系. 这使得其可以对特征矢量的每一维进行非常细致的划分而不必担心效率问题, 进而提高不同类别的多维金字塔表达矢量的区分力.

由于用局部特征描述图像, 一些不相关信息很可能对最终的表达造成干扰. 我们使用 AdaBoost 从多维金字塔表达矢量中选择少量具有区分力的维用于高分辨率 SAR 图像的城区场景分类.

我们分别在一个高分辨率 TerraSAR-X 图像库和一张大幅 TerraSAR-X 图像上比较了基于金字塔表达的 AdaBoost 和基于多维金字塔表达的 AdaBoost 的性能. 实验结果均显示, 后者的分类效率远远高于前者而且保持较好的分类性能.

## 1 多维金字塔表达算法

这部分分两个小节: 第 1.1 节简要说明相关工作, 第 1.2 节详细说明多维金字塔表达算法.

### 1.1 相关工作

最近出现了很多关于高分辨率 SAR 图像的研究成果.

Stasolla 等<sup>[7]</sup> 提出了在高分辨率 SAR 图像中提取建筑物的方法. 这个方法的思想是首先检测独立的感兴趣区域(如建筑), 然后把它们聚类成复杂的区块结构. 他同时指出在高分辨率 SAR 图像中居民区表现出不均匀性, 即同一个物体的不同部分可能表现出不同的散射特性, 因此基于像素的分析方法不合适, 而基于分割的方法更有效. Soergel 等<sup>[1]</sup> 利用分辨率达到分米级的高分辨率 SAR 图像重建人工建筑物. Greco 等<sup>[8]</sup> 对不同的植被覆盖区的高分辨率 SAR 图像建模. Bennett 等<sup>[9]</sup> 在高分辨率 SAR 图像上提取建筑物及其阴影, 利用获得的散射信息和阴影信息来推测建筑物的高度.

针对高分辨率 SAR 图像的城区场景分类算法大多是利用高分辨率 SAR 图像和其他数据的融合实现的<sup>[10-12]</sup>. 单纯利用高分辨率 SAR 图像(单极化非干涉)完成这一任务的算法并不多见, 具有代表性的如 Tison 等<sup>[13]</sup> 提出的利用 Fisher 分布对场景

中各个成分的分布建模并利用 Markov 场分类的方法. 其思想秉承传统的对均匀成分建立统计模型的思想. 与基于其他分布 (如 K, Weibull, Log-normal, Nakagami-Rice, Alpha stable 等) 的分类算法的思路相同. 但是, 这个方法不能用来处理我们所要解决的问题, 因为我们要处理的类别是复杂的场景, 每个场景包含了多种均匀成分.

下面将详细说明本文提出的多维金字塔表达算法的思想.

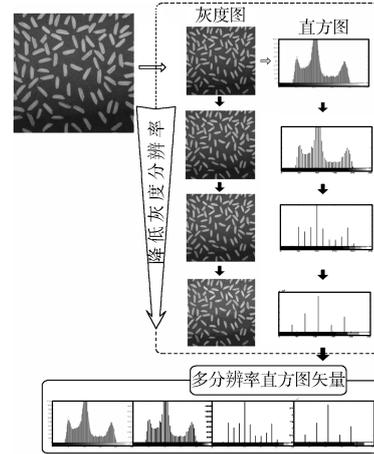
### 1.2 多维金字塔表达

图 1 具体展现了多分辨率直方图、金字塔表达和多维金字塔表达三者的计算方法以及输出矢量的形式. 它们之间的关系是: 多分辨率直方图<sup>[14]</sup> 在灰度特征空间的多分辨率下计算直方图; 金字塔表达是多分辨率直方图在多维特征空间上的扩展; 多维金字塔表达是由每一特征维的金字塔表达连接起来构成的; 金字塔表达和多维金字塔表达的目的都是将特征矢量集合转化成一个矢量.

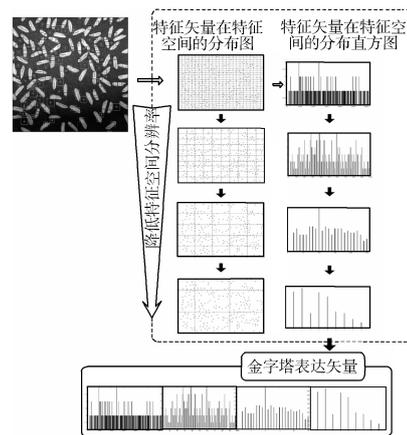
图 1 (a) 描述了多分辨率直方图的思想. 多分辨率直方图算法思想是: 逐渐降低灰度图像的灰度分辨率, 同时统计每次分辨率降低后图像的灰度直方图, 将得到的灰度直方图连接起来构成一个矢量, 称为多分辨率直方图矢量. 文献 [14] 是对这一算法的详细描述.

图 1 (b) 描述了金字塔表达算法思想: 首先在图像上提取兴趣点, 每个兴趣点用一个特征矢量描述. 将图像上计算出的特征矢量映射到特征空间中. 此时, 每个特征矢量对应于特征空间的一个点. 对特征空间划分网格, 在图 1 (b) 中自上而下共划分了 4 次网格, 网格逐渐变大, 即特征空间的分辨率逐渐变小, 同时在图 1 (b) 的右边统计了在特征空间每种分辨率下特征矢量的分布直方图 (此分布直方图的 bin 对应于该分辨率下特征空间的网格, 网格内特征矢量的数目是 bin 对应的频数). 将得到的分布直方图连接起来构成一个矢量, 称为金字塔表达矢量. 这里“金字塔”是对特征空间分辨率不断降低这一处理的形象描述.

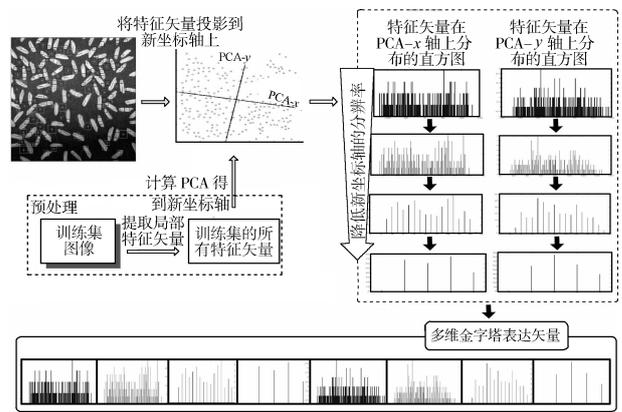
图 1 (c) 描述了多维金字塔表达算法思想: 首先处理局部特征矢量. 图 1 (c) 所描述的是使用 PCA 处理特征矢量后, 将特征矢量映射到新的坐标系下, 即以 PCA-x 为横轴, PCA-y 为纵轴的新的坐标系. 在新坐标轴上分别统计特征矢量的分布直方图并逐渐降低分布直方图的分辨率 (如图 1 (c) 中分辨率自上而下逐渐降低). 最后在各轴上得到多个分辨率下的直方图的连接矢量, 再将所有轴的矢量连接起来构成多维金字塔表达矢量.



(a) 多分辨率直方图算法示意图  
(a) Multi-resolution histogram



(b) 金字塔表达算法示意图  
(b) Pyramid representation



(c) 多维金字塔表达算法示意图  
(c) Multi-dimensional pyramid representation

图 1 多分辨率直方图、金字塔表达和多维金字塔表达算法示意图

Fig. 1 Overview of multi-resolution histogram, pyramid representation, and multi-dimensional pyramid representation

虽然从理论上讲, 在使用多维金字塔表达时, 有必要事先对原始局部特征矢量进行处理 (例如利用 PCA 或者流型方法), 使得处理后的矢量不仅各维间独立而且保留原始各维间的有用关系<sup>[15]</sup>, 但是, 在实验中我们发现即使不做任何处理, 多维金字塔表达仍然能够获得较好的性能。

下面给出多维金字塔表达的数学描述。

设  $I$  表示一张图像, 使用兴趣点检测算法从  $I$  中检测到  $m$  个显著区域  $\{p_1, p_2, \dots, p_m\}$ , 在每个区域  $p_i$  中计算  $D$  维局部特征矢量  $\mathbf{x}_i \in F \subseteq \mathbf{R}$ , 图像  $I$  用特征矢量集合  $X$  描述:

$$X = \{\mathbf{x}_1, \dots, \mathbf{x}_m\} \quad (1)$$

使用 PCA (保留各维之间的线性关系) 或流型 (保留各维之间的非线性关系) 处理  $X$ , 设得到的新特征矢量集合为  $\tilde{X} = \{\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_m\}$ , 其中,  $\tilde{\mathbf{x}}_i \in \tilde{F} \subseteq \mathbf{R}^d$ ,  $d \leq D$ . 处理后, 新特征矢量各维被独立看待。在每一维上分别计算金字塔表达矢量。

下面以第  $i$  维特征为例, 说明在一维特征空间上金字塔表达矢量  $\Psi$  是如何计算的。

首先, 计算  $\tilde{X}$  的第  $i$  维特征 (用  $\tilde{X}(i)$  表示) 的分布直方图  $\mathbf{H}_0(\tilde{X}(i)) = [h_{0, \min(\tilde{X}(i))}, h_{0, \min(\tilde{X}(i))+1}, \dots, h_{0, \max(\tilde{X}(i))}]$ , 设  $\min(\tilde{X}(i))$  和  $\max(\tilde{X}(i))$  分别为  $\tilde{X}(i)$  经过缩放到某个精度, 然后取整得到的最小值和最大值,  $h_{0,j}$  表示  $j$  在  $\tilde{X}(i)$  中出现的概率。

其次, 将  $\mathbf{H}_0(\tilde{X}(i))$  的分辨率降为原来的  $1/2$ , 得到的分布直方图为  $\mathbf{H}_1(\tilde{X}(i))$ . 再将  $\mathbf{H}_1(\tilde{X}(i))$  的分辨率下降  $1/2$ , 得到分布直方图  $\mathbf{H}_2(\tilde{X}(i))$ . 以此类推, 共可以进行  $\lceil \log_2(\max(\tilde{X}(i)) - \min(\tilde{X}(i))) \rceil$  次直方图分辨率下降  $1/2$  的操作。

最后, 将直方图连接起来, 得到一维特征空间上金字塔表达矢量  $\Psi(\tilde{X}(i))$

$$\Psi(\tilde{X}(i)) = [\mathbf{H}_0(\tilde{X}(i)), \dots, \mathbf{H}_{L-1}(\tilde{X}(i))] \quad (2)$$

$L$  为金字塔的层数

$$L = \lceil \log_2(\max(\tilde{X}(i)) - \min(\tilde{X}(i))) \rceil + 1 \quad (3)$$

$\mathbf{H}_j(\tilde{X}(i))$  的维数为

$$r_j(i) = \frac{\max(\tilde{X}(i)) - \min(\tilde{X}(i))}{2^j} \quad (4)$$

$I$  的多维金字塔表达矢量为

$$\mathbf{MP}(\tilde{X}) = [\Psi(\tilde{X}(1)), \Psi(\tilde{X}(2)), \dots, \Psi(\tilde{X}(d))] \quad (5)$$

多维金字塔表达矢量的长度 (维数) 计算如下:

$$R = \sum_{i=1}^d \sum_{j=0}^{L-1} r_j(i) \quad (6)$$

多维金字塔表达矢量的长度 (即维数) 与局部特征的维数  $d$  呈线性关系. 这使得其可以对每一特征维进行非常细致的划分而不必担心效率问题, 进而提高不同类别的多维金字塔表达矢量的区分力。

## 2 基于多维金字塔表达的 AdaBoost

本节说明如何将多维金字塔表达和 AdaBoost 结合构建分类器。

由于用局部特征描述图像, 一些不相关信息很可能对最终的表达造成干扰<sup>[16]</sup>. 要克服干扰的影响, 主要有两种方式: 一种是使用特征选择<sup>[17]</sup> 或是 Boosting<sup>[18]</sup> 选择有区分力的特征维用于构建出色的分类器; 另一种方法就是设计具有高区分力的核<sup>[19-21]</sup>. 本文采用第一种方法 (关于第二种方法我们将在后续的工作中实现), 并使用 AdaBoost 算法<sup>[22]</sup> 自动选择有区分力的特征维. AdaBoost 是基于 Boosting 思想的众多改进算法之一, 近年来在遥感图像的物分类中得到广泛的应用<sup>[23-24]</sup>.

下面具体说明如何利用 AdaBoost 得到强分类器  $F_y(\mathbf{MP}(\tilde{X}))$ . 假设数据库中有  $K$  类, 共  $N$  幅图像, 每幅图像用多维金字塔表达算法表达, 共得到  $N$  个多维金字塔表达矢量, 表示为:  $\mathbf{MP}(\tilde{X}) = \{\mathbf{MP}(\tilde{X}_1), \dots, \mathbf{MP}(\tilde{X}_N)\}$ , 其中每个矢量都有一个类别标号  $y \in Y = \{1, \dots, K\}$ , 利用 AdaBoost 训练强分类器  $F_y(\mathbf{MP}(\tilde{X}_i)), y = 1, \dots, K$ .

强分类器计算如下<sup>[23-24]</sup>:

$$F_y(\mathbf{MP}(\tilde{X}_i)) = \begin{cases} \sum_{t=1}^T \log \frac{1}{\beta_t}, & \text{若 } h_t(\mathbf{MP}(\tilde{X}_i)) = y \\ 0, & \text{否则} \end{cases} \quad (7)$$

其中,  $T$  是迭代次数,  $\beta_t = \varepsilon_t / (1 - \varepsilon_t)$ ,  $\varepsilon_t = \sum_{i=1:N; h_t(\mathbf{MP}(\tilde{X}_i)) \neq y} \omega_t(i)$ .  $\omega_t(i)$  是第  $t$  次迭代后第  $i$  个训练矢量的权重.  $h_t$  是第  $t$  次迭代选择的最佳弱分类器 (我们使用决定树作为弱分类器). 一般来说, 迭代次数  $T$  远远小于多维金字塔表达矢量的维数  $R$ .

## 3 实验

本节分别用两个实验测试并比较基于多分辨率直方图的 AdaBoost 算法、基于金字塔表达的 AdaBoost 算法和基于多维金字塔表达的 AdaBoost 算法. 第一个实验在一个包含 5 类地物的 TerraSAR-X 数据集上, 分别使用两种特征测试特征维数对算法性能的影响; 第二个实验在一张大幅的 TerraSAR-X 图像上测试三种算法分别进行城区场景分类的性能。

### 3.1 在 TerraSAR-X 数据集上的实验

这个数据集包含 511 张图片分别属于: 居民区 (100 张)、工业区 (113 张)、草地 (102 张)、森林 (100 张) 和裸地 (96 张). 这些图像都来自 TerraSAR-X 拍摄的中国广州的图像 (大小为 20 000 像素  $\times$  20 000 像素, 分辨率为 1.25 米). 图像集中每张图像的大小为 128 像素  $\times$  128 像素. 根据相同区域的高分辨率光学图像、对该地域的先验知识以及已知的 SAR 图像对各种地物表现的特点, 手工标注了各图像所属的类别. 从每类随机选择 50 张图像, 将这 250 张图像作为训练集, 剩下的所有图像作为测试集, 从而将数据集分成两组, 如此独立进行 10 次分组并测试, 最后得到的分类准确率是 10 次运算结果的平均.

将每张图像分成 16 像素  $\times$  16 像素的图像子块, 子块间重叠 50%. 分别从每个子块中提取两种局部特征: Gabor 特征 (48 维) 和灰度直方图特征 (32 维). 我们的实验分别测试了局部特征和特征维数对算法的影响. 使用 PCA 将局部特征降到不同维, 然后在降维后的特征矢量集合上分别构建金字塔表达矢量和多维金字塔表达矢量. 再将这两种表达矢量分别与 AdaBoost 结合, 在训练集上学习出强分类器后在测试集上分类. 所有实验中 AdaBoost 的迭代次数设为 100, 给出的分类准确率是 100 次迭代的最高值. 需要声明的是: 1) 由于金字塔表达矢量是稀疏的, 实验中给出的金字塔表达矢量的维数都是将始终为零的维去除后的维数; 2) 实验中所给出的金字塔的层是以分辨率最低的一层为第一层, 分辨率次低的为第二层, 以此类推, 分辨率最高的一层为最后一层 (即刚好和式 (2) 的顺序相反).

#### 3.1.1 使用 Gabor 特征实验

使用 3 种尺度, 8 个方向 (分别计算了均值和方差) 共 48 维的 Gabor 特征提取 16 像素  $\times$  16 像素的图像子块内的纹理特征. 再使用 PCA 分别按照保留原始特征能量的 90%、80%、70%、60% 和 50% 对 48 维 Gabor 特征降维, 降维后的维数用  $Dim$  表示. 在降维后的特征矢量集合上分别构建金字塔表达矢量和多维金字塔表达矢量, 得到的矢量的维数分别用  $PR_{dim}$  和  $MPR_{dim}$  表示 (参数  $Dim$ 、 $PR_{dim}$  和  $MPR_{dim}$  的定义下同). 表 1 给出了使用 Gabor 特征时,  $Dim$ 、 $PR_{dim}$ 、 $MPR_{dim}$  和分类准确率的对应值. 对于金字塔表达和多维金字塔表达, 表 1 所示是使用前 4 层计算的结果 (由于多分辨率直方图是定义在灰度空间的, 所以表 1 没有基于多分辨率直方图算法的实验结果).

观察表 1 发现:

1) 无论降维与否, 基于多维金字塔表达的 Ad-

aBoost 的性能都高于基于金字塔表达的 AdaBoost 的性能. 同时发现, 多维金字塔表达在不降维时得到最佳的性能 (如表 1 中黑体所示), 这说明本文的算法不受降维与否的限制.

2) 多维金字塔表达矢量的维数远远小于金字塔表达矢量的维数, 所以基于多维金字塔表达的 AdaBoost 在获得高性能的同时也获得了比基于金字塔表达的 AdaBoost 高得多的计算效率.

表 1 使用 Gabor 特征时两种算法性能比较

Table 1 Comparison of classification performances using Gabor feature

保留能量	$Dim$	$PR_{dim}$	分类准确率	$MPR_{dim}$	分类准确率
100%	48	66 776	48.70%	454	<b>75.42%</b>
90%	20	74 240	51.94%	262	69.00%
80%	13	55 902	58.22%	203	69.49%
70%	10	41 811	61.29%	169	68.55%
60%	7	15 255	<b>64.11%</b>	121	69.78%
50%	6	8 147	63.80%	107	70.07%

#### 3.1.2 使用灰度直方图特征的实验

本文使用 32 维的灰度直方图特征 (对原始直方图均匀量化得到) 提取 16 像素  $\times$  16 像素的图像子块内的特征. 再使用 PCA 分别按照保留原始特征能量的 90%、80%、70% 和 60% 对 32 维特征降维. 表 2 列出了各参数的对应值. 表 2 所示是使用前 5 层计算的结果. 同时, 表 2 也给出了基于多分辨率直方图的 AdaBoost 的分类性能. 这里的多分辨率直方图是对图像的全局表达. 多分辨率直方图矢量的维数用  $MH_{dim}$  表示.

表 2 使用灰度直方图特征时三种算法性能比较

Table 2 Comparison of classification performances using gray histogram feature

保留能量	$Dim$	$PR_{dim}$	分类准确率	$MPR_{dim}$	分类准确率	$MH_{dim}$	分类准确率
100%	32	64 856	86.74%	377	90.88%	63	<b>88.20%</b>
90%	4	1 261	<b>91.14%</b>	112	<b>91.20%</b>	7	82.59%
80%	3	617	90.82%	92	90.67%	6	79.49%
70%	2	186	90.40%	64	90.22%	3	75.80%
60%	1	33	81.97%	36	83.40%	1	71.26%

观察表 2 发现:

1) 对于多分辨率直方图, 表达矢量的区分力随着降维数的增加而下降. 而且基于全局表达的 AdaBoost 的分类性能低于基于局部表达的 AdaBoost 的分类性能.

2) 基于多维金字塔表达的 AdaBoost 分类性能和基于金字塔表达的 AdaBoost 的分类性能只有微

小的差别,但是多维金字塔表达矢量的维数只是金字塔表达矢量的 1/10.

3) 对于高维局部特征 (32 维), 多维金字塔表达矢量也具有好的区分力.

4) 使用灰度直方图特征得到的分类性能要明显好于使用 Gabor 特征时的性能.

### 3.1.3 小结

分析 TerraSAR-X 数据库上的实验结果, 发现:

1) 当降维后的特征是高维特征时, 基于多维金字塔表达的 AdaBoost 的性能比基于金字塔表达的 AdaBoost 的性能好、效率高 (如表 1 所示).

2) 当降维后的特征是低维特征时, 基于多维金字塔表达的 AdaBoost 的性能与基于金字塔表达的 AdaBoost 的性能相仿 (如表 2 所示), 但是计算效

率比后者高.

3) 基于多维金字塔表达的 AdaBoost 受降维数的影响较小, 而基于金字塔表达的 AdaBoost 受降维数的影响较大, 而且, 通过降维通常可以提高基于金字塔表达的 AdaBoost 的分类性能. 对于基于金字塔表达的 AdaBoost 而言, 确定降维数是关键, 用户应在保证降维后信息不过多丢失的前提下, 使降维后的维数为低维 (即  $d \leq 10$ ), 才能保证金字塔表达矢量具有较好的区分力.

### 3.2 对 TerraSAR-X 图像城区场景分类的实验

这部分将用基于多分辨率直方图的 AdaBoost、基于金字塔表达的 AdaBoost 和基于多维金字塔表达的 AdaBoost 分别对一幅 TerraSAR-X 城区场景图像 (如图 2(a) 所示) 分类.



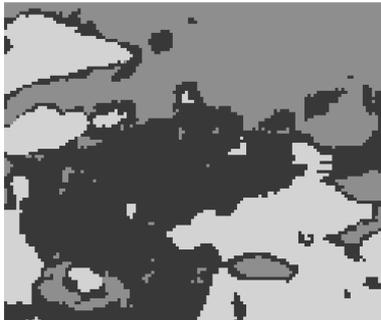
(a) SAR 图像

(a) SAR image



(b) 手工标注图 (其中第一类用中等灰色表示; 第二类用浅灰色表示; 第三类用深灰色表示)

(b) Ground truth (Class 1: middle gray; Class 2: light gray; Class 3: deep gray)



(c) 使用基于多分辨率直方图的 AdaBoost 的分类结果

(c) Classification result of AdaBoost based on multi-resolution histogram



(d) 使用基于金字塔表达的 AdaBoost 的分类结果

(d) Classification result of AdaBoost based on pyramid representation



(e) 使用基于多维金字塔表达的 AdaBoost 的分类结果

(e) Classification result of AdaBoost based on multi-dimensional pyramid representation

图 2 SAR 原图、标注图和分类结果

Fig. 2 SAR image, ground truth, and classification results

这张图像是 TerraSAR-X 拍摄的加拿大的多伦多市某处, 我们只截取了原图的一部分. 这张图像的大小是 1680 像素  $\times$  1984 像素, 分辨率为 1.25 米. 我们根据相同区域的高分辨率光学图像以及该区域的地图, 将此区域大致分成三类: 第一类包括草地、裸地、混合植被和墓地; 第二类主要包括一些人工建筑的公共场所, 如商场、学校和医院; 第三类主要是住宅区. 各类不再是“特征一致”的类别而是“特征不一致”的复杂类别, 这对三种算法是一种挑战.

由于 AdaBoost 属于监督分类, 因此训练样本的选择非常重要. 在选择训练样本的时候, 我们在原始大图像 (图 2(a) 只是其中一部分) 上除图 2(a) 以外的区域, 根据已知的 SAR 图像对各种地物表现的特点以及参照 Google Earth 的光学图像选择训练样本. 每一类 100 个训练样本, 每个样本是 128 像素  $\times$  128 像素的图像.

图 2(b) 是对图 2(a) 手工标注的结果. 这个结果同样是根据已知的 SAR 图像对各种地物表现的特点、该区域的先验知识以及参照 Google Earth 的光学图像手工标注得到的.

将图 2(a) 分解成 128 像素  $\times$  128 像素的子图像, 子图像间重叠 87.5%. 用均匀网格将每张子图像分解成 16 像素  $\times$  16 像素的子块, 块间重叠 50%. 根据第 3.1 节的实验结果, 我们使用每个子块内均匀量化的 32 维灰度直方图, 并使用 PCA 降到 4 维, 作为子块的局部特征并计算 5 层金字塔的金字塔表达矢量和多维金字塔表达矢量, 得到的金字塔表达矢量的维数为 664, 多维金字塔表达矢量的维数为 85. 同时使用 PCA 降到 4 维的灰度特征作为全局特征, 构建多分辨率直方图, 得到的多分辨率直方图矢量的维数是 6. 图 2(c)~2(e) 分别给出了基于多分辨率直方图的 AdaBoost, 基于金字塔表达的 AdaBoost 和基于多维金字塔表达的 AdaBoost 分类的结果. 三种算法中 AdaBoost 都迭代 100 次, 取在训练集上分类误差最小的强分类器在图 2(a) 上分类. 我们使用错分率来评价三种算法的性能. 基于多分辨率直方图的 AdaBoost 的错分率为 20.41%, 基于金字塔表达的 AdaBoost 的错分率为 13.66%, 基于多维金字塔表达的 AdaBoost 的错分率为 14.73%. 可见, 基于多分辨率直方图的 AdaBoost 的分类准确率最低, 即与全局特征表达相比, 局部特征表达更适合描述高分辨率 SAR 图像. 基于多维金字塔表达的 AdaBoost 的分类准确率比基于金字塔表达的 AdaBoost 的分类准确率降低了 1%. 但是, 由于多维金字塔表达矢量的维数只是金字塔表达矢量维数的 1/8, 而 AdaBoost 的计算时间与表达矢量的维数呈线性关系, 所以与基于金字塔表达的 AdaBoost 相比, 基于多维金字塔表达的 AdaBoost 将计算效

率提高了 8 倍.

## 4 总结与展望

本文提出了多维金字塔表达算法, 并与 AdaBoost 结合, 完成高分辨率 SAR 图像的城区场景分类任务. 多维金字塔表达能够将特征矢量集合转化成一个矢量, 因此它可以被广泛地应用在基于局部特征表达的应用框架中. 多维金字塔表达矢量的长度 (即维数) 与局部特征矢量的维数呈线性关系. 这使得其可以对每一特征维进行非常细致的划分而不必担心效率问题, 进而提高不同类别的多维金字塔表达矢量的区分力. 我们分别在一个 TerraSAR-X 数据集和一张大幅 TerraSAR-X 图像上比较了基于金字塔表达的 AdaBoost 和基于多维金字塔表达的 AdaBoost 的分类性能, 实验结果显示后者的计算效率显著高于前者同时保证了良好的分类性能.

我们在今后的工作中将继续研究如何利用基于多维金字塔表达的核方法进行高分辨率 SAR 图像的城区场景分类.

## References

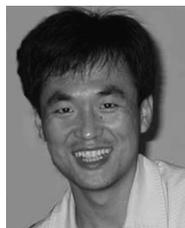
- 1 Soergel U, Thoennessen U, Brenner A, Stilla U. High-resolution SAR data: new opportunities and challenges for the analysis of urban areas. *IEEE Proceedings — Radar Sonar and Navigation*, 2006, **153**(3): 294–300
- 2 Li F F, Perona P. A Bayesian hierarchical model for learning natural scene categories. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE, 2005. 524–531
- 3 Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE, 2006. 2169–2178
- 4 Grauman K L. Matching Sets of Features for Efficient Retrieval and Recognition [Ph. D. dissertation], Massachusetts Institute of Technology Computer Science and Artificial Intelligence Laboratory, USA, 2007
- 5 Bordes J B, Maitre H. Semantic annotation of satellite images. In: *Proceedings of the 5th International Conference on Machine Learning and Data Mining in Pattern Recognition*. Leipzig, Germany: Springer, 2007. 120–133
- 6 Verleysen M, Francois D. The curse of dimensionality in data mining and time series prediction. In: *Proceedings of the 8th International Workshop on Artificial Neural Networks*. Barcelona, Spain: Springer, 2005. 758–770
- 7 Stasolla M, Gamba P. Spatial indexes for the extraction of formal and informal human settlements from high-resolution SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2008, **1**(2): 98–106
- 8 Greco M S, Gini F. Statistical analysis of high-resolution SAR ground clutter data. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, **45**(3): 566–575

- 9 Bennett A J, Blacknell D. The extraction of building dimensions from high resolution SAR imagery. In: Proceedings of the International Radar Conference. Washington D. C., USA: IEEE, 2003. 182–187
- 10 Klonus S, Ehlers M. Pansharpening with TerraSAR-X and optical data. In: Proceedings of the 3rd TerraSAR-X Science Team Meeting. Darmstadt, Germany: German Aerospace Center, 2008. 25–26
- 11 Tison C, Tupin F, Maitre H. A fusion scheme for joint retrieval of urban height map and classification from high-resolution interferometric SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, **45**(2): 496–505
- 12 Gabrielle L, Marine C, Florence T, Celine T, Guillaume O. Combining SAR and optical features in an SVM classifier for man-made structures detection. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Cape Town, Africa: IEEE, 2009. 873–876
- 13 Tison C, Nicolas J M, Tupin F, Maitre H. A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, **42**(10): 2046–2057
- 14 Hadjidemetriou E, Grossberg M D, Nayar S K. Multiresolution histograms and their use for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(7): 831–847
- 15 Shlens J. A tutorial on principal component analysis [Online], available: <http://www.cs.cmu.edu/~elaw/papers/pca.pdf>, March 30, 2010
- 16 Zhang J, Marszaek M, Lazebnik S, Schmid C. Local features and kernels for classification of texture and object categories: a comprehensive study. *International Journal of Computer Vision*, 2007, **73**(2): 213–238
- 17 Dorko G, Schmid C. Object Class Recognition Using Discriminative Local Features, Technical Report RR-5497, Institute National of Research on Information ET on Automation, France, 2005
- 18 Opelt A, Fussenegger M, Pinz A, Auer P. Weak hypotheses and boosting for generic object detection and recognition. In: Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer, 2004. 71–84
- 19 Grauman K, Darrell T. Efficient image matching with distributions of local invariant features. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2005. 627–634
- 20 Lyu S. Mercer kernels for object recognition with local features. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2005. 223–229
- 21 Wallraven C, Caputo B, Graf A. Recognition with local features: the kernel recipe. In: Proceedings of the 9th IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE, 2003. 257–264
- 22 Meir R, Ratsch G. An introduction to boosting and leveraging. *Advanced Lectures on Machine Learning*. New York: Springer, 2003. 118–183
- 23 Friedl M A, Brodley C E, Strahler A H. Maximizing the land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, **37**(2): 969–977
- 24 McIver D K, Friedl M A. Estimating pixel-scale land cover classification confidence using nonparametric machine learning methods. *IEEE Transactions on Geoscience and Remote Sensing*, 2001, **39**(9): 1959–1968



殷 慧 武汉大学电子信息学院博士研究生。主要研究方向为图像处理和模式识别。

E-mail: kittyxuanxuan.yinhui@gmail.com  
(YIN Hui Ph.D. candidate at the Signal Processing Laboratory, Electronic Information School, Wuhan University. Her research interest covers image processing and pattern recognition.)



曹永锋 武汉大学电子信息学院讲师。主要研究方向为图像处理和模式识别。本文通信作者。

E-mail: yongfengcao.cyf@gmail.com  
(CAO Yong-Feng Lecturer at the Electronic Information School, Wuhan University. His research interest covers image processing and pattern recognition. Corresponding author of this paper.)



孙 洪 武汉大学电子信息学院教授。主要研究方向为图像分析、统计和通信信号处理。E-mail: hongsun@whu.edu.cn  
(SUN Hong Professor at the Electronic Information School, Wuhan University. Her research interest covers image analysis, statistical and communication signal processing.)