

基于模糊加权 SVM 的 SAR 图像水体自动检测

程明跃¹, 叶 勤^{1,2}, 张绍明¹, 陈映鹰², 李 伟¹

(1. 同济大学测量与国土信息工程系, 上海 200092; 2. 同济大学遥感与空间信息技术研究中心, 上海 200092)

摘要: 提出一种 SAR 图像水体自动检测方法。该方法采用树型小波对 SAR 图像进行分解, 提取样本图像与待检测图像的中频信息, 并进行纹理分析, 利用模糊加权支持向量机对样本图像的纹理进行训练, 从而获得判别水体的决策函数, 通过该决策函数能够检测出图像的水体区域。该方法结合了图像的灰度与纹理信息, 减少了 SAR 图像中的噪声影响, 能够适用于大幅面范围的 SAR 图像水体自动检测。

关键词: SAR 图像; 水体检测; 纹理分析; 树型小波; 模糊加权支持向量机

Water Automatic Detection from SAR Image Based on Fuzzy Weighted SVM

CHENG Ming-yue¹, YE Qin^{1,2}, ZHANG Shao-ming¹, CHEN Ying-ying², LI Wei¹

(1. Dept. of Surveing and Geoinformatics, Tongji University, Shanghai 200092;

2. Research Center of Remote Sensing & Saptial Imformation Technology, Tongji University, Shanghai 200092)

【Abstract】 A method of water automatic detection for SAR images is presented, which uses tree wavelet to decompose the SAR images, obtains the medium-frequency imformation from the sample image and the test image, and analyzes the texture. The texture information of the sample image is trained by using Fuzzy Weighted Support Vector Machine(FW-SVM), and the water decision function is got, through which the water region can be detected. This method combines the gray and texture information of the SAR image, reduces the noise effect, and is suitable for the water automatic detection of large area SAR image.

【Key words】 SAR image; water detection; texture analysis; tree wavelet; Fuzzy Weighted Support Vector Machine(FW-SVM)

1 概述

利用合成孔径雷达(Synthetic Aperture Radar, SAR)图像进行水体目标检测是 SAR 技术应用的一个主要方向。传统 SAR 图像水体目标检测方法往往是基于灰度的单一阈值分割方法, 这种方法对某些低噪声、图幅较小的 SAR 图像的水体检测有较好效果, 然而由于雷达反射回波信号的影响, SAR 图像中通常都存在大量乘性噪声, 因此单纯地采用基于灰度的方法会由于噪声而影响水体目标检测的效果。另外, 对于大范围(5 000×5 000 pixel)的 SAR 图像, 图像中的灰度分布十分复杂, 采用阈值分割方法的错分概率很大。为克服这些困难, 可以对 SAR 图像中的纹理等非灰度信息加以研究, 并运用机器学习方法分类, 优化 SAR 图像中水体目标检测效果。

前人针对纹理和机器学习方法分别进行了大量研究, 如 Haralick 于 1973 年提出的灰度共生矩阵法, 该方法计算量巨大, 若进行灰度量化的会导致损失大量纹理信息, 难以在大幅图像纹理分割中的使用。为提高纹理分析效率, 并尽量保留原图像信息, 本文采用树型小波纹理分解。目前应用最广的机器学习方法是神经网络算法, 能够高速寻找复杂问题的优化解, 但神经网络具有很高的风险性, 容易出现局部极小点和泛化能力差等问题。本文采用具有高抗风险性的模糊加权支持向量机(Fuzzy Weighted Support Vector Machine, FW-SVM)方法。

2 理论基础

2.1 树型小波纹理分析

纹理是指由按照某种规则排列的基本元素组成的一个宏观区域, 如果图像中的某个特定区域的局部统计特性和其他

统计特征是一致的、变化缓慢的、或近似周期的, 那么该区域就具有一致的纹理^[1]。水体目标在 SAR 图像中具有较低灰度值、像素分布较相似, 因此, SAR 图像中的水体可以认为具有相同的纹理, 利用纹理分析的分类方法可以实现水体目标检测。近年来, 小波变换法被逐渐应用于纹理分析中, 小波变换具有良好的时频分析能力, 能够描述图像各个尺度上的多方向纹理特征。常用的小波变换算法是 Mallat 提出的塔式算法, 该方法对每一级图像的低频成分进行分解, 得到 LL, HL, LH 和 HH 4 个分量, 构成类似金字塔形状的图像数据结构。而树型小波分解对每个分解层次中的所有分量进行有选择的再次分解, 如图 1 所示, 如果对所有分量都进行分解则称为完全树型小波分解, 否则称为不完全树型小波分解^[2]。

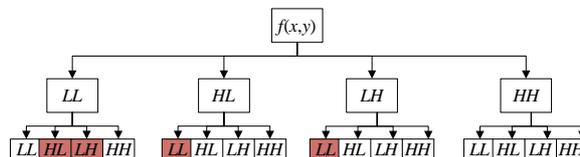


图 1 二维信号的树型小波分解

基金项目: 国家“863”计划基金资助项目“特大城市中心密集区复杂地物群信息的自动提取与定量分析技术”(2007AA12Z178); 开放课题基金资助项目(ES_SBSM_(07)_09)

作者简介: 程明跃(1984 -), 男, 硕士研究生, 主研方向: 数字图像处理, 模式识别; 叶 勤, 副教授; 张绍明, 博士; 陈映鹰, 教授、博士生导师; 李 伟, 硕士

收稿日期: 2008-05-10 **E-mail:** cmyxl@163.com

塔式小波分解方法应用广泛,但是它并不适用于分解纹理图像,现通过实验进行分析。对一张普通的 Cameraman 图像和一张 Brodatz 纹理库中的 D52 图像进行一次小波分解,分别得到它们的 LL , HL , LH 和 HH 4 个分量图,然后利用公式 $E = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N |f(x,y)|$ 计算各分量的能量,实验结果分别如图 2 和表 1 所示。

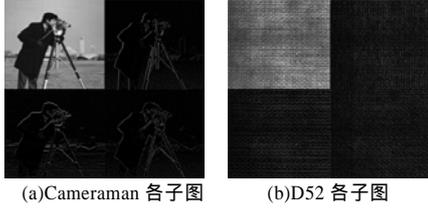


图 2 普通图像与纹理图像小波分解结果

表 1 Cameraman 和 D52 的各子图能量统计

	能量				中频比率(%)
	LL	LH	HL	HH	
D52	134.80	42.34	42.41	10.85	36.78
Cameraman	237.45	13.00	11.75	3.39	9.32

从表 1 可以看出,对于普通图像而言,图像的能量几乎完全集中于图像的低频部分,中高频部分所占能量很小,因此,对普通图像使用塔式小波非常合适。但对于纹理图像,低频部分能量所占比重降低,中频部分能量比重增加,这时如果使用塔式小波,就会因完全抛弃中频信息而降低纹理图像的分类正确率。对于纹理图像,使用树型小波能增加图像信息的利用率,从而提高纹理图像的正确分类概率。本文在分解 SAR 纹理图像时,选择富含能量信息的分量进行再分解,并选取如图 1 所示阴影的 4 个二次分解分量,即可较好地描述 SAR 图像的纹理信息。

2.2 基于模糊加权 SVM 的分类方法

SAR 图像水体目标检测即是图像分为水体与非水体 2 个部分,因此,采用机器学习方法最主要的就是选择一种合适的分类器对水体和非水体的样本进行学习,以训练用于水体分类的决策函数。为克服传统的机器学习方法具有的高风险性问题,文献[3]提出了基于结构风险最小化原理的 SVM 分类算法,其基本思想是通过某种事先选定的非线性映射,将输入向量映射到一个高维特征空间,在该空间构造一个最优分类平面,使之能够尽可能多地将 2 类数据正确分开,并使这 2 类数据距离分类平面最远。

由于 SAR 图像水体数据有限,而非水体目标样本的数量很大,因此采用标准 SVM 会使水体的错分率大于非水体的错分率。但水体类的重要性大于非水体类,对水体进行正确的识别是机器学习的目的所在。因此,必须考虑对标准 SVM 进行改进,即使牺牲整体的错分率也要减小水体样本的错分率。本文采用加权支持向量机来训练决策函数,对水体样本使用大的惩罚系数,提高其在训练中的权重,虽然总体错分率有所上升,但能够降低水体的错分率。

对于可分的样本集 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$, $x_i \in R^n$, $y_i \in \{1, -1\}$, $i = 1, 2, \dots, l$, 分类超平面方程为 $\langle w, x \rangle + b = 0$, 判别函数为 $g(x) = \langle w, \varphi(x) \rangle + b$ 。SVM 的优化问题可描述为

$$\begin{aligned} \min & \frac{1}{2} \langle w, w \rangle + \frac{1}{l} C \sum_{i=1}^l \lambda_i \xi_i \\ \text{s.t.} & \langle w, \varphi(x_i) \rangle + b - 1 + \xi_i \leq 0 \quad i = 1, 2, \dots, l \end{aligned} \quad (1)$$

其中, C 为常数,控制对错分样本的惩罚程度;对于线性不可分的数据,要增加一个松弛项 $\xi_i > 0$; λ_i 为类别权重。

对不同类别进行加权区分不同类别的重要程度。在水体样本中,有的样本基本全为水体区域,有的样本可能存在部分噪声,因此,本文对每个水体样本计算隶属度 μ 。利用隶属度对样本再次进行加权,得到模糊加权 SVM。模糊加权 SVM 的最优化问题描述如下^[4]:

$$\begin{aligned} \min & \frac{1}{2} \langle w, w \rangle + \frac{1}{l} C \sum_{i=1}^l \lambda_i \mu_i \xi_i \\ \text{s.t.} & \langle w, \varphi(x_i) \rangle + b - 1 + \xi_i \leq 0, \quad \xi_i \geq 0 \quad i = 1, 2, \dots, l \end{aligned} \quad (2)$$

其中, μ_i 样本权重,即样本隶属度。

此时,模糊加权 SVM 的对偶问题可描述为

$$\begin{aligned} \max & \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \langle \varphi(x_i), \varphi(x_j) \rangle \\ \text{s.t.} & \sum_{i=1}^l \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \lambda_i \mu_i, \quad i = 1, 2, \dots, l \end{aligned} \quad (3)$$

其决策函数为

$$f(x) = \text{sgn} \left[\sum_{x_i \in SV} y_i \alpha_i^* K(x_i, x) + b \right] \quad (4)$$

其中, $K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle$; $\alpha^* = (\alpha_1, \alpha_2, \dots, \alpha_l)$ 为式(2)的最优解。

3 基于树型小波与 FW-SVM 的自动水体检测

本文方法包括 2 个过程:先获取 SAR 图像纹理信息,使用树型小波对样本和待检测图像进行分解,构成描述图像纹理信息的向量,然后将样本图像的纹理向量装入 SVM 分类器中,通过分类器训练得到决策函数,利用决策函数对待检测图像的纹理向量进行分类判别。其主要流程如图 3 所示。

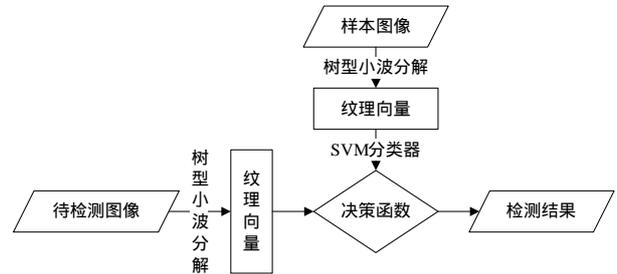


图 3 算法流程

使用树型小波对 SAR 纹理图像进行第 1 次分解,得到纹理图像的 LL , HL , LH 分量,然后分别对这 3 个分量进行第 2 次树型小波分解。保留由 LL 分量分解得到的 HL , LH , 由 LH 分解得到的 LL 以及由 HL 分解得到的 LL 这 4 个二次分解分量,它们保留了原图像的大部分信息,也防止了仅使用低频分量带来的噪声污染。为更有效地描述图像的纹理信息,对于每个分量子图都以一个 $n \times n$ 的模板遍历,用公式 $E = \frac{1}{n \times n} \sum \sum f(x)^2$ 来计算特征信息。这样,对于图像中的每个像素,都有一个 4 维的向量 $V = [E_{LL_1}, E_{LL_2}, E_{LH}, E_{HL}]$ 来描述该像素的纹理信息。为更完备地描述图像的灰度分布信息,再加上图像像素的灰度值就可以组成一个 5 维向量 $V = [E_{LL_1}, E_{LL_2}, E_{LH}, E_{HL}, gray]$, 以描述该像素的信息。将样本图像的 5 维信息向量装入 SVM 分类器中,通过训练生成决策函数。

使用 SVM 分类器将 SAR 图像分类为水体与非水体的过程是先通过 SVM 分类器对水体样本与非水体样本的学习,

产生一个专用于水体分类的决策函数，然后利用这个决策函数对待检测图像每个像素的信息进行判别，决定该像素是水体还是非水体。具体的 SVM 分类步骤如下：

(1)将样本图像的纹理向量装入 SVM 分类器中进行训练，得到决策函数；

(2)用决策函数对待检测图像的纹理向量进行判别，得到水体分类的结果。

4 实验与分析

本文实验所用测试图像为一幅包括水体区域的吴江地区 RadarSat-1 SAR 图像，从测试图像中水体区域与非水体区域中各提取 2 块区域作为样本图像，如图 4 所示。

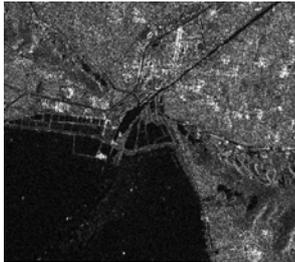


图 4 测试图像

实验 1：使用纹理信息结合灰度信息 $V=[E_{LL}, E_{LLc}, E_{LH}, E_{HL}, gray]$ ，在提取纹理信息时模板为 3×3 ，SVM 分类器核函数类型为高斯径向基核函数，gamma 系数为 1，惩罚系数 $C=50$ 。水体样本数和非水体样本数均为 20 000 个，样本经过 SVM 分类器训练后，总的支持向量数(SV)为 1 035 个，SV 比例为 5.17%，水体 SV 数为 502 个，非水体 SV 数为 533 个，实验结果如图 5(a)所示。

实验 2：仅使用纹理信息 $V=[E_{LL}, E_{LLc}, E_{LH}, E_{HL}]$ ，其他配置同实验 1。样本经过 SVM 分类器训练后，SV 数为 1 809 个，SV 比例为 9.05%，水体的 SV 数为 895 个，非水体 SV 数为 914 个。实验结果如图 5(b)所示。



(a)实验 1 结果



(b)实验 2 结果

图 5 水体检测实验结果

为进一步验证本文中方法的分类正确性，从测试图像中分别取出一块完全为水体区域的图像与一块完全为非水体区域的图像，仍然使用与实验 1 相同的配置，进行分类检测。对于水体区域，分类正确率为 99.40%，对于非水体区域，分类正确率为 99.95%，测试结果如图 6 所示。

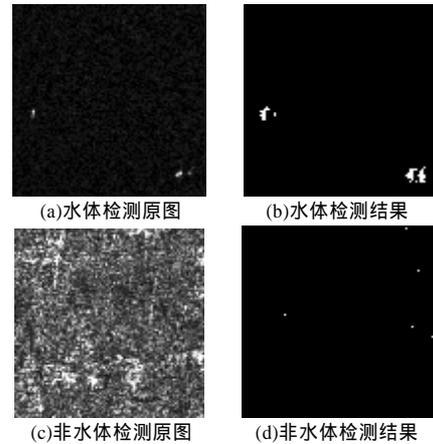


图 6 分类检测结果

根据实验结果可以得出：

(1)本文方法对于 SAR 图像中的水体区域具有很高的正确检测率，基本上检测出了图像中的水体区域，尤其是对于 SAR 图像中的面状水体区域有良好的检测效果；

(2)引入图像灰度信息虽然引入了部分噪声，但是增强了图像纹理信息的描述，提高了水体区域的正确检测概率；

(3)由于非水体区域灰度分布的复杂性，因此图像中非水体存在少量的误检斑点，这很容易通过后续处理消除；

(4)由于采用 3×3 的模板，因此水体区域中高亮噪声周围像素在处理时包括了这些噪声的信息，扩大了水体区域中高亮噪声附近的错误检测概率。

5 结束语

本文提出一种用于检测 SAR 图像中水体区域的方法。该方法使用树型小波方法对 SAR 图像进行纹理分析，有效地利用 SAR 图像中的纹理信息，结合图像中的灰度信息，较好地描述了 SAR 图像中水体区域与非水体区域的信息。文中采用了具有强鲁棒性的 FW-SVM 分类方法，通过提高 SVM 分类器的惩罚系数，提高了对水体目标的正确分类概率。使用纹理信息结合 SVM 的方法，有效解决了阈值法不适合于具有复杂灰度分布的大范围图像的问题。实验表明，本文的方法对 SAR 图像中的水体区域具有良好的检测效果。

参考文献

- [1] Coggins J M. A Framework for Texture Analysis Based on Spatial Filtering[D]. Michigan, USA: Michigan State University, 1982.
- [2] Yukinobu M, Kuniaki U. Texture Analysis and Classification with Tree Structured Wavelet Transform[J]. IEEE Trans. on Image Processing, 1993, 3(4): 429-441.
- [3] Cortes C, Vapnik V. Support Vector Networks[J]. Machine Learning, 1995, 20(3): 273-297.
- [4] Branbanter J D. Weighted Least Squares Support Vector Machines: Robustness and Sparse Approximation[J]. Neurocomputing, 2002, 48(1): 85-100.