

基于多特征融合的彩色图像声呐目标检测

王 晓 邹泽伟 李勃勃 王 静

(云南大学信息学院 昆明 650000)

摘 要 随着国内对河流、湖泊和海洋近岸浅水区域水下工作的深入开展,潜水员进行水下打捞、定位以及勘探等水下工程建设变得意义重大。本实验室开发的专利产品 TKIS-I 头盔式彩色图像声呐获得中国海军航行保障部认可,目前已有 20 多台服务于部队并持续获得部队订货。但是,在复杂的水下环境中,潜水员进行水下作业具有较大的风险,所以期望今后能利用水下机器人实现自动水下目标检测,从而把潜水员从危险的复杂水下活动中解放出来。为此,文中针对声呐图像的特点,在颜色、形状、纹理 3 个方面分别采取了 HSV 颜色空间、梯度直方图(HOG)、局部二值模式(LBP)的特征提取方法,并且改进了多特征融合的方式,使用优化后的支持向量机(SVM)进行分类,旨在快速检测出水下目标,为以后水下机器人的自动目标检测奠定基础。

关键词 方向梯度直方图,彩色图像声呐,支持向量机,局部二值模式,HSV 颜色空间,多特征融合
中图分类号 TP399 文献标识码 A

Target Detection in Colorful Imaging Sonar Based on Multi-feature Fusion

WANG Xiao ZOU Ze-wei LI Bo-bo WANG Jing

(School of Information Science and Engineering, Yunnan University, Kunming 650000, China)

Abstract With the in-depth development of underwater work in rivers, lakes and offshore near-shore shallow water areas, diver's underwater engineering construction such as underwater salvage, positioning and exploration becomes significant. The TKIS-I helmet-mounted colorful imaging sonar developed by this lab has been acknowledged by Navigation and Warranty Department of Chinese Navy. Currently, there are more than two dozens of TKIS-I in service. However, under the complex underwater environment, divers usually perform underwater operations with great risks, so it is expected to use underwater robots to achieve automatic underwater target detection in the future. Aiming at the feature of sonar image, this paper adopted feature extraction methods of HSV color space, Histogram of Oriented Gradient(HOG) and Local Binary Pattern(LBP) respectively in the aspects of color, shape and texture. Besides, the paper improved multi-feature fusion method and used optimized support vector machine(SVM) for classification, aiming to quickly detect underwater targets to lay the foundation for robots' underwater automatic target detection in the future.

Keywords Histogram of oriented gradient(HOG), Color image sonar, Support vector machine(SVM), Local binary pattern(LBP), HSV color space, Multi-feature fusion

随着国内对河流、湖泊和海洋近岸浅水区域的不断探索和开发,水下环境变得越来越错综复杂,水下搜救、河底打捞等水下作业面临着更大的挑战^[1]。水下目标探测在水下作业中发挥了重要的作用。水下声呐成像技术作为水下探测的一个分支,有着不可或缺的作用。但是由于声呐数据集在计算机视觉领域的众多数据集中比较稀少,因此目前在视觉检测中很少出现水下声呐目标检测的研究。而本文采集声呐目标数据的设备——TKIS-I 头盔式彩色图像声呐获得了中国海军航行保障部认可。头盔式彩色图像声呐是一种具有高分率的图像声呐,能够轻易获取水下目标的形状和表面细节,方便了人工智能在水下声呐目标检测方向的研究。因为这套声呐设备需要潜水员佩戴进行水下作业,潜水员的安全始终都会受到威胁,所以本文根据水下图像声呐的成像特点,分别在

颜色方面使用 HSV 颜色空间提取,在形状方面使用对边缘提取效果表现很好的 HOG 特征,在纹理方面使用 LBP 特征。但是由于特征提取后的冗余信息太多,维度较高,因此我们需要采用主成分分析(PCA)对其进行降维,然后通过改进后的加权特征融合的方法有选择性地融合 3 种特征,以更好地表现出声呐图像中目标的特性。在分类器的选择上,本文采用支持向量机,并且在训练过程中加入难分样本挖掘(Hard Negative Mining)的思想去优化 SVM 模型,并把得到的模型运用到声呐目标检测中,为将来用水下机器人目标自动检测代替潜水员水下作业奠定基础。

1 彩色图像声呐及其成像特点

TKIS-I 图像声呐系统是一种潜水员单独携带进行水下

本文受国家自然科学基金(K1020546),云南省教育厅基金(K1050674)资助。

王 晓(1993—),男,硕士生,主要研究方向为计算机视觉、信号与信息处理,E-mail:wx20690@163.com;邹泽伟(1993—),男,硕士生,主要研究方向为高频水下图像声呐系统;李勃勃(1993—),男,硕士生,主要研究方向为水声信号处理;王 静(1970—),女,博士,副教授,主要研究方向为成像声呐、水下匹配场被动定位。

工作的小型高分辨率的彩色图像声呐。如图1所示,该系统主要由4个部分组成:腰间控制盒、视频眼镜、声呐探头和陆上工控机。在水下工作过程中,目标图像显示在视频眼镜中,潜水员可以根据目标显示的方位进行搜寻。图2为广州南海救助局声呐水下实验现场,图3为潜水员佩戴声呐的图像。



图1 TKIS-I图像声呐系统

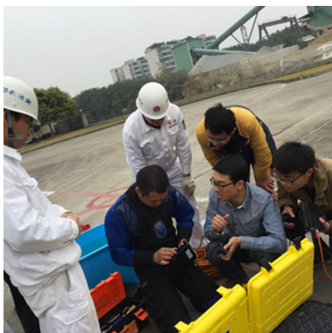


图2 广州南海救助局声呐水下实验现场

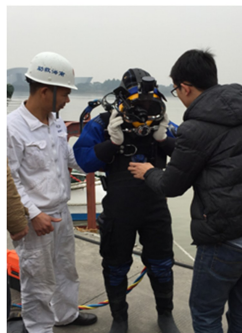


图3 潜水员佩戴 TKIS-I 图像声呐

在这次实验中,我们主要选择了学校的木桩水池,在水池的中间立有8根粗细不同的木桩,如图4所示。图5是TKIS-I彩色图像声呐获取的一帧图像,从图中可以清楚地看出木桩在这一帧图像中的确切位置和形状。但是在实际过程中,探测目标和成像声呐存在一起运动的情况,这就会导致目标在声呐中的成像发生平移、旋转等变化。因此,本次实验分别在不同位置、不同的方向和不同量程下搜集出目标成像图片。根据这种成像的特点,我们选取了多特征融合的提取方式,该方式对目标检测的效果很好。



图4 实验现场

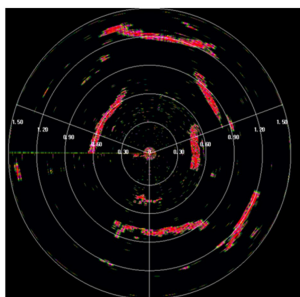


图5 TKIS-I图像声呐获取木桩的原始成像图

2 多特征提取与降维

2.1 颜色特征的提取

由于声呐成像的木桩在屏幕中主要显示为红色和绿色这两种颜色的混合色,不会随着位置的移动、方向的变化而改变,在背景为黑色的情形下具有很高的辨识性和区别性,所以本文采用这种混合颜色特征作为检测图像的一种特征。

在颜色特征提取中,RGB与HSV是最常用的模型。但是这两种模型中,HSV特征能更容易且直观地区分出颜色的属性。HSV颜色直方图的3个参数分别为色调、饱和度和亮度。在对图像提取HSV直方图的时候,因为得到的彩色图像声呐是RGB颜色空间中的图像,RGB图像本身也有自己的3个参数(分别对应红、绿、蓝3个颜色通道),所以需要将其转化到HSV空间中。首先将RGB颜色空间中的3个参数值归一化,然后利用以下转换公式得到HSV空间的3个参数值,从而完成转化。

$$h = \begin{cases} 0^\circ, & \max = \min \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, & \max = r \text{ and } g \geq b \\ 60^\circ \times \frac{g-b}{\max-\min} + 360^\circ, & \max = r \text{ and } g < b \\ 60^\circ \times \frac{b-r}{\max-\min} + 120^\circ, & \max = g \\ 60^\circ \times \frac{r-g}{\max-\min} + 240^\circ, & \max = b \end{cases} \quad (1)$$

$$s = \begin{cases} 0, & \max = 0 \\ \frac{\max-\min}{\max} = 1 - \frac{\min}{\max}, & \text{else} \end{cases} \quad (2)$$

$$v = \max \quad (3)$$

实验中,在提取HSV特征时也采用HOG特征分块的方法,将图像统一到 128×128 的大小,以 16×16 作为block的大小分割,每个block的滑动增量为 8×8 ,每幅图像就包含225个block,每个block中HSV的维数为3,所以得到整幅图像的特征维数为675维。

2.2 形状特征提取

在声呐图像中我们需要提取木桩的形状特征。形状特征的提取一般使用边缘轮廓,局部归一化的HOG描述子能够很好地描述目标的边缘轮廓情况。HOG在计算机视觉中主要是作为行人检测的算法,是Dalal和Triggs^[2]于2005年提出的。HOG是在网格密集、大小统一的细胞单元上进行计算,而且为了提高性能,还采用了重叠的局部对比度归一化。

在计算HOG特征的时候,首先将彩色声呐图像转换为灰度图像并将其统一到 128×128 的大小中;然后用梯度算子对原图像做卷积运算,分别得到水平方向和垂直方向的梯度分量;再用以下公式计算该像素点的梯度大小和方向。 $G_x(x,y)$ 和 $G_y(x,y)$ 分别表示输入图像在像素点 (x,y) 处的水平方向梯度和垂直方向梯度,则 $G(x,y)$ 和 $\alpha(x,y)$ 分别为像素点 (x,y) 的梯度幅值和梯度方向。

$$G_x(x,y) = I(x+1,y) - I(x-1,y) \quad (4)$$

$$G_y(x,y) = I(x,y+1) - I(x,y-1) \quad (5)$$

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \quad (6)$$

$$\alpha(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \quad (7)$$

之后把图像平均分成若干个正方形的胞元 (cell), 每个胞元中包含 8×8 个像素, 每个胞元中又把梯度方向 $0^\circ \sim 180^\circ$ 分成 9 个方向块, 对 cell 内每个像素用梯度方向在直方图中进行加权投影, 将其映射到对应的角度范围块内, 就可以得到这个 cell 的梯度方向直方图了。 2×2 个 cell 形成一个 block, 这样一个 block 就形成 36 维的特征向量, 再利用 L2-范数对整个 block 进行归一化, 从而得到最终的特征向量。

实验中将 16×16 作为 block 的大小分割, 在每个 block 中进行归一化处理。每个 block 中包含 2×2 的 cell, 所以每个 cell 所占的像素为 8×8 , 每个 block 的滑动增量为 8×8 , 那么每幅图像就包含 225 个 block。每个 block 包含 4 个 cell, 每个 cell 的维度为 9, 所以每个统一的 128×128 窗口图片的 HOG 特征向量维度为 8100。

2.3 纹理特征提取

LBP 是一种用来描述图像局部纹理特征的算子, 它是由 Ojala 等^[3]在 1994 年提出。原始的 LBP 特征使用的是固定邻域内的灰度值, 因此当图像的尺度发生变化时, LBP 特征的编码将会发生错误, LBP 特征将不能正确地反映像素点周围的纹理信息, 这显然不能满足不同尺寸和频率纹理的需要。为了适应不同尺度的纹理特征, 并满足灰度和旋转不变性的要求, 改进后的 LBP 算子将 3×3 邻域扩展到任意邻域, 并用圆形邻域代替正方形邻域, 改进后的 LBP 算子允许在半径为 R 的圆形邻域内有任意多个像素点。经过上述操作, 得到了诸如半径为 R 的圆形区域内含有 P 个采样点的 LBP 算子。使用符号 $LBP_{\frac{P}{R}}$ 表示 $LBP_{\frac{P}{R}}$ 邻域的一种扩展的均匀模式编码, 即该邻域的 LBP 二进制码对应的二进制数从 0 到 1 或从 1 到 0 最多有两次跳变, 如 11111111 (0 位转变), 11100111 (2 位转变)。图 6 给出几种均匀模式的 LBP 算子。

本文将 128×128 的统一窗口图像按照 16×16 的大小分割, 然后针对 16×16 的子图像, 根据 $LBP_{\frac{P}{R}}$ 算子求出其纹理

特征, 再将特征向量转化为 59 维的特征向量。与 HOG 中的 block 归一化相似, 需要对提取出来的 59 维特征向量进行归一化。根据实验中的归一化因子对比, 本实验最后都选取了 L2-normal 归一化因子。因为实验统一的窗口图像为 128×128 , 所以检测窗口有 64 个 cell, 每个 cell 的特征向量为 59 维, 因此最终可以得到 3776 维的 LBP 特征向量。

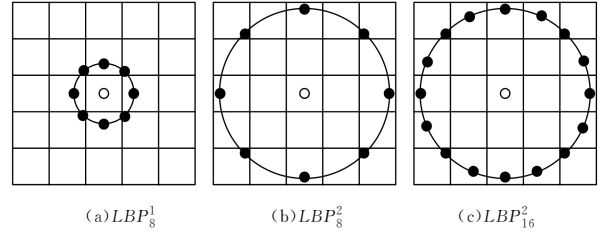


图 6 几种均匀模式的 LBP 算子

2.4 PCA 降维

PCA 又称主成分分析法, 它是将原本的样本空间进行空间变换, 把原本的数据投影到维度较低的空间中。这种方法可以减少原始样本中大量冗余的信息和噪音, 提高检测率, 加快训练的速度, 因此我们在特征融合前先把 3 种特征各自降维, 得到检测率最高的维数。

本文采用协方差矩阵将高维度空间的特征向量做线性投影, 形成新的特征向量。通过训练样本的特征均值 α 和需要降维的特征 x_i , 利用式 (18) 求出协方差矩阵 U^T :

$$U^T = \frac{1}{N} \sum_{n=1}^N (x_n - \alpha)(x_n - \alpha)^T \quad (8)$$

实验中取出协方差矩阵前 m 个主成分, 利用式 (9) 对训练样本中的每个特征值进行降维, 得到主成分 y , 其中通过不同 m 对应的检测率 (见表 1—表 3) 大小得出最佳的维度 m 。

$$y = U^T (x_i - \alpha) \quad (9)$$

从表 1—表 3 中可以看出, HSV 特征、HOG 特征、LBP 特征分别在 300 维度、1100 维度和 300 维度时的检测率最高, 同时相比于未使用 PCA 降维的原本特征, 检测率都有所提高, 所以这 3 种维度作为本实验进行 PCA 降维后的最佳维度。

表 1 不同维度 HSV 特征的检测率比较

HSV-PCA 维度										
维度	20	30	40	50	100	200	300	400	500	600
检测率	0.7895	0.7963	0.8013	0.8064	0.8215	0.8316	0.8487	0.8358	0.8349	0.8338
未使用 PCA 降维的原始 HSV 特征维度										
维度	675									
检测率	0.8363									

表 2 不同维度 HOG 特征的检测率比较

HOG-PCA 维度										
维度	200	300	400	500	700	900	1100	1300	1500	1700
检测率	0.8775	0.8867	0.8923	0.9062	0.9112	0.9168	0.9247	0.9188	0.9179	0.9128
未使用 PCA 降维的原始 HOG 特征维度										
维度	8100									
检测率	0.8967									

表 3 不同维度 LBP 特征的检测率比较

LBP-PCA 维度										
维度	20	30	40	50	100	200	300	400	500	1000
检测率	0.8134	0.8280	0.8319	0.8375	0.8466	0.8587	0.8684	0.8596	0.8521	0.8529
未使用 PCA 降维的原始 LBP 特征维度										
维度	3776									
检测率	0.8369									

3 多特征融合

3.1 实现背景

HOG 特征虽然能较好地反映边缘轮廓信息,但当背景覆盖声呐目标图像的噪声边缘时,HOG 处理的效果就会变得比较差;而 LBP 可以滤出噪声,并且结合边缘形状信息和纹理信息,能够很好地记录声呐图像中木桩的外表。HSV 特征能够直观反映出木桩在声呐图像中的混合颜色特征,并且提取的是图像的全局信息,这 3 种特征结合在一起更具有鲁棒性。

传统的多特征融合一般是将几种特征串连在一起,相当于几种特征直接相加,这种方法忽略了图像本身对每种特征的表达能力的不同,所以我们需要将这 3 类特征单独放入各自的分类器中训练,最后检测,以观察哪一种特征对图像的表达能力更强,从而分配不同的权重。

很多已有加权特征融合的方法是把正确率作为权值进行融合的,其实正确率并不能很好地反映检测的效果,比如当正类数目要比负类数目少很多时,分类的结果就会集中于负类,此时的正确率较高,但是分类错误。在二分类的问题中,正类被检测为正类的个数为 TP ,负类被检测为正类的个数为 FP ,正类被检测为负类的个数为 FN ,负类被检测为负类的个数为 TN ,则正确率 (ACC)、真正类率 (TPR)、负正类率 (FPR) 的计算公式为:

$$ACC = (TP + TN) / (TP + TN + FP + FN) \quad (10)$$

$$TPR = TP / (TP + FN) \quad (11)$$

$$FPR = FP / (FP + TN) \quad (12)$$

本文对此作出改进,采用 ROC 曲线中的 AUC 值作为特征融合的权值。ROC 曲线被称为受试者工作特性曲线,横坐标代表 FPR,纵坐标轴代表 TPR,AUC 值就是该曲线下的面积。从式(10)~式(12)可以看出,真正类率刻画的是分类器所检测出的正类占所有实际正类的比例,负正类率刻画的是分类器错误正类的负类占所有负类的比例。因此,曲线上的各个点在受到一种特征刺激后,通过真正类率和负正类率这两种判定指标来得出结果。AUC 作为该曲线下的面积,可以更加直观、精确地反映出每个特征对分类器的效果。如果一个模型是随机猜测模型,则 AUC 的值为 0.5;如果是完美模型,则 AUC 的值为 1。因此模型分类检测的好坏可以通过比较 AUC 值的大小判断,AUC 越大,模型越好。另外,ROC 曲线具有很好的稳定特性,当测试集的正负样本分布发生变化(比如负样本的数量要远远多于正样本的数量,或者数据中正负样本的分布随时间发生变化)时,ROC 曲线都能保持不变。

多特征融合示意图如图 7 所示。

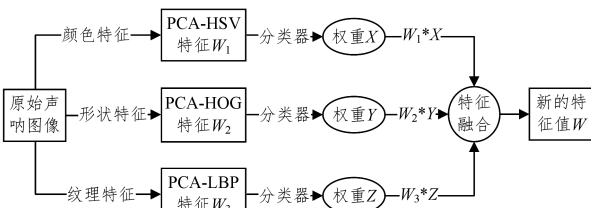


图 7 多特征融合的系统框图

原始声呐图片分别经过这 3 种特征被分类,然后根据各自在测试集中所得到的 AUC 值分别为 M_1, M_2, M_3 ,这 3 个值的和为 M ,则特征融合时,颜色特征 W_1 所占权值 X 为 M_1/M ,形状特征 W_2 所占权值 Y 为 M_2/M ,纹理特征 W_3 所占权值 Z 为 M_3/M ,并且其值分别等于 X, Y, Z 。融合后的总特征表示为:

$$W = W_1 \cdot X + W_2 \cdot Y + W_3 \cdot Z \quad (13)$$

3.2 实现过程

首先把已经经过 PCA 降维后的 3 种特征分别通过分类器分类,对得到的模型再利用测试集进行测试。测试集共有 4250 张图片,其中含有正测试样本和负测试样本,然后将 3 条曲线进行多项式回归拟合,从而得到 3 种特征的 ROC 曲线,如图 8 所示。可以看到,HOG 表现形状特征的 AUC 值最大,然后是 LBP 表现的纹理特征,HSV 表现的颜色特征的 AUC 值最小。最后根据 3 种特征表现的 ROC 曲线中的 AUC 值进行上述权值分配,加权融合后,得到最后的总特征。

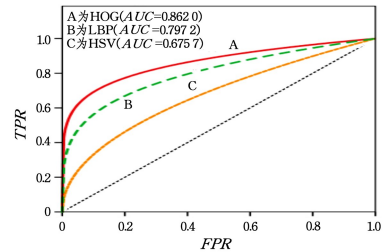


图 8 3 种特征各自表现的 ROC 曲线

4 分类器的分类与优化

本文选择使用支持向量机对正负样本进行训练。支持向量机适用于解决小样本及高维的模式识别,运用核函数,可以将低维度的空间转化成高维度的空间,然后在高维空间中求得最优的分隔超平面,从而使得数据集线性可分,同时 SVM 分类器的时间开销较小,计算速度快。输入空间的数据为 x ,支持向量机的一般形式为:

$$g(x) = \omega \cdot x + b \quad (14)$$

最优分隔超平面要求距离超平面最近的样本距离尽可能大,那么检测样本的检测率就会越大,分类的效果也就越好,所以要求:

$$y_i [(\omega \cdot x) + b] \geq 1, i = 1, 2, \dots, n \quad (15)$$

其中,样本点为 (x_i, y_i) ,样本的类别标识 y_i 为 1 或者 -1。当 ω 的范数最小时,间隔就最大,然后在满足式(15)的限制条件下求出 ω 和 b 。

本文选用径向基函数 (RBF) 作为 SVM 的核函数,因为 RBF 核函数可以把输入的数据映射到高维空间中去选择最优的分类面,实现非线性映射,分类效果很好。

为了使检测的结果更好,在分类的优化方面我们选择使用难分样本挖掘 (Hard Negative Mining)^[4] 的思想,先用初始的正负样本训练分类器得到一个模型,然后再用训练出来的模型对所有负样本进行滑窗预测。如果该窗口的预测结果为正,那么将该窗口加入训练集,指定标签为负,对样本再进行分类,如此反复,直到达到停止条件(比如分类器性能不再提升),所以该方法在原始正负样本的选择方面,负样本的数量要远远大于正样本的数量。因此,我们选取了 8000 张 128×128 的负样本图片,并将得到的优化后的分类模型作为我们最终的目标检测模型,最后使用目标检测模型在测试集中进行检测。

5 数据集制作与实验结果分析

5.1 训练与检测的数据

根据上文描述,我们选取池塘中的木桩作为本次实验目标检测的对象。本文训练样本使用的是实验室 TKIS-I 头盔式彩色图像声呐采集的数据,首先利用该声呐设备得到木桩

在声呐图像中的原始图片,然后利用自己编写的标记软件进行标注,正样本中的目标标签保存在 xml 格式文件中。负样本是通过自己编写的程序将原始正样本木桩框外的图片进行截取,然后经过人为筛选得到。该数据中共有 2000 张正样本图片和 8000 张负样本图片。测试集中共有 850 张正样本图片和 3400 张负样本图片。样本数据集总体比较清晰,木桩的方向和大小都有不同,部分正负样本如图 9 所示。

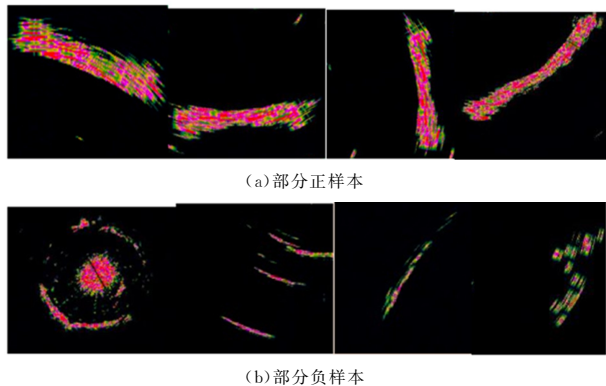


图 9 部分正负样本

5.2 实验结果分析

本文将最后训练出来的模型放在测试集中检测,并且加入对比实验,与单独的特征提取进行检测正确率的对比,得到的 AUC 值如表 4 与表 5 所列。

表 4 不同特征提取的检测率比较

特征种类	检测率
HSV	0.8487
HOG	0.9247
LBP	0.8684
直接串连 HSV+HOG+LBP	0.9793
本文融合方法 HSV+HOG+LBP	0.9868

表 5 不同特征提取的 AUC 值比较

特征种类	AUC
HSV	0.6757
HOG	0.8620
LBP	0.7972
直接串连 HSV+HOG+LBP	0.9176
本文融合方法 HSV+HOG+LBP	0.9578

从表 4 和表 5 中可以看出,特征融合后检测的正确率和 AUC 值都比单一特征的高。通过每个特征各自的正确率和 AUC 值可以得出,对声呐图像木桩表现最强的特征是形状特征 HOG,其次是纹理特征 LBP,最后是颜色特征 HSV。观察正确率发现:直接特征串连的方法与本文加权特征融合方法的 AUC 没有很大的差别,但是从最后 AUC 值的大小中看出本文的融合方法更为准确。图 10 为最后声呐木桩图像检测结果示意图,可以看出作为目标的木桩基本被正确框出。



图 10 声呐图像木桩检测图

结束语 本文采用改进的多特征融合的提取方式和 PCA 降维加上 SVM 分类器训练与优化的方法对声呐图像目标检测,得到了较好的效果,为以后水下机器人自动目标检

测取代潜水员下水作业进行奠定了基础。

目标检测滑框法得到的检测目标有时候并未完全被框中,而且当目标周围有许多干扰或者噪音时目标的形状变得不清晰,从而影响检测的准确度,这也是本文后续需要改进和探索的方向。

本文研究选择的水下目标为木桩,该目标可以在声呐图像中清晰可见,并得到了比较好的效果。但是声呐探测的实际目标类型很多,比如人、船只或者其他不规则形状的物体,由于实验室目前所具备的图像资料较少,而部队使用中的资料又难以获取,因此本文的研究暂时难以做到多类目标检测。随着获取的水下目标的声呐图像越来越多,本文的研究也将逐步深入和完善。

参 考 文 献

- [1] 谭晓岚. 论海洋经济发展的总体趋势[J]. 海洋开发与管理, 2009,26(7):12-16.
- [2] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2005:886-893.
- [3] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002,24(7):971-987.
- [4] 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, 2010,32(9):1627-1645.
- [5] WANG X, HAN T X, YAN S. An HOG-LBP human detector with partial occlusion handling[C]// 2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009:32-39.
- [6] CANNY J. A Computational Approach to Edge Detection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1986,8(6):43-57.
- [7] LOWE D G. Distinctive image features from scale-invariant key points [J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [8] BELONGIE S, MALIK J, PUICHIA J. Shape context: a new descriptor for shape matching and object recognition[C]//INPS'00. Cambridge: MIT press, 2000:831-837.
- [9] WU B, NEVATIA R. Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors[C]// Tenth IEEE International Conference on Computer Vision, 2005 (ICCV 2005). IEEE, 2005:90-97.
- [10] BELL J M, PETILLOT Y R, LEBART K, et al. Target recognition in synthetic aperture and high-resolution sidescan sonar[C]// 2006 IET Seminar on High Resolution Imaging and Target Classification. IET, 2006:99-106.
- [11] BLONDEL P. The Handbook of Sidescan Sonar[M]// Berlin: Springer, 2009.
- [12] BENGIO Y, LECUN Y. Scaling learning algorithms towards ai [J]. Large-scale Kernel Machines, 2007, 34(5):1-41.
- [13] BENGIO Y. Learning deep architectures for ai[J]. Foundations and trends in Machine Learning, 2009, 2(1):1-127.
- [14] ANZAI Y. Pattern recognition and machine learning[M]. Springer, 2012.