

## 基于图像显著性区域的遥感图像机场检测

王鑫<sup>1)</sup>, 王斌<sup>1,2)</sup>, 张立明<sup>1)</sup>

<sup>1)</sup>(复旦大学电子工程系 上海 200433)

<sup>2)</sup>(复旦大学波散射和遥感信息教育部重点实验室 上海 200433)

(wangbin@fudan.edu.cn)

**摘要:** 针对已有方法对图像逐像素进行分析的不足, 将人眼的注意力选择计算模型引入到遥感图像的机场目标检测中, 提出一种基于图像显著性区域的遥感图像中机场目标检测与识别的方法, 以提高自动目标检测的效率. 首先利用霍夫变换对遥感图像中是否存在机场目标进行初步筛选, 然后利用改进后的基于图像的视觉显著性模型提取显著性区域, 根据区域上的尺度不变特征变换特征并结合多层分类回归树完成机场目标的识别. 实验结果表明, 该方法比现有的其他机场检测方法具有速度快、识别率高、虚警率低的特点, 同时对噪声有较强的鲁棒性.

**关键词:** 视觉注意; 显著性区域; 机场检测; 尺度不变特征变换; 多层分类回归树; 霍夫变换

中图分类号: TN911.73

## Airport Detection Based on Salient Areas in Remote Sensing Images

Wang Xin<sup>1)</sup>, Wang Bin<sup>1,2)</sup>, and Zhang Liming<sup>1)</sup>

<sup>1)</sup>(Department of Electronic Engineering, Fudan University, Shanghai 200433)

<sup>2)</sup>(The Key Laboratory of Wave Scattering and Remote Sensing Information (Ministry of Education), Fudan University, Shanghai 200433)

**Abstract:** In this paper we propose a new airport detection and recognition method for remote sensing images based on salient areas with the introduction of visual attention models to improve the efficiency of automatic target detection. Firstly, Hough transform is used to judge the existence of an airport, and then the improved graph-based visual saliency (GBVS) model is used to extract regions of candidates (ROCs). According to the features with scale-invariant feature transform (SIFT), which are extracted from ROCs and classified by hierarchical discriminant regression (HDR) tree, the airport areas are recognized. Experimental results show that our method is faster with higher recognition rate and lower false alarm rate than other current methods, and is robust against white noise.

**Key words:** visual attention; salient areas; airport detection; scale-invariant feature transform; hierarchical discriminant regression tree; Hough transform

遥感图像中典型目标的检测与识别是当前自动目标检测与识别研究的热点. 机场作为一类特定目标, 在军事和民用领域都具有十分重要的意义, 受到人们越来越多的关注. 然而, 机场所处的区域背景往往很复杂, 这给它的检测与识别带来了困难. 通过对遥感图像的观察, 我们总结出遥感图像中机场检测问题的要点为如下: 首先是背景复杂, 机场或处于

市区中, 或处于偏远郊区, 城镇、山脉、公路等自然或人工景物具备一些与机场相似的特征; 其次, 机场最主要的特点在于它有 2 条平行的或纵横交错的跑道, 且具有一定的灰度、长度和宽度; 此外, 遥感图像覆盖范围大, 图像像素数多, 所以对检测和识别方法的速度、复杂度都有一定要求, 以符合实时性需要.

迄今为止, 国内外学者针对机场检测问题提出

收稿日期: 2011-05-14; 修回日期: 2011-08-11. 基金项目: 国家自然科学基金(61071134) 王鑫(1987—), 男, 硕士研究生, 主要研究方向为遥感图像目标识别; 王斌(1964—), 男, 博士, 教授, 博士生导师, 主要研究方向为信号和图像处理及其他的应用; 张立明(1943—), 女, 教授, 博士生导师, 主要研究方向为神经网络模型及其在图像识别上的应用.

了许多方法,可以将这些方法大致分为 2 大类<sup>[1]</sup>:基于边缘提取的方法和基于区域分割的方法.前者是根据机场跑道的平行长直线特性,从图像中提取它的边缘,然后通过 Hough 变换<sup>[2]</sup>等方法进行直线检测并辅以纹理识别,以达到机场的确认<sup>[1, 3-4]</sup>.后者则是基于机场与周围纹理的差异性,通过图像分割的方法达到提取备选区域的目的,再在备选区域内进行识别<sup>[5-7]</sup>.

上述 2 类方法各自有优点也有不足.基于边缘提取的方法速度快、复杂度低,但由于背景的干扰,有许多非机场目标的区域会被检测出来,如公路、山脉等等,这些干扰物都会影响识别的准确性,且仅使用 Hough 变换的结果来定位机场是不精确的.而基于区域分割的方法虽然在刻画机场区域方面要做得更好,但其在图像分割时需要逐个像元进行分析,因而速度慢、复杂度高,识别效果依赖于区域分割阈值的选取以及用于识别的特征的鲁棒性好坏.值得一提的是,目前关于机场检测的文献中对于识别的虚警率往往避而不谈.当对各幅不同的图像进行识别时,不仅要保证目标不被遗漏,还要确保没有目标的地方不被识别出目标,这样才有实际的应用价值.

近年来,注意力选择<sup>[8]</sup>被广泛地应用到模式识别领域中,并且在自然图像的目标识别方面已经收到了很好的效果.所谓视觉注意力,就是指视觉图像场景中引起我们关注的区域.人可以在短时间内找出眼前具有显著特征的物体,机器视觉却无法做到,这是因为人的注意力选择机制在发挥着作用.注意力选择过程可分为自底向上和自顶向下 2 部分:自底向上的过程是指由输入图像生成显著图的过程,显著区域即那些具有较强对比度的区域或与周围有明显差异的区域,这一过程是数据驱动的,它不受主观意识支配;而自顶向下的过程则是与人的先验知识及判断有关的,有助于从显著区域中确定目标区域.针对机场检测问题,考虑机场与周围的差异性比较大,可以认为其是显著的,因此,本文将注意力选择机制引入到遥感图像机场目标的检测中来,以达到提取机场备选区域的目的;在克服了以往方法对图像逐像素进行分析导致速度过慢的同时,解决了机场定位不精确的问题,在保证较高检测和识别率的同时,提高自动目标检测的效率.

目前已有许多学者对注意力选择机制进行了建模.早期 Itti 等<sup>[9]</sup>和 Walther 等<sup>[10]</sup>在计算结构上模仿人脑形成视觉显著性的神经机制,分别提出了 2 种注意力选择计算模型,但其计算量大、复杂度高,

难以应用在工程上.为简化计算复杂度,基于频域的视觉显著性计算方法被提出来,包括 Hou 等<sup>[11]</sup>提出的残留谱法、Guo 等<sup>[12-13]</sup>提出的傅里叶变换相位谱法和四元数的傅里叶变换相位谱法,以及 Ding 等<sup>[14]</sup>提出的双四元数法.这类模型运算速度很快,在自然图像上收到了不错的效果.但实验后发现,上述模型在遥感图像复杂背景的情况下,识别效果不甚理想.因为自然图像的背景往往是平滑单一的,目标很容易被分离出来,而遥感图像杂乱无章的地表背景严重干扰了识别效果.与之不同的是, Harel 等<sup>[15]</sup>的基于图的视觉显著性 (graph-based visual saliency, GBVS) 模型相对于其他方法有较好的效果,但也有不足,且由于其运行速度过慢而难以实时操作.

本文根据机场检测这一任务,对 GBVS 模型进行改进,使其在运算速度、检测效果上都有大幅度提升.然后将改进后的 GBVS 模型引入到遥感图像机场目标检测中来,生成原始灰度图像对应的显著图,并根据显著性大小(即显著图内像素值的大小)得到几个可能存在目标的候选区域,再对每个候选区提取尺度不变特征变换 (scale-invariant feature transform, SIFT) 特征点和特征向量<sup>[16]</sup>.本文方法会事先在训练数据上提取一些具有机场和非机场(如房屋、植被、山脉)的 SIFT 特征向量,用多层分类回归 (hierarchical discriminant regression, HDR) 树<sup>[17]</sup>来训练,并将先前候选区域上提取的 SIFT 特征代入到训练好的 HDR 树中进行识别.最后根据各个候选区域的显著性大小和区域上具有机场特征的特征点数的多少,确定一个候选区域为机场区域.

## 1 改进的 GBVS 模型

GBVS 模型是在 Itti 的模型<sup>[9]</sup>的基础上运用马尔可夫随机场的特点构建二维图像的马尔可夫链,通过求其平衡分布而得到显著图.具体步骤如下:

Step1. 获取多尺度的亮度信息.对输入的灰度图像  $I$ ,使用高斯金字塔低通滤波器对其进行滤波.高斯金字塔的每一阶都是一个二维高斯低通滤波器

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right);$$

其中,  $(x, y)$  为图像中点的位置,  $\sigma$  称为尺度因子.所谓金字塔,就是指对原始图像不断地进行 1/2 降采样和高斯低通滤波,滤波的  $\sigma$  随图像的变小而不断减小,得到一组不同尺度下的滤波结果.这一组滤波结果表示亮度通道.

Step2. 获取多尺度的方位信息. 用 Gabor 金字塔滤波器组对原始灰度图像  $I$  进行滤波, 以得到方向上的信息. 二维 Gabor 滤波器表示为

$$H(x, y, \sigma, \theta) = \frac{1}{\sigma^2} \exp\left(-\pi \frac{x^2 + y^2}{\sigma^2}\right) \left[ \exp(i2\pi f(x \cos \theta + y \sin \theta)) - \exp\left(-\frac{\pi^2}{2}\right) \right];$$

其中,  $\sigma$  为尺度因子,  $f$  为正弦波频率,  $\theta$  为方位. 一般情况下可以取  $\theta = \left[0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right]$ , 即在 4 个方向上滤波. 这样, 与亮度信息类似, 可以得到 4 组不同尺度下的滤波结果, 表示方向通道. 在文献[15]中, 亮度和方向通道均使用 3 个尺度进行计算.

Step3. 求不同尺度和不同特征图的马尔可夫平衡. 对上述每个通道内每个尺度的滤波结果, 根据其像素间的差异和欧氏距离建立各自的马尔科夫链, 然后求其马尔科夫平衡分布.

Step4. 计算显著图. 将所有组、所有尺度的滤波结果依次计算出平衡分布后, 将结果按照通道叠加起来并归一化, 得到综合显著图, 其大小与原始图像一致.

以往的文献中普遍认为 GBVS 最大的问题在于运算速度<sup>[12-13]</sup>, 考虑到这一问题, 再结合机场检测这一特定任务, 本文对 GBVS 模型做出了以下几点改进:

1) 因为机场具有一定范围的大小且机场之间大小差异不是很大, 故将文献[15]中亮度和方向通道的 3 个尺度改为 2 个尺度; 也就是说金字塔只有 2 层, 并且使用的原始图像也经过降采样处理, 以使运算速度得到保证.

2) 由于机场跑道的存在, 机场具有很好的方向性. 为此本文将方向滤波器的数目从 4 个扩展到 8 个, 以便更好地抓住这一特点.

3) 本文对原始图像进行 Hough 变换后将其作为单独的一个通道加入到模型中, 将检测到的直线用值为 '1' 的像素表示, 并用高斯函数平滑, 可以得到 Hough 通道的特征图. 最后将其与亮度通道、方向通道的特征图等权重相加并归一化, 得到最终显著图.

4) 考虑到大部分机场都具有中等以上的亮度, 为了排除较暗背景(如河流、峡谷)的干扰, 本文将之前的显著图与经过一次滤波的原始灰度图像  $I$  相乘, 得到新的显著图, 并作为最终结果.

## 2 SIFT 特征

如何找到一个鲁棒性强的、有代表性的特征来

描述区域特性, 是模式识别的一个重要问题. 常见的一些特征多以统计特征为主, 例如均值、方差、各阶矩、旋转动量等, 但它们或者代表性不强, 或者受仿射变换影响大, 或者计算速度缓慢, 不适合复杂背景下的快速目标检测. 文献[5]采用了 SIFT 特征作为机场位置确定的依据, 收到了很好的效果, 本文也引入 SIFT 特征来识别机场目标.

SIFT 特征算法于 2004 年由 Lowe 提出<sup>[16]</sup>, 它根据图像中的极值点及其周围区域的信息来生成相应的特征向量, 较其他统计性特征, SIFT 特征具有旋转、缩放、平移不变性, 对于不同大小、方向的机场有着较好的适应性, 并且该算法具有较快的计算速度. 具体的 SIFT 特征提取算法参见文献[16]. 在 SIFT 特征提取过程中, 尺度的不同保证了它的缩放适应性, 而坐标轴的旋转保证了其旋转不变性. 实验中发现, 对含有机场的图像提取 SIFT 特征时, 机场上的特征点有不少落在跑道拐角、直线交叉的区域, 而跑道的交叉、拐角也是大多数机场所共有的特征, 可见 SIFT 特征可以较好地反映机场的特点.

## 3 HDR 树

数据训练是目标识别不可缺少的一部分, 训练的的目的是为测试的特征提供分类的依据. 支撑向量机(support vector machine, SVM)是一种常用的分类器, 文献[3, 5]都将它用作机场纹理识别的工具, 而本文使用 HDR 树来进行训练. HDR 树是一种具有快速学习和取回能力的记忆树, 可以将它用于模拟人脑视觉记忆的 2 种行为: 储存和取回. 相比 SVM, HDR 树由于其树状结构同样具备很快的识别速度, 而且还可以不断地用新样本更新训练树, 因此不用像 SVM 那样需要重新训练所有样本. HDR 树的具体结构详见文献[17]. 当有新的样本要参与训练时, 新样本只需顺着树枝寻找与自己最相似的叶子节点, 并与之合并或分裂.

## 4 基于显著性区域的机场检测

本文运用改进后的 GBVS 模型实现机场检测. 首先用 Hough 变换对输入图像是否含有机场目标进行判断, 然后对可能存在目标的图像使用改进的 GBVS 模型提取显著区域, 再对每个显著区域提取 SIFT 特征并输入到训练好的 HDR 树中进行识别, 以确定机场的位置. 其整个流程如图 1 所示.

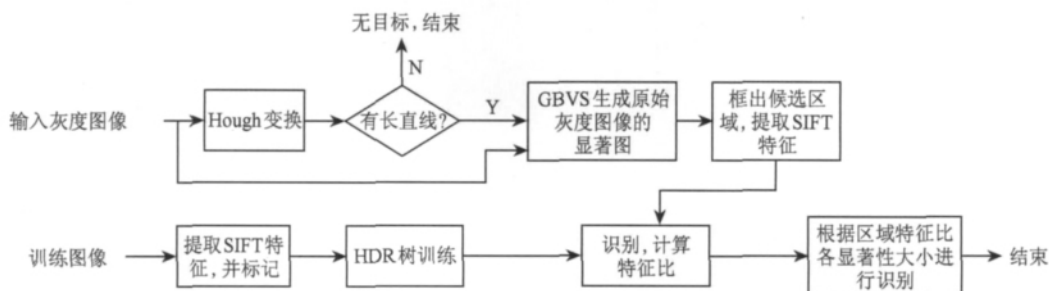


图 1 算法流程示意图

下面将整个方法的具体步骤描述如下:

Step1. 训练. 从原始的灰度图像中选取一些作为训练样本并提取 SIFT 特征,把得到的 SIFT 特征对应的特征向量分成 2 类:属于机场区域的标记为 1,属于背景区域的标记为 0;随后把这些带有标记的训练向量输入到 HDR 树中进行训练.图 2 所示为一个训练样本示例.

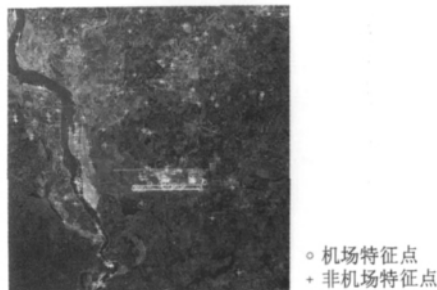


图 2 SIFT 特征提取与训练示例

Step2. 预处理. 对于一幅待识别的图像,首先对其是否存在机场目标做一个初步的判断.这样做可以减少对无机场目标图像不必要的分析处理过程,加快了检测速度,同时还可以降低虚警率,排除掉一部分可能导致虚警发生的输入图像.本文采用 Hough 变换实现筛选功能,并把变换后的结果单独作为一个通道加入到 GBVS 中.这是因为 Hough 变换可以较好地反映机场跑道长直线的特性,且其运算速度快.具体的做法是对原始图像进行二值化处理后做 Hough 变换,如果存在一定范围长度内的直线,则认为可能存在机场并进行显著性分析;否则直接按无目标跳过. Hough 变换的参数设置如下:取覆盖像素最多的 4 条直线,线段间填充值为 4 个像素,最短直线长度为 40.本文方法中,Hough 变换仅仅用来判断机场是否存在,并不用于机场定位,因此也不会涉及文献[3]中的 Hough 变换定位不精确的问题.

Step3. 显著图生成. 对于含有目标的图像,用改进的 GBVS 模型生成它的显著图.

Step4. 候选区生成. 记显著图上最大像素值为  $I_{max}$ :

Step4. 1. 从显著图中最亮的一个点向周围 8 邻域方向生长,如果某个方向上有某个像素值小于最亮点像素值的  $\alpha$  倍,则该方向停止生长,其中,  $\alpha$  为取值在  $[0, 1]$  间的阈值.这一过程直到所有方向上的生长都停止为止;

Step4. 2. 将生长完成的区域的外接矩形框画出来,

此即为一个候选区域,然后将该候选区域内的所有像素值清零,以免被再次选中;

Step4. 3. 计算上个候选区清零后的显著图中的最亮点的像素值.如果其小于  $I_{max} \times \alpha$ ,则表明所有候选区已提取完毕,进行 Step5;否则回到 Step4. 1,以提取新的候选区.

Step5. 特征识别. 对先前提取得到的候选区依次提取 SIFT 特征和对应的特征向量,并把特征向量代入到训练好的 HDR 树中进行识别,获得每个特征点的标记(1 为机场,0 为背景).待所有特征点标识完毕,计算各个候选区的特征比,即

$$\text{特征比} = \frac{\text{标记为 1 的特征点数}}{\text{该区域内特征点总数}}$$

Step6. 确定机场区域. 先前的候选区提取是按照显著性从大到小的顺序确定候选区域的,那么每个候选区域有显著性次序和特征比 2 个评价参数.我们先只观察特征比而不考虑区域显著性,如果只有一个区域的特征比大于设定的阈值  $\beta$ ,则该区域即为机场区域,不再考察其他区域;若没有这样的区域,则按照区域的显著性大小次序依次考察各个候选区,只要考察的区域内存有标记为 1 的特征(即特征比不为 0),便将该区域判定为机场区域.

## 5 实验结果

将本文方法用于实际的遥感图像中进行实验,并与文献[3,5]方法做比较.实验数据来自 Google Earth. Google Earth 上的数据是卫星影像与航拍数据的整合,其中卫星影像多为 Landsat-7 的数据.参考以往方法<sup>[1,3-7]</sup>中选取的原始图像覆盖范围,本文以  $30\text{ m} \times 30\text{ m}$  的分辨率从世界各地截取了 200 幅  $400 \times 400$  大小的彩色图像,其中包含机场与不包含机场的各有 100 幅.对于覆盖范围更大的遥感图像,应先把它分割成几个  $400 \times 400$  大小的图像块;然后再逐块进行分析,这样可以节约计算内存空间,相应地降低计算硬件性能要求.本文先仅考虑单波段遥感图像的机场检测问题,故对彩色图像的 R, G, B 三通道像素值直接求和后再求平均,并将所得到的灰度值作为实验数据.此外,因为每幅图像上能够提取出大量的 SIFT 特征向量,所以仅选取其中包含

机场的 7 幅作为训练数据就足够了,其余的 193 幅都作为测试数据.所有实验均在 Inter Core 2 Duo CPU 2.53 GHz 主频,2 GB 内存的 Windows 平台上进行,程序代码是在 Matlab R2008a 上编写的.

### 5.1 参数分析

设测试图像中含机场的为  $n$  幅,其中被正确识别出来的为  $x$  幅,那么识别率  $r$  计算公式为

$$r = \frac{x}{n} \times 100\% \quad (1)$$

现在对本文方法中涉及到的 2 个参数  $\alpha$  和  $\beta$  对识别率的影响进行分析,以确定后续实验中所选择的参数.参数  $\alpha$  的取值主要影响候选区域的大小和个数,选取  $\alpha$  从 0.1 变化到 0.9,步长 0.05,分别计算每个  $\alpha$  取值下的识别率,绘制出相应的识别率曲线如图 3 所示.

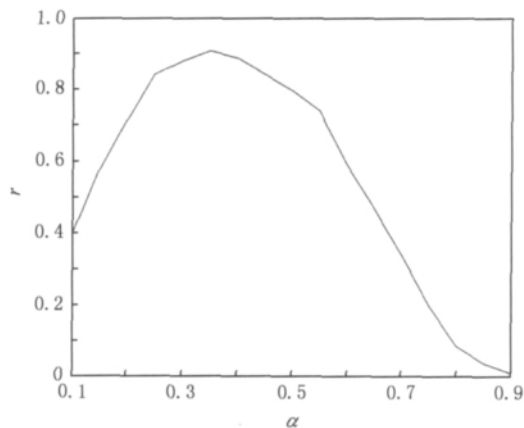


图 3 参数  $\alpha$  对识别率的影响

从图 3 可以看出, $\alpha$  取值在 0.3~0.4 之间有着比较好的识别结果.当  $\alpha$  取值比较小时,候选区的数目会增加,增大了识别的困难度,同时候选区的面积会很大;即使挑选出了含有机场的候选区域,该区域内也会有大面积的非机场区域,我们认为这样不精确的定位是属于识别失败的一类.而当  $\alpha$  取值比较大时,候选区域数目的减小使得一些显著性不是最高的机场被排除在外;此外候选区面积的减小会导致候选区只覆盖到机场很小的一部分,这样的识别也是失败的.

再考虑阈值  $\beta$  对识别率的影响. $\beta$  的作用主要在于排除那些显著性高于机场且具有一定的机场特征的非机场区域,使得机场区域在不是第一显著性的情况下也能被正确识别.选取  $\beta$  从 0 变化到 0.3,步长 0.02,分别计算每个  $\beta$  取值下的识别率,绘制出相应的识别率曲线如图 4 所示.

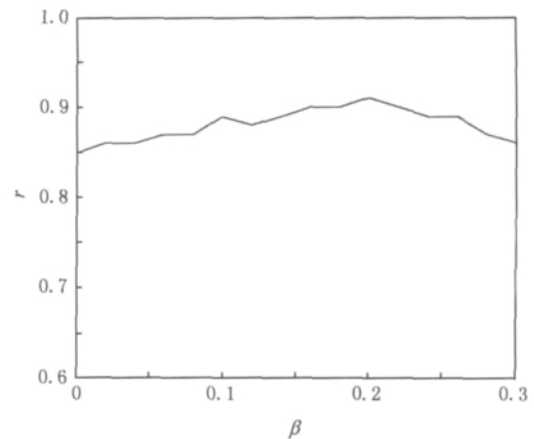


图 4 参数  $\beta$  对识别率的影响

图 4 表明,参数  $\beta$  对识别率的影响不是很大,识别率稳定在 85% 以上.这说明测试图像的显著区域的提取效果比较好,机场总能位于第一显著区而被正确识别.当  $\beta$  较小时,大多数图像的候选区域的特征比都比  $\beta$  大,基本上起不到筛选作用;而  $\beta$  较大时,只有少数几幅图的某些候选区域会有比  $\beta$  大的特征比,作用不明显.上述 2 种情况的识别率主要取决于候选区域的显著性程度,故  $\beta=0$  和  $\beta=0.3$  时具有相近的识别率.可以看到,在  $\beta=0.2$  时,识别率达到最大值 91%.这说明测试数据集中有一些图像的机场不是最显著的,但它拥有较高的特征比,可在按显著性次序识别之前被提前识别出来.

### 5.2 识别结果

在本文方法中,由输入到显著图、候选区域、最终的识别区域一共 4 个主要过程,每个过程都有一张

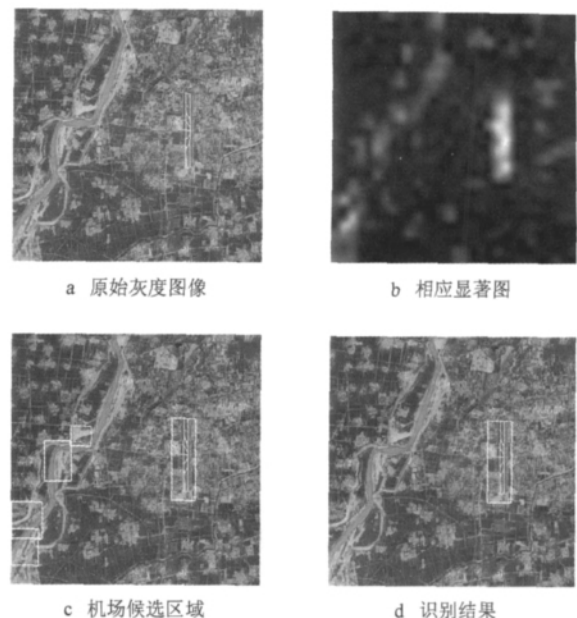


图 5 逐步实验结果示例:中国济南机场

对应的结果图.如图 5 所示,根据前面的分析,这里选取实验参数为  $\alpha=0.35$  和  $\beta=0.2$ .从图 5 b 中可以看到,机场区域具有极高的显著性,并被判别为第一候选区域,虽然机场周围河流的存在使得沿河的一些区域也被确定为候选区域,但经过 SIFT 特征的识别,可以很好地把这些候选区域与机场区域区分开来.

一些其他机场检测结果如图 6 所示.图 6 a 所示的机场被云层大面积覆盖,但仍旧能够从中提取出机场区域,机场区域露出一条白直线起到了作用,无论是在方向通道、Hough 通道都会有一定的滤波

输出,从而使其更显著.图 6 c 所示机场地处山区,山脉、山脊会给机场检测带来干扰,它们也是显著的,能够提取出很多 SIFT 特征,但这些特征中具备机场特征的不多,导致它的特征比明显低于机场区域的特征比,因此没有对识别结果造成影响.图 6 d 所示的机场位于海边,它的主要干扰源来自亮度较高且细长的海岸线,其与机场的跑道有相似之处,但因为缺少交叉点,造成其特征比低于机场区域的特征比,同样避免了误识的发生.从图 6 中可以看到,本文方法在不同的复杂背景下均可以排除干扰,准确地定位出机场区域.

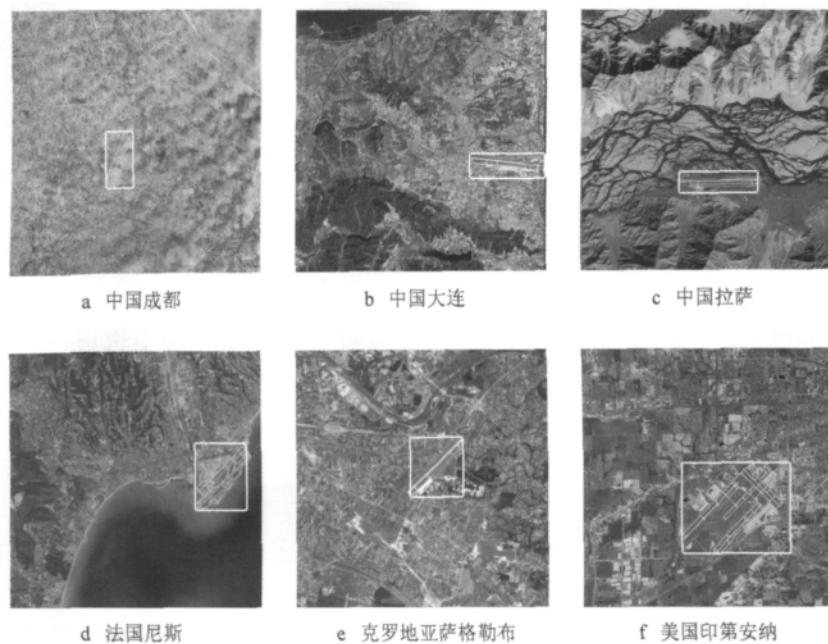


图 6 一些机场的识别结果

另外,图 7 所示为 2 个虚警的情况.可以看到,笔直的或者交叉的道路会对机场存在性的判断产生干扰,这也是仅使用 Hough 变换来筛选的一个弊端.当图像中存在一些较长且具有一定宽度的较亮的直线,可能会出现误判,且后续的 SIFT 特征识别也会因其与跑道的某些相似性而错误地识别目标.

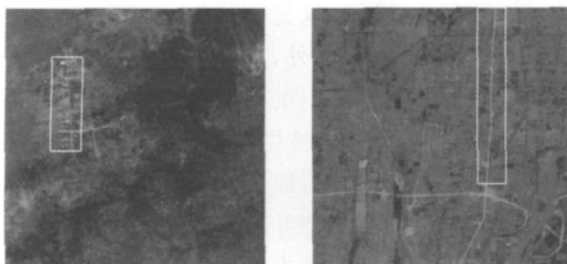


图 7 2 个虚警情况示例

### 5.3 方法比较

将本文方法与文献[3,5]方法进行比较,部分结果如图 8 所示.文献[3]是基于边缘提取的算法的一个代表,它对输入图像进行边缘检测,然后去除小的、弯的边缘并进行 Hough 变换求取长直线,再将直线附近区域的纹理代入 SVM 识别.而文献[5]方法则是区域分割类算法的代表,它于 2011 年提出,具有较满意的识别结果.它首先对待识别图像进行区域分割;然后对整幅图像提取 SIFT 特征并据其密度来聚群;之后,将每个群的位置与区域分割的结果相对应得到若干候选区域;最后在候选区域上提取诸如均值、方差、各阶矩等统计特征,并代入到 SVM 中识别以确定机场区域.文献[5]中的参数选取为  $R=120, S=4$ ,其中  $R$  为 SIFT 空间聚群的最大半径, $S$  为群内最少的 SIFT 特征数,这里根据实验

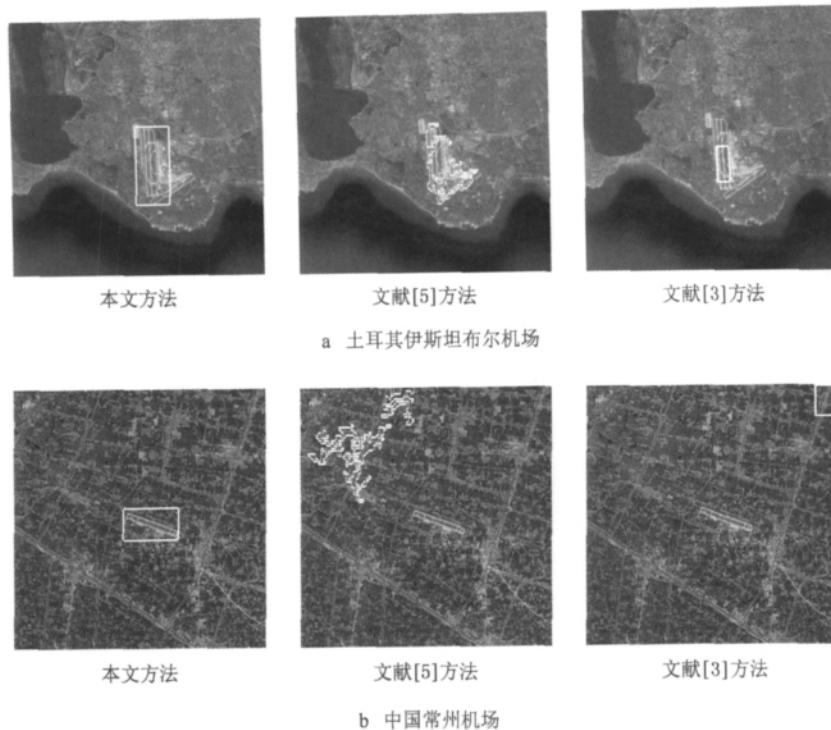


图 8 3 种方法识别结果对比示例

数据调整其至最佳, 选取  $R=80, S=3$ .

图 8 所示为一组采用 3 种方法都识别出来的和一组只有本文方法识别出来的例子. 对于那些具有明显机场特征的、干扰少的机场区域, 3 种方法都可以识别出来, 一旦机场区域机场类特征点不够多, 识别效果就受到影响. 现将 3 种方法对所有测试图像的识别率、识别虚警率和平均运行时间进行计算, 其中后 2 种方法都使用了 30 幅图像作为训练样本, 以保证 SVM 的分类效果. 设不含机场的为  $m$  幅, 其中被误识别出机场的为  $y$  幅, 那么识别虚警率  $R_{FP}$  计算公式为

$$R_{FP} = \frac{y}{m} \times 100\% \quad (2)$$

3 种方法的计算结果如表 1 所示.

表 1 3 种方法识别率、识别虚警率和运行时间的比较

方法	$r/\%$	$R_{FP}/\%$	平均运行时间/s
文献[5]	66	28	>600
文献[3]	65	18	2.43
本文	91	10	2.59

从表 1 可以看出, 本文方法在识别率、识别虚警率方面具有明显的优势, 时间上与文献[3]相接近. 文献[5]存在的主要问题以及本文方法和它的主要差别在于: 1) 文献[5]依靠 SIFT 特征密度来确定候选区, 但有些机场上能提取到的 SIFT 特征点数很

有限, 不能很好地覆盖机场, 而且在 SIFT 空间聚群时, 群的大小、半径也是因机场大小而异的, 找不到一个固定的参数; 而本文方法使用注意力选择机制来提取候选区, 并不涉及特征点的分布情况, 且对机场的大小差异不敏感. 2) 因为实际机场有大有小, 文献[5]在图像分割时阈值不好选取, 并且从单个像素合并生长的方法十分耗时; 而本文方法不是逐像素分析的, 速度快且对机场差异适应性强. 3) 也是文献[5]最大的问题, 就是其在 SVM 识别时所选取的统计特征不具备良好的抗旋转性和抗缩放性. 它所使用的统计特征中均值、方差代表性不强, 而 Zernike 矩<sup>[18]</sup>只用了第一阶, 这是远远不够的, 一般要到 4, 5 阶矩才有较好的描述性, 但是高阶 Zernike 矩计算速度很慢; 本文则使用 SIFT 特征作为识别依据, 因为其有较好的仿射不变性且提取速度快, 故能带来较好的识别效果和较高的识别效率. 文献[3]在速度上是不错的, 但其对水平、垂直方向上的机场的识别远远好于倾斜的机场. 这是因为影响该算法因素除了复杂背景的直线干扰外, 还有它在四邻域找长直线, 那么如果机场是倾斜的方向, 就可能因为垂直、水平方向的边缘过短而被切割成好几段. 也就是说, 文献[3]方法必须事先将原始图像旋转至机场方向才行. 而注意力选择模型并不会因为方向的差异导致检测效果的下降, 因此本文方法可以将各个方向的机场都检测出来.

接收者操作特性 (receiver operating characteristic, ROC) 曲线常被用来衡量检测效率,其横坐标是检测虚警率(false positive rate, FPR),纵坐标为检测正确率(true positive rate, TPR). 这里的检测正确率与前面的识别率不同,它针对的是有没有目标的问题,而不关心目标的位置是否正确,FPR 也是相对 TPR 而言. 设预处理过程中实际存在机场且被判断为存在机场的图像数为  $x_1$ ,而实际不存在机场却被判断为存在机场的图像数为  $y_1$ ,则 FPR 和 TPR 的计算方式分别为

$$R_{FP} = \frac{y_1}{m} \times 100\% , R_{TP} = \frac{x_1}{n} \times 100\% ;$$

其中,  $m, n$  如式(1)(2)定义. 将不同判断标准下的 FPR 和 TPR 绘制成一条曲线,曲线下面积越大,则检测效率越高. 3 种方法根据各自判别标准变化的 ROC 曲线绘制如图 9 所示.

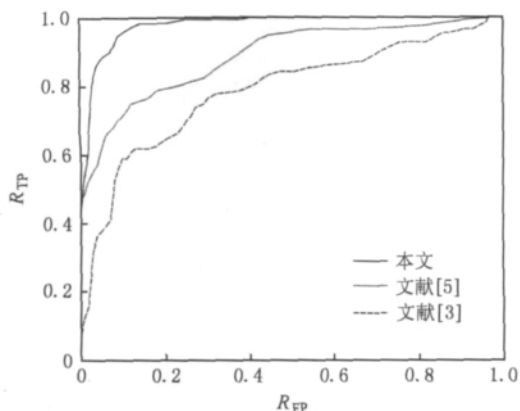


图 9 3 种方法各自的 ROC 曲线

从图 9 中可以看到,本文方法的 ROC 面积明显大于其他 2 种,在较低的虚警率下(比如 10%)已经拥有 90% 以上的 TPR. 而文献[5]方法相对文献[3]要好一些,这是因为文献[5]通过同时具有机场和非机场特征的 SIFT 特征密度来判断否有目标存在,而在很多没有机场的图像上能提取道的 SIFT 特征点数目很少,具有机场特征的就更少了,所以在相同检测率下,虚警率要低一些.

在遥感图像的拍摄中,加性噪声的存在是不可避免的,因此本文将对 3 种方法在人为加入高斯白噪声下的识别率进行测试. 我们在含有机场的 100 幅图像上依次添加信噪比从 0~30 dB 变化的高斯白噪声,生成 31 组含噪声的图像,然后分别运用的 3 种方法对这 31 组图像进行识别,并计算出不同噪声下各自的识别率,绘制出识别率曲线如图 10 所示.

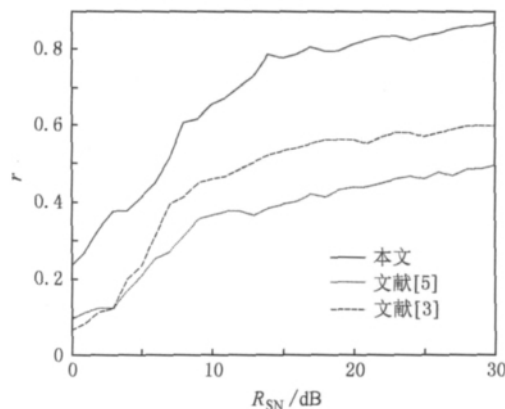


图 10 本文方法识别率随输入信噪比变化的曲线

从图 10 可以看出,3 种方法在 10 dB 信噪比  $R_{SN}$  之前识别率快速上升,10 dB 以后上升幅度逐渐减小. 本文方法在 15 dB 以后趋于稳定,在 30 dB 处已经接近无噪图像的识别率 91%,拥有较好的抗噪性能;而文献[5]方法由于噪声对区域分割、纹理识别的影响导致大噪声时识别率低,识别率随噪声减小上升速度缓慢,直到 30 dB 时才能到 50% 左右的识别率. 文献[3]方法在 30 dB 时已经接近 60% 的识别率,它之所以优于文献[5]方法,是因为在噪声不是很大时,噪声点对 Hough 变换寻找长直线的影响很小,偶尔的断点或者多余的点都不会影响最终的判断,而影响文献[3]方法识别率最主要的因素还是在纹理识别上,散落的噪声点会影响统计特性.

## 6 总 结

本文将注意力选择机制运用到遥感图像机场检测中以提取显著性区域,克服了以往对图像逐像素进行分析的方法的缺点,提高了机场目标的自动检测的识别率和效率. 首先利用注意力选择计算模型分析待测图像的显著性,然后根据显著程度提取目标的候选区域,并在候选区域上提取 SIFT 特征,最后把提取的特征输入到用训练数据训练好的 HDR 树中进行特征识别来确定出机场区域. 实验结果表明,与传统的区域分割、边缘提取的方法相比,本文方法具有更高的识别率、更低的识别虚警率,同时还具有更好的抗噪声能力、更快的运算速度,非常适用于复杂背景下的实时目标检测,对于实际应用具有较大的意义. 如何进一步降低周围景物对机场检测的干扰,并让本文方法对较暗的机场也能拥有同样好的检测效果,是我们今后要研究的一个内容.



## 参考文献 (References):

- [1] Wang W, Liu L, Hu C B, *et al.* Airport detection in SAR image based on perceptual organization [C] //Proceedings of International Workshop on Multi-Platform/Multi-sensor Remote Sensing and Mapping. Los Alamitos: IEEE Computer Society Press, 2011: 1-5
- [2] Duda R O, Hart P E. Use of the Hough transformation to detect lines and curves in pictures [J]. *Communications of the ACM*, 1972, 15(1): 11-15
- [3] Qu Y Y, Li C H, Zheng N N. Airport detection base on support vector machine from a single image [C] //Proceedings of 5th International Conference on Information, Communications and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2005: 546-549
- [4] Pi Y M, Fan L H, Yang X B. Airport detection and runway recognition in SAR images [C] //Proceedings of IEEE International Geoscience and Remote Sensing Symposium. Los Alamitos: IEEE Computer Society Press, 2003, 6: 4007-4009
- [5] Tao C, Tan Y H, Cai H J, *et al.* Airport detection from large IKONOS images using clustered SIFT keypoints and region information [J]. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(1): 128-132
- [6] Liu D H, He L H, Carin L. Airport detection in large aerial optical imagery [C] //Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2004, 5: 761-764
- [7] Zhou Nuo, Chen Wei, Zhao Hui, *et al.* Airport detection algorithm in large area satellite borne SAR images [J]. *Systems Engineering and Electronics*, 2009, 31(8): 761-764 (in Chinese)  
(周 诺, 陈 炜, 赵 晖, 等. 大幅面星载 SAR 图像中机场检测算法[J]. *系统工程与电子技术*, 2009, 31(8): 761-764)
- [8] Desimone R, Duncan J. Neural mechanisms of selective visual attention [J]. *Annual Reviews in Neuroscience*, 1995, 18: 193-222
- [9] Itti L, Koch C, Niebur E, *et al.* A model of saliency-based visual attention for rapid scene analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11): 1254-1259
- [10] Walther D, Koch C. Modeling attention to salient proto-objects[J]. *Neural Networks*, 2006, 19(9): 1395-1407
- [11] Hou X D, Zhang L Q. Saliency detection: a spectral residual approach [C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2007: 1-8
- [12] Guo C L, Ma Q, Zhang L M. Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform [C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2008: 1-8
- [13] Guo C L, Zhang L M. A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression [J]. *IEEE Transactions on Image Processing*, 2010, 19(1): 185-198
- [14] Ding Zhenghu, Yu Ying, Wang Bin, *et al.* Visual attention-based ship detection in multispectral imagery [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2011, 23(3): 419-425 (in Chinese)  
(丁正虎, 余 映, 王 斌, 等. 选择性视觉注意机制下的多光谱图像舰船检测[J]. *计算机辅助设计与图形学学报*, 2011, 23(3): 419-425)
- [15] Harel J, Koch C, Perona P. Graph-based visual saliency [C] //Proceedings of Advances in Neural Information Processing Systems. Cambridge: The MIT Press, 2007: 545-552
- [16] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [17] Hwang W S, Weng J. Hierarchical discriminant regression [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1277-1293
- [18] Khotanzad A, Hong Y H. Invariant image recognition by Zernike moments [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(5): 489-497