

基于支持向量机的显著性建筑物检测

曲延云^{1,2} 郑南宁¹ 李翠华² 袁泽剑¹ 叶聪颖²

¹(西安交通大学人工智能与机器人研究所 西安 710049)

²(厦门大学计算机科学系 厦门 361005)

(yyqu@xmu.edu.cn)

Salient Building Detection Based on SVM

Qu Yanyun^{1,2}, Zheng Nanning¹, Li Cuihua², Yuan Zejian¹, and Ye Congying²

¹(Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049)

²(Department of Computer Science, Xiamen University, Xiamen 361005)

Abstract This paper focuses on detecting salient buildings in a scenery image. A method based on bottom-up attention mechanism is proposed to detect salient buildings. Firstly, Haar wavelet decomposition is used to obtain the enhanced image which is the sum of the square of LH sub-image and HL sub-image. Secondly, the enhanced image is projected in the vertical direction to obtain the projection profile, and building candidates are separated from the background based on multi-level thresholding. Thirdly, the structure statistic features of buildings are extracted based on Sobel operator. The feature vector is formed by the number of long horizontal edges and that of vertical edges. Finally, linear support vector machines are used to classify buildings and the others. The proposed approach has been experimented on many real-world images with promising results.

Key words building detection; bottom-up attention mechanism; Haar wavelet decomposition; SVM

摘要 提出了一种针对自然图像中显著性建筑物的检测方法。首先,采用自底向上的注意力机制,对图像进行 Haar 小波分解,对得到的 HL , LH 分量进行平方求和,得到增强图像,然后对该增强图像在垂直方向上进行侧投影,基于得到的投影曲线进行多层阈值分割,找到显著性建筑物候选区域。进而,利用 Sobel 算子进行水平边缘与垂直边缘的检测,并统计较长的水平边缘与垂直边缘的数目,组成特征向量。最后利用线性支持向量机对特征进行分类。实验证明了所提算法的有效性。

关键词 建筑物检测;自底向上的注意力机制;Haar 小波分解;支持向量机

中图法分类号 TP391.4

建筑物检测在自动导航、虚拟现实、图像理解、图像检索等方面都起着重要作用。本文主要研究自然景物图像中显著性建筑物的检测。我们用显著性建筑物表示那些在图像中较其他对象更能引起人们视觉注意的建筑物。比如在天空衬托下的高大建筑、在建筑群中的突出建筑等。建筑物的检测所遇到的困难有:把建筑物候选区域从背景中分割出来;对建筑物进行表示。

建筑物检测是近年来一个活跃的研究领域。有航拍图像的建筑物检测、有地面拍摄的光学图像的建筑物检测。前者已有了一些系统,如美国 Carnegie

Mellon 大学的 Mapslab 小组相继完成了 4 个建筑物提取的系统:BUILD^[1], BUILD + SHAVE^[2], VHBUILD^[3] 和 PIVOT^[4,5]。Southern California 大学对航拍图像建筑物的检测也做了大量工作,见文献[6]。这些系统多采用了屋顶存在性假设,并采用边缘检测、线连接及连接形状分析等图像的底层信息。在得到这些信息的基础上,使用基于几何或统计的模型进行聚类。此外上述系统还利用了光照信息、阴影信息、高度信息或 DEM 信息等。后者的工作多用于图像检索、场景分析。对建筑物结构特征的提取和对建筑物的表示是学者们关注的焦点。文献[7]构建高级

收稿日期:2005-10-09;修回日期:2006-08-22

基金项目:国家自然科学基金项目(60635050,60405004);国家自然科学基金创新研究群体基金项目(60021302);

语意结构,进行场景分析,检测建筑物.文献[8]利用基于感知的方法检测建筑物.文献[9]利用基于因果多尺度随机场的方法对自然图像中建筑物进行检测,上面提到的工作都结合了图像的底层信息,如角点、直线统计等.

本文的任务是针对自然景物图像中的显著性建筑物进行检测,与航拍图像的建筑物检测的研究对象不同.与文献[8]的区别是后者是建筑物图像,而本文研究的是自然景物图像,其对象除了建筑物外还有其他的对象,如山、树、湖泊等.本文与文献[7,9]都是研究复杂背景下的自然景物图像,通过自底向上的注意力机制,把 Haar 小波分解得到的 HL , LH 分量进行平方求和,得到增强图像.然后对该增强图像进行垂直方向投影,对得到的曲线进行多阈值划分,找到显著性建筑物候选区域.考虑到在自然场景图像中影响显著性建筑物检测的常常是高大的树冠、山峰等,我们把这些干扰对象作为负样本,把建筑物作为正样本.根据我们的经验,建筑物一般有规则的几何特征,如平行的直线簇、矩形框等.而负样本一般不具备这一特征,因此我们提取候选区域的基于 Sobel 算子的结构统计特征.特征矢量是由长的水平边缘的数目与垂直边缘的数目组成.利用支持向量机对特征矢量进行分类.

1 采用自底向上的注意力机制提取感兴趣区域

把建筑候选区域从图像中提取出来是我们遇到的一个难题.其原因如下:图像的成像比例无法确定,建筑物成像的大小不定;建筑物颜色受光照影响较大;图像中建筑物可能部分被遮挡;建筑物周围有很多干扰,如树木、建筑群等,把建筑物分割出来较困难.

本文采用自底向上的注意力选择机制来解决该问题.根据我们的经验,高大的物体与背景对比度

强的物体较容易引起我们视觉的注意.而提取建筑物候选区域过程好比我们对环境的感知,那些对光线反射强的物体最先被眼睛感知到,那么我们就知道了有东西存在于那个地方,然后才看该物体到底是什么.由于显著性建筑物多以天空为衬托,通过对图像的处理,使天空和建筑区域分别成为背景和前景,以背景反衬前景.然后通过对天空边界的搜索,把前景从背景中分离出来.传统的图像分割方法^[10]一般分为基于边缘的方法和基于区域的方法,分割出的区域一般都是均质的,如区域具有一致的颜色、纹理等,但是建筑区域一般不具有这样的性质,很多建筑屋顶与墙体颜色不一致,墙体的纹理、颜色一般也不一致,采用常用的分割方法会把建筑物过分割成不同的区域,造成只见树叶不见森林的后果,无法得到整体信息.本文利用小波分解的信息对可能存在建筑物的区域进行增强.

对图像 $I(x, y)$ 进行 Haar 小波分解可以得到 4 个子图像: LL 子图像、 LH 子图像、 HL 子图像、 HH 子图像. LL 子图像是在水平垂直两个方向上进行低通滤波得到的图像,相当于得到原图像的一个光滑图像; LH 子图像是在水平方向进行低通滤波、垂直方向进行高通滤波得到的图像,该图像对水平边缘有增强作用; HL 子图像是在水平方向进行高通滤波、垂直方向进行低通滤波得到的图像,该图像对垂直边缘有增强作用; HH 子图像是在水平垂直两个方向都进行高通滤波得到的图像.我们仅对 LH 和 HL 子图像感兴趣.考虑到分割所消耗的时间,我们采用 2 层 Haar 小波分解,令在第 2 层上得到 LL 子图像为 $A_2(x, y)$,在第 2 层上的 LH 子图像为 $H_2(x, y)$,在第 2 层上的 HL 子图像为 $V_2(x, y)$,按式 (1) 构造一新图像:

$$B(x, y) = H_2(x, y)^2 + V_2(x, y)^2. \quad (1)$$

式 (1) 的作用是对存在水平边缘与垂直边缘的区域进行增强.采用该做法可以得到较好的天空与非天空分割图,如图 1 所示:

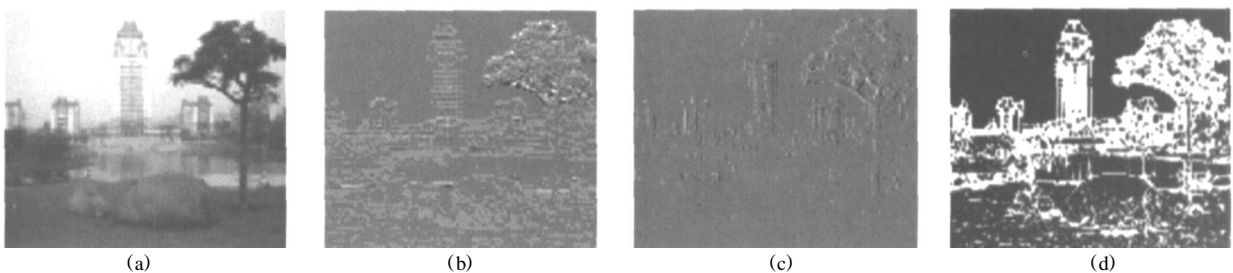


Fig. 1 The sub-images after Haar wavelet decomposition and the enhanced image. (a) LL sub-image; (b) LH sub-image; (c) HL sub-image; and (d) The enhanced image.

图 1 基于小波分解的图像增强处理效果图. (a) LL 分量图; (b) LH 分量图; (c) HL 分量图; (d) 增强后的效果图

在得到增强图像后,我们对它进行垂直方向投影,找到天空的边界,然后对投影曲线进行阈值分割,找到建筑候选区域,其算法步骤如下:

1) 对处理后的图像 $B(x, y)$ 进行去噪,使得天空部分尽量少的受到噪声的干扰;

2) 对图像 $B(x, y) \in R^{m \times n}$ 进行逐列投影,找到第 1 个非零点的位置:

$$S(y) = \min_x \{ x \mid B(x, j) > 0, j = y, 1 \leq y \leq m \}; \tag{2}$$

3) 对投影轮廓进行多阈值分割,找到曲线各凸包的左右边界;

4) 利用各凸包的左右边界,并结合投影曲线的值,划分出建筑候选区域

算法第 3 步是完成提取建筑物候选区域的关键

步骤,我们先采用一较大阈值截取投影曲线,得到孤立的凸包,通过检查点与点之间连续性,判断凸包的位置及个数.图 2(a)是通过图 1 的增强图像进行垂直方向投影得到的投影曲线,从图中可以看到 4 个明显的凸包;图 2(b)为采用投影曲线最大值的一半作为阈值,进行阈值划分后得到的曲线 $S_1(y)$,与图 2(a) 比较发现少了两个相对不明显的凸包.然后将已确定凸包的投影值降低到最低的天空边界值.接着降低阈值,重复上面的步骤,直到找不到大凸包为止.图 2(c)为降低(a)中两个最显著凸包的值后进行第 2 次阈值划分得到的曲线 $S_2(y)$.图 3 为建筑物分割图像,其中黑色矩形框部分为建筑候选区域.之所以采用该方法,是因为阈值的大小很难一次选定,如阈值选得过大造成建筑候选区域的丢失,如阈值选的过小造成阈值分割的困难

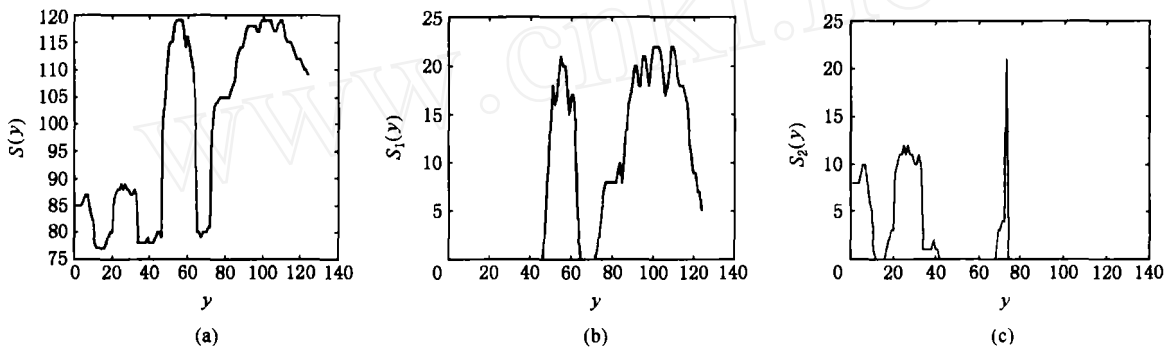


Fig. 2 Vertical Projection profile and thresholding. (a) The vertical projection profile of Fig. 1 (d); (b) The projection profile after the first thresholding; and (c) The projection profile after the second thresholding.

图 2 垂直投影曲线及其阈值分割. (a) 图 1(d) 的垂直方向投影曲线; (b) 第 1 次阈值分割后得到的投影曲线; (c) 第 2 次阈值分割后得到的曲线

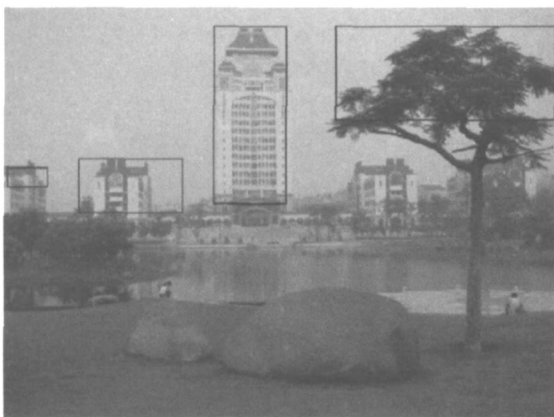


Fig. 3 Extraction of ROI. 图 3 提取建筑物候选区域

2 建筑物结构特征描述

建筑物的特征选择在建筑物检测中起着重要作

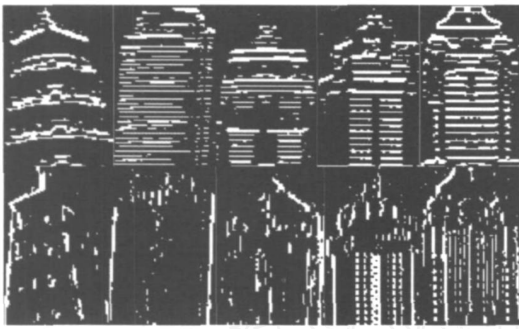
用. 在自然场景图像中,对建筑物检测造成干扰的往往是高大的树冠、山峰等,与这些干扰物比较,建筑物一般都有规则的几何形状,其墙体常常具有横的或竖的直线,而干扰物多不具备这一特征,例如树冠比较杂乱,山峰具有较一致的纹理等.我们利用 Sobel 算子分别计算建筑物的水平方向结构特征与垂直方向结构特征.

建筑物图像用 $I \in R^{m \times n}$ 表示,其与 Sobel 算子 $H = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$ 卷积得到二值图像 $S_h \in R^{m \times n}$, 与 H^T 卷积得到二值图像 $S_v \in R^{m \times n}$. 对图像 S_h, S_v 分别计算其水平投影和垂直投影:

$$P_r(i) = \sum_{j=1}^n S_h(i, j), 1 \leq i \leq m, \tag{3}$$

$$P_c(j) = \sum_{i=1}^m S_v(i, j), 1 \leq j \leq n, \tag{4}$$

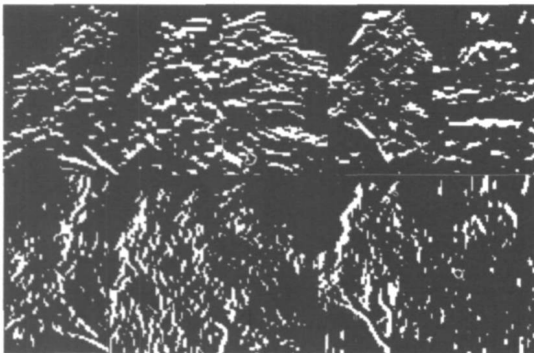
其中, $0 \leq P_r(i) \leq n, 0 \leq P_c(j) \leq m$. 令 h_m 表示数组 $P_r(i)$ 中值大于 $\frac{n}{3}$ 的个数, v_m 表示数组 $P_c(j)$ 中值大于 $\frac{m}{3}$ 的个数. 定义建筑物几何统计特征为 $f_m = [h_m, v_m]$, 该特征表示了感兴趣区域的规则性. 图 4 显示了建筑物的水平边缘检测与垂直边缘检测的例子. 图 5 显示了树冠边缘检测的例子. 比较图 4、图 5 可以看出, 建筑物具有较规则的几何特征, 而树冠则不具备这一特征.



The top row are the detected horizontal edge images;
The bottom row are the detected vertical edge images.

Fig. 4 Structure features of buildings based on Sobel operator.

图 4 基于 Sobel 算子的建筑物结构特征



The top row are the detected horizontal edge images;
The bottom row are the detected vertical edge images.

Fig. 5 Structure features of negative samples based on Sobel operator.

图 5 基于 Sobel 算子的非建筑物结构特征

3 基于支持向量机的建筑物判别

我们采集了 40 幅建筑物图像, 并把它们做镜像变换共得到 80 幅正样本图像. 考虑到干扰建筑物检测的多来自树木、山的纹理等, 搜集了 64 幅非目标图像作为负样本. 部分正样本和负样本见图 6、图 7. 计

算训练样本集的几何统计特征, 用 $\{x_i\}_{i=1}^N, x_i \in \mathbb{R}^2$ 表示它们, 对应的标号用 $\{y_i\}_{i=1}^N, y_i \in \{-1, 1\}$ 表示. 训练样本集的几何统计特征分布如图 8 所示, 从图中可以看出, 正样本特征矢量与负样本特征矢量之间可以用线性函数分开, 因此我们采用线性支持向量机函数来训练样本. 用于判断建筑物的最优分类函数为 $f(x) = \text{sgn}(w^* \cdot x + b)$, (5)

$$\text{其中, } \text{sgn}(x) = \begin{cases} +1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

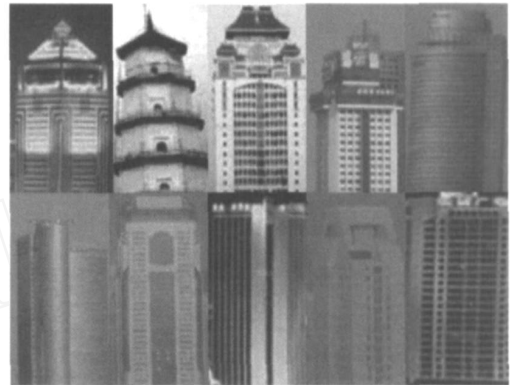


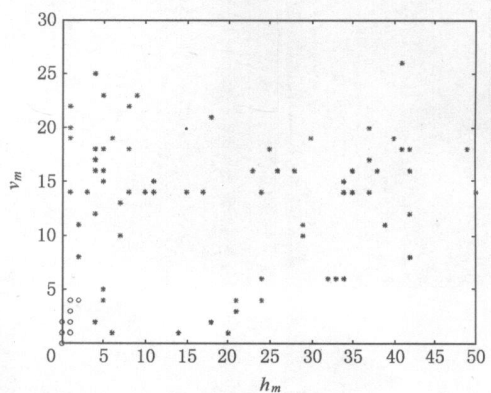
Fig. 6 Part of building samples.

图 6 部分建筑物样本



Fig. 7 Part of negative samples.

图 7 部分非建筑物样本



“*” stands for features of building,
“o” stands for negative samples’ features

Fig. 8 Feature distribution of training samples.

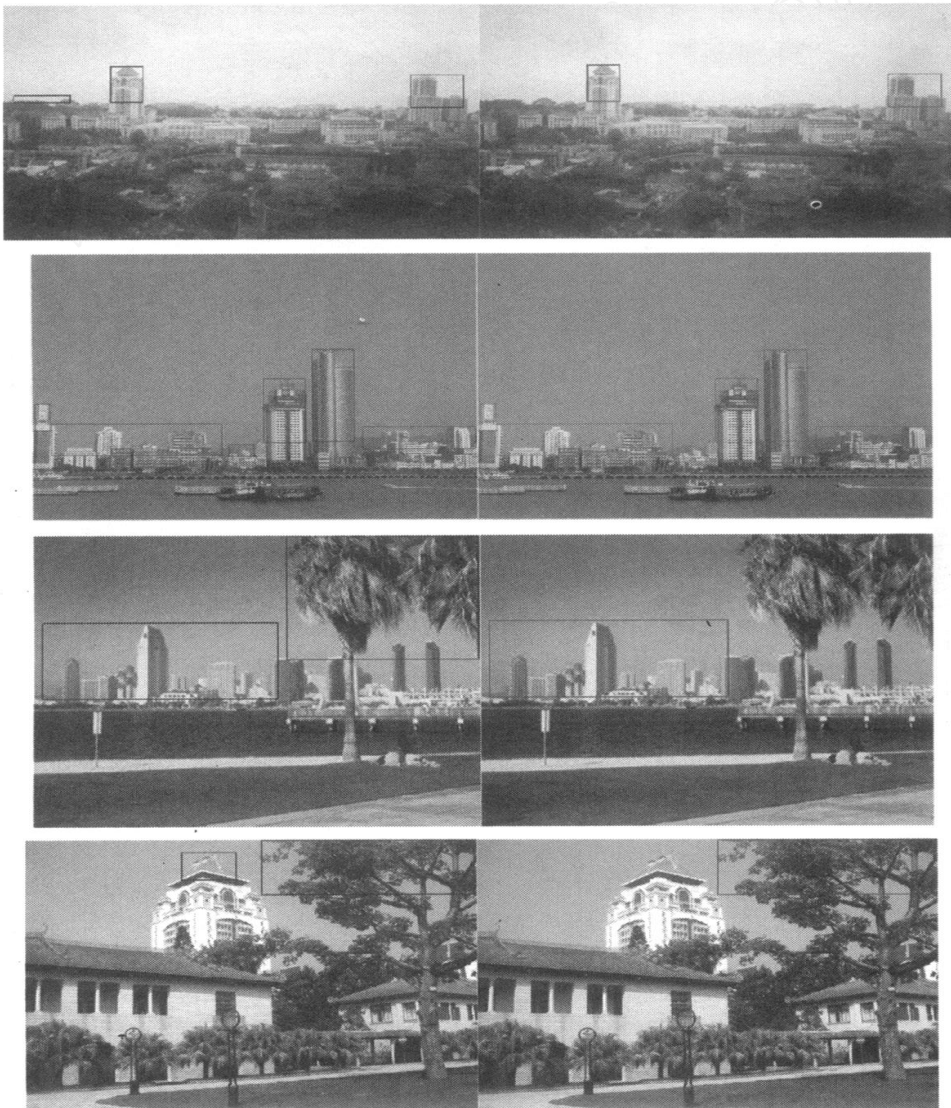
图 8 训练集中样本统计特征的分布

4 实验结果

我们对搜集到的几十幅自然景物图像进行实验,在 Matlab6.5 上执行.对图 1 的原图像进行建筑物检测的结果见图 9,算法最后只找到一个建筑物区域.与图 3 比较,少了两个较小的建筑候选区域.我们选择了部分自然场景图像的建筑物检测的实验结果见图 10.比较建筑物分割图像与最终的建筑物检测结果图像,我们发现部分建筑物候选区域在最后的結果中被去除,在图 9 的检测中是建筑物的候选区域最终没有被判別为建筑物,属于漏检,在图 10 的第 1 行图像中,左侧有一小块建筑候选区域为



Fig. 9 Detection result for the original image in Fig. 1.
图 9 对图 1 的原图像进行建筑物检测的结果



On the left, extraction of ROI which are indicated by black rectangles.
On the right, building detection results, and the detected buildings are indicated by red rectangles.

Fig. 10 Experimental Results.
图 10 部分实验结果

山峰,最终通过 SVM 进行正确的去除. 图 10 第 4 行图像没有正确检测到建筑物,为误检情况. 建筑物区域没有被检测到有如下原因:一是由于遮挡,建筑物没有被分割出来,或者说部分遮挡物占据图像的显著部分. 二是由于我们搜集的样本集合较小,我们对建筑物统计特征的选择来自于我们的观察与经验,我们认为建筑物具有规则的几何特征,墙体具有水平或垂直的平行线结构是建筑物普遍存在的现象,但是我们所提及的特征仅仅是建筑物特征的一方面,并不是建筑物检测的充分而必要的条件,这也限制了我们的检测效果,关于建筑物特征的描述是建筑物检测的难点之一.

5 结 论

本文对自然场景图像中显著性建筑物进行检测. 首先采用自底向上的注意力机制,利用 Haar 小波分解把建筑物候选区域从背景中提取出来. 考虑到建筑物墙体具有几何规则性的特点,我们采用 Sobel 算子提取建筑物的几何统计信息,组合成特征矢量,作为建筑物的结构特征,利用支持向量基函数判断建筑物. 实验证明了本文所提算法的有效性. 如何选择更好的建筑特征对建筑物进行表示是我们下一步的工作.

参 考 文 献

- [1] D M McKeown. Toward automatic cartographic feature extraction. In: Mapping and Spatial Modelling for Navigation [C]. NATO ASI Series. Berlin: Springer-Verlag, 1990, F65: 149-180
- [2] R B Irvin, D M McKeown. Methods for exploiting the relationship between buildings and their shadows in aerial imagery [J]. IEEE Trans on Systems, Man, and Cybernetics, 1989, 19(6): 1564-1575
- [3] J C McClone, J A Shufelt. Projective and object space geometry for monocular building extraction [C]. IEEE Conf on Computer Vision and Pattern Recognition, Washington, USA, 1994
- [4] J A Shufelt. Exploiting photogrammetric methods for building extraction in aerial images [J]. Int J Archives of Photogrammetry and Remote Sensing, 1996, 31(B6/S): 74-79
- [5] J A Shufelt. Projective geometry and photometry for object detection and delineation [R]. Carnegie Mellon University, Tech Rep: CMU-CS-96-164, 1996
- [6] C Lin, R Nevatia. Building detection and description from a single intensity image [J]. Computer Vision and Image Understanding, 1998, 72(2): 101-121
- [7] Yunqi Song, Aidong Zhang. Analyzing scenery images by monotonic tree [J]. ACM Multimedia Systems Journal, 2002, 8(6): 495-511

- [8] A Iqbal, J K Aggarwal. Applying perceptual grouping to content-based image retrieval: Building images [C]. IEEE Int J Conf on CVPR, Fort Collins, Colorado, 1999
- [9] S Kumar, M Hebert. Man-made structure detection in natural images using a causal multiscale random field [C]. IEEE Int J Conf on CVPR, Madison, USA, 2003
- [10] R C Gonzalez, R E Woods, S L Eddins. Digital Image Processing Using Matlab [M]. Englewood Cliffs, NJ: Prentice-Hall, 2004



Qu Yanyun, born in 1972. She is a lecturer of Xiamen University and a Ph. D. candidate of the Institute of Artificial Intelligence and Robotics at Xi'an Jiaotong University. Her research interests include image processing, pattern recognition, machine learning and so

on.

曲延云, 1972 年生, 博士研究生, 讲师, 主要研究方向为图像处理、模式识别、机器学习等.



Zheng Nanning, born in 1952. He is currently a professor, Ph. D. supervisor and the director of the Institute of Artificial Intelligence and Robotics at Xi'an Jiaotong University. His research interests include computer vision, pattern recognition, computational intelligence, image processing, and hardware implementation of intelligent systems. Since 2000, he has been China representative of the Governing Board of the International Association for Pattern Recognition. He presently serves as executive editor of the Chinese Science Bulletin. He became a member of the Chinese Academy of Engineering in 1999. He is a senior member of IEEE.

郑南宁, 1952 年生, 教授, 博士生导师, 主要研究方向为计算机视觉、模式识别、计算智能、图像处理、智能系统的硬件实现.



Li Cuihua, born in 1960. He is currently a professor, Ph. D. supervisor and the director of Computer Science Department at Xiamen University, senior member of China Computer Federation. His research interests include video processing and image processing.

李翠华, 1960 年生, 教授, 博士生导师, 中国计算机学会高级会员, 主要研究主向为视频与图像处理.



Yuan Zejian, born in 1971. He is currently a lecturer of the Institute of Artificial Intelligence and Robotics at Xi'an Jiaotong University. His research interests include pattern recognition and machine learning.

袁泽剑, 1971 年生, 讲师, 主要研究方向为模式识别、机器学习等.



Ye Congying, born in 1979. He received master degree from the Computer Science Department at Xiamen University. His research interests include image processing and pattern recognition.

叶聪颖,1979年生,硕士,主要研究方向为图像处理与模式识别。

Research Background

Man-made structure detection is important in many applications such as automatic navigation, scene understanding, virtual reality, image understanding etc. In this paper, we think it as a universal phenomenon that buildings have long parallel horizontal edges and vertical edges. So we firstly use Haar wavelet decomposition to obtain the enhanced image. Secondly, we project the enhanced image in the vertical direction and obtained projection profile. We use multi-level threshold to extract the building candidates. We count the number of horizontal edges and that of vertical edges which form the feature vector. We use linear SVM to classify the building candidates. Our work is supported by the Chinese National Natural Science Foundation No. 60635050, No. 60405004 and the Foundation of Creative Research Group Programming No. 60021302.

2007 全国理论计算机科学学术年会 征文通知

由中国计算机学会理论计算机科学专业委员会主办、广西大学计算机与电子信息学院承办的“2007 全国理论计算机科学学术年会”将于 2007 年 11 月在广西南宁市召开。本次年会录用的论文将由《计算机研究与发展》(增刊)和《计算机工程与科学》(正刊)发表,欢迎大家积极投稿。有关征文要求通知如下:

1) 为进一步提高学术水平,本次年会将采取以下措施: 评选优秀论文并将其推荐给一级学报; 除征集尚未在其他刊物或学术会议上正式发表过的论文外,还将征文扩大到 2006、2007 年中国大陆学者在国际一流学术刊物上已发表的论文(注明出处),对后者经审稿录用后,邀请作者出席年会并作报告(但论文不再收入论文集),根据报告情况择优评选最佳论文并进行奖励; 将从与会代表所作的报告中评选出年会的最佳报告并进行奖励; 鼓励在年会召开期间组织讲座(Tutorial),有意者请与广西大学钟诚教授联系。

2) 稿件请按《计算机研究与发展》、《计算机工程与科学》的论文格式编排,要求用计算机打印,正文采用五号宋体。稿件中的图形、图像要求工整、清晰、紧凑,尺寸要尽量小,图表中文字采用六号宋体。稿件全文不超过 6000 字。标题、作者姓名、作者单位、摘要、关键词五部分要求中、英文齐全。稿件正文部分依次为:1 引言;2 ...;最后是结语。附录放在参考文献之后;参考文献限列已公开发表的,参考文献格式为:

序号 作者. 书名. 出版地:出版社名称,出版年份

序号 作者. 论文题名. 期刊名称,年份,卷号(期号):起止页码

序号 作者. 论文题名. 会议论文集名称,出版地:出版社名称,出版年份,起止页码

务必附上第一作者简历(姓名、性别、出生年月、出生地、职称、学位、研究方向等)、通信地址、邮政编码、联系电话和电子信箱。并注明论文所属领域。来稿一律不退,请自留底稿。

3) 征文范围

程序理论(程序逻辑、程序正确性验证、形式化开发方法等)

计算理论(算法设计与分析、复杂性理论、可计算性理论等)

语言理论(形式语言理论、自动机理论、形式语义学、计算语言学等)

人工智能(知识工程、机器学习、模式识别、机器人、数据挖掘、进化计算等)

逻辑基础(数理逻辑、多值逻辑、模糊逻辑、模态逻辑、直觉主义逻辑、组合逻辑等)

数据理论(演绎数据库、关系数据库、面向对象数据库、并行分布数据库、空间数据库等)

计算机数学(符号计算、数学定理证明、计算几何、组合数学等)

并行算法(分布式并行算法、大规模并行算法、网格算法、量子计算并行算法等)

4) 征文投稿截止日期:2007 年 4 月 1 日;论文录用通知日期:2007 年 5 月 1 日

5) 论文投稿需提交激光打印稿一式 2 份和电子版 WORD 文件,论文投寄地址和电子信箱如下:

(530004)广西南宁市大学东路 100 号 广西大学计算机与电子信息学院 钟诚、李陶深教授收

Email:tcs2007@sina.com

6) 联系人和联系电话及 Email 信箱:

钟诚:0771-3236396,13607819333, chzhong@gxu.edu.cn

李陶深:0771-3236627,13768301390, tshli@gxu.edu.cn