

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **5**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年5月 第17卷 第5期(总第193期)

目次

综述

- 中国图像工程:2011 章毓晋(603)
- 植物叶片表面质感建模与真实感绘制研究进展 田原,赵春江,陆声链,郭新宇(613)

图像处理和编码

- 高位有效位概率算术解码的图像密写改进算法 马丽红,吕先明,高小满(621)
- 滑动平均和改进权重函数的快速非局部平均图像去噪算法 熊波,尹周平(628)
- 提升小波的同态滤波在图像烟雾弱化中的应用 范有臣,李迎春,韩意,张来线(635)

图像分析和识别

- 基于彩色模型的重构标记分水岭分割算法 张桂梅,周明明,马珂(641)
- 均衡化特征匹配的非刚体细胞形态跟踪 陈莹,艾春璐(648)
- 图像斑状特征位置与尺寸的自动检测 王志衡,刘红敏(656)
- 改进极化白化滤波的边缘检测 邓少平,张继贤,李平湘,黄国满(665)
- 联合特征在行人检测中的应用 杨阳,杨静宇(671)
- Gabor 相位特征的人脸光照不变量提取 范春年,张福炎(676)

图像理解和计算机视觉

- 保持几何特征的自适应弹性二次曲线模型 蒋建国,郝世杰,郭艳蓉,詹曙,李鸿(682)
- 局部颜色映射的彩色夜视算法 钱小燕,张天慈,王帮峰,黄圣国(689)
- 结合分支定界法和线性规划的摄像机位姿估计 马文娟(694)

金字塔评分改进主方向模板匹配的实时目标检索 洪朝群,朱建科,李娜,卜佳俊,陈纯(700)

计算机图形学

类曲率在曲线相似性判定中的应用 于昊,赵乃良,陈小雕(707)

虚拟现实与增强现实

人机系统中视域评估的可视化 李倩,吉晓民,林文周(715)

遥感图像处理

小波方向子带偏微分方程遥感图像去噪 王相海,李放,王爽(721)

遥感图像理想均衡化及图像质量定量评价 孟天佑,汪云甲(729)

地理信息技术

点要素扇形缓冲区的设计与应用 侯景伟,孔云峰,张迪,吕可文(740)

中国图象图形学报

刊名题字: 宋 健

月刊(1996年创刊)

第17卷 第5期

2012年5月16日出版

主管单位 中国科学院
主 办 中国科学院遥感应用研究所
 中国图象图形学学会
 北京应用物理与计算数学研究所
主 编 李小文
编辑出版 《中国图象图形学报》编辑出版委员会
 北京9718信箱 邮编 100101
 电子信箱:jig@irsa.ac.cn
 电话:010-68407995 010-82614429
 网 址:www.cjig.cn
印刷装订 北京北林印刷厂
广告经营许可证 京朝工商广字第0346号
总 发 行 北京报刊发行局
订 购 全国各地邮局
国外发行 中国国际图书贸易总公司
 (中国国际书店)
 (北京399信箱 邮编 100044)

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing Application,
 CAS China Society of Image and Graphics
 Institute of Applied Physics and Computational
 Mathematics
Chief editor LI Xiaowen
Editor, Publisher Editorial and Publishing Board
 of Journal of Image and Graphics
 (P. O. Box 9718, Beijing 100101, China)
 E-mail:jig@irsa.ac.cn
Distributed by Beijing Bureau for Distribution of Newspapers
 and Journals
Domestic All Local Post Offices in China
Foreign China International Book Trading Corporation
 (P. O. Box 399, Beijing 100044, China)
Printed by Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00元

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 5 May 2012

Contents

Review

- Image engineering in China: 2011 Zhang Yujin (603)
- Advances in appearance modeling and photorealistic rendering of plant leaves
..... Tian Yuan, Zhao Chunjiang, Lu Shenglian, Guo Xinyu (613)

Image Processing and Coding

- Most significant bits probability arithmetic decoding for improved image steganography
..... Ma Lihong, Lv Xianming, Gao Xiaoman (621)
- Fast non-local means for image de-noising on moving average and modified weight function Xiong Bo, Yin Zhouping (628)
- Weakening of smoke for homomorphic filtering Fan Youchen, Li Yingchun, Han Yi, Zhang Laixian (635)

Image Analysis and Recognition

- Image segmentation algorithm for reconstruction labeling watershed in color space Zhang Guimei, Zhou Mingming, Ma Ke (641)
- Non-rigid cell contour tracking method for balanced feature matching Chen Ying, Ai Chunlu (648)
- Automatically detecting position and size of blob features in images Wang Zhiheng, Liu Hongmin (656)
- Improved polarimetric whitening filter for edge detection Deng Shaoping, Zhang Jixian, Li Pingxiang, Huang Guoman (665)
- Pedestrian detection based on compound feature Yang Yang, Yang Jingyu (671)
- Illumination invariant extraction on Gabor phase Fan Chunnian, Zhang Fuyan (676)

Image Understanding and Computer Vision

- Adaptive geometrical-feature-preserving elastic quadratic wire model
..... Jiang Jianguo, Hao Shijie, Guo Yanrong, Zhan Shu, Li Hong (682)
- Color night vision algorithm based on local color mapping ... Qian Xiaoyan, Zhang Tianci, Wang Bangfeng, Huang Shengguo (689)
- Camera pose estimation using branch and bound method with linear programming Ma Wenjuan (694)
- Real-time object retrieval with dominant orientation template matching improved by pyramid scoring
..... Hong Chaoqun, Zhu Jianke, Li Na, Bu Jiajun, Chen Chun (700)

Computer Graphics

- Quasi-curvature and its application in similarity measurement of curves
..... Yu Hao, Zhao Nailiang, Chen Xiaodiao (707)

Virtual Reality and Augmented Reality

- Visualization of the visual range assessment in man-machine system Li Qian, Ji Xiaomin, Lin Wenzhou (715)

Remote Sensing Image Processing

- Remote sensing image de-noising on partial differential equation in wavelet directional subband
..... Wang Xianghai, Li Fang, Wang Shuang (721)
- Ideal equalization of remote sensing images and quantitative assessment of image quality Meng Tianyou, Wang Yunjia (729)

Geoinformatics

- Design and applications of sector buffers for point feature
..... Hou Jingwei, Kong Yunfeng, Zhang Di, Lv Kewen (740)

中图法分类号: TP391.9 文献标志码: A 文章编号: 1006-8961(2012)05-0700-07

论文引用格式: 洪朝群, 朱建科, 李娜, 卜佳俊, 陈纯. 金字塔评分改进主方向模板匹配的实时目标检索[J]. 中国图象图形学报, 2012, 17(5): 700-706

金字塔评分改进主方向模板匹配的实时目标检索

洪朝群, 朱建科, 李娜, 卜佳俊, 陈纯

浙江大学计算机科学与技术学院, 杭州 310027

摘要: 传统的基于内容的目标检索方法一般通过提取和比较局部特征, 往往计算复杂度较高而且需要进行离线训练。因此, 以最新的主方向模板(DOT)特征为基础提出新的实时目标检索方法。同时, 利用搜索窗口中网格得分的空间分布情况构造相似度映射图, 并采用多层次金字塔评分去除网格得分较为分散的错误窗口。利用以上改进实现的目标检索方法, 真正能够无须离线训练即可从输入视频直接得到检索结果。实验结果表明, 本文方法在提升检索性能的同时仍然保持主方向模板匹配的实时处理能力。

关键词: 基于内容的目标检索; 主方向模板; 相似度映射图; 金字塔评分策略

Real-time object retrieval with dominant orientation template matching improved by pyramid scoring

Hong Chaoqun, Zhu Jianke, Li Na, Bu Jiajun, Chen Chun

College of Computer Science & Technology, Zhejiang University, Hangzhou 310027, China

Abstract: Traditional methods for Content-based Object Retrieval usually employ extracting and comparing local features, which are too computational demanding and require offline training process. Therefore, a novel real-time CBOR method based on dominant orientation templates (DOT) is proposed. Besides, the proposed method makes use of the spatial distributions of grid scores to construct likelihood maps and utilizes a pyramid-scoring strategy to dispose the false windows in which the grid scores are dispersed. With the above improvement, the implementation for CBOR is able to retrieve the query target directly from the videos without offline training. The results show the proposed method improves the retrieval performance while retaining real-time processing ability of DOT matching.

Key words: content-based object retrieval; dominant orientation templates; likelihood map; pyramid-based scoring strategy

0 引言

目前互联网上信息呈爆炸式增长, 搜索引擎的重要性不言而喻。而信息检索大多仍针对文本信息, 甚至图像检索也是使用文本标签而非图像本身的内容。因此, 研究基于内容的图像检索(CBIR)对

互联网时代的信息检索具有重要意义^[1]。CBIR的目的在于根据与检索目标的相似程度, 对数据集内的图像进行排序, 选择相似度最高的图像。传统的CBIR方法的局限性在于其表示方法通常只描述图像整体, 当检索目标仅与图像的一部分匹配时, CBIR无法获取正确结果。与CBIR不同, 基于内容的目标检索(CBOR)能够寻找包含检索目标的图

收稿日期: 2011-08-06; 修回日期: 2011-10-11

基金项目: 国家自然科学基金项目(60873124, 61103105); 浙江省自然科学基金项目(Y1090516); 中央高校基本科研业务费专项资金资助基金项目(2009QNA5015); 国家科技支撑计划基金项目(2008BAH26B00); 浙江省服务机器人重点实验室基金项目(2008E10004)

第一作者简介: 洪朝群(1984—), 男, 浙江大学计算机科学与技术专业在读博士研究生, 主要研究方向为计算机视觉与模式识别。E-mail: hcq@zju.edu.cn

像,并且能够定位目标,从而根据检索图像的局部与数据集内图像的相关性进行排序^[2]。

针对海量图像或视频进行基于内容的目标检索是CBOR的研究热点之一。Sivic和Zisserman实现的Video Google系统是较为成功的方法之一^[3]。在他们的工作中,图像采用局部特征(如SIFT^[4]或者SURF^[5])进行表示。经过这样的处理,可以认为图像是局部特征的集合。如果把局部特征看作是文本中的词语,那么CBOR就被转化成文本检索问题。CBOR与文本检索的区别只是在于描述内容的特征不同,而由特征的匹配实现检索目标的匹配这一思想则是通用的。采用这一思想定位目标,目前常用的方法有滑动窗口搜索(sliding window searching)^[6-7]、分支定界搜索(branch-and-bound searching)^[1,8]等。滑动窗口搜索在被测试图像中自左向右、自上而下地滑动搜索窗口并且比较窗口中所包含部分的特征与检索目标的特征,从而判断窗口中的内容与检索目标的匹配程度。分支定界则将图像分割为较小的子图,并且将不可能包含检索目标的部分排除。以上过程递归地进行,终将得到一个足够小的子图,其中含有检索目标。但是基于局部特征的检索方法需要耗费大量时间进行特征的提取,一般还需要进行聚类学习从而减少信息量^[9]。

近年来,模板匹配(template matching)受到了很多关注。模板匹配一般使用全局特征,而检索目标的特征描述被称为模板。其优点在于能够通过添加和学习不同的模板而适应不同场合的需要^[6]。它采用的定位方法是滑动窗口。根据前面的描述,滑动窗口搜索的特点是定位过程比较简单,但是遍历图片的运算复杂度较高,因此需要快速有效地表示和匹配模板的特征。早期的模板匹配方法采用边缘信息作为特征,并且采用倒角距离作为相似度的衡量标准^[10]。这一方法的主要缺点在于其性能严重受限于边缘提取方法的性能。而边缘提取方法一般需要手工设定阈值,本身就不具备非常强的鲁棒性。模板匹配的方法进一步发展,出现了利用图像梯度作为特征的方法^[11],其中最为成功的特征是梯度直方图特征(histograms of gradients)^[12]。它将图像分割为网格,并捕捉图像梯度在网格中的分布进行统计。然而提取HOG特征计算复杂度较高,难以适应实时应用的要求。

为了加快特征的提取与匹配速度,Hinterstoisser等人在HOG的基础上开创了一种新的思路,提出了

主方向模板(dominant orientation template)特征^[13]。它既保留了局部统计方法的优点,又将滑动窗口分割为网格。在对网格中的梯度进行统计的时候,DOT将梯度的方向量化到7个,并且只保留网格中梯度较大的方向,使得DOT的特征描述得到充分的压缩。此外,DOT还使用了按位的表示法和计算,从而大幅度降低了特征提取与比较的计算复杂度。为了进一步加速匹配,DOT特征匹配使用的函数利用SSE2指令集进行并行处理。虽然DOT显示了非常好的性能,它仍然存在一些局限。首先,它容易受到噪声干扰,因为在DOT特征的匹配过程中,得分只有匹配与不匹配两种,而没有中间结果对匹配的程度进行评价,容易产生误匹配和漏匹配。其次,它简单采用了搜索窗口的全局得分作为评价搜索窗口是否存在检索目标的标准,这使得它无法排除全局得分较高但局部得分较为分散的错误窗口,而且会错误地排除全局得分较低但局部得分较为集中的正确窗口。

本文提出的改进方法受到HOG-LBP的启发^[14]。其基本思想在于针对每个值得怀疑的搜索窗口构造遮挡映射图。在这个图中,HOG特征的每个小块所得到的响应都得到记录。因此,这个图反映了每个小块的可视程度。接着,HOG-LBP特征采用Mean-shift方法对遮挡映射图进行空间聚类。HOG-LBP的遮挡映射图是根据SVM(支持向量机)分类得分构造。由于不同的可见程度和噪声影响,每个小块的得分为 $[-1, 1]$ 范围内的某一个浮点数。通过正确的设置,Mean-shift算法能够聚合获得相对高分的块和获得相对低分的块,从而帮助HOG-LBP方法区分被遮挡区域和未被遮挡区域。最后,局部检测器对遮挡映射图进行最终分类,从而判断当前窗口是否存在物体。这个方法依赖于SVM的分类得分,需要对每个块单独训练SVM分类器,因此其计算复杂度非常高,只能适用于在图片中进行人体检测等非实时应用。

以模板匹配方法为基础,通过改进DOT特征匹配过程,提出了新的CBOR方法,有效克服传统基于局部特征匹配的CBOR方法的固有缺陷。其贡献主要有3点:

- 1)保留搜索窗口中各个网格的得分,并且按照网格原有的顺序排列,构造DOT相似度映射图。映射图体现了搜索窗口的网格局部得分情况以及目标的可见部分;

2) 分析 DOT 相似度映射图, 利用中值滤波过滤映射图中的噪声, 并采用多层金字塔评分排除全局得分较高但局部得分较为分散的错误窗口, 排除噪声和遮挡的干扰, 实现新的模板匹配方法, 称为 Pyd-DOT (pyramid-scoring dominant orientation templates);

3) 拓展了模板匹配的应用领域, 将 Pyd-DOT 模板匹配方法应用到 CBOR, 实现直接在视频中进行实时检索。

1 Pyd-DOT 模板匹配过程

1.1 改进方法的思路

在通常情况下, 目标检索总是会受到很多干扰。首先由于受到噪声的影响, 检索目标在图像中并不总是非常清晰地得到呈现; 其次现实场景中总是存在大量物体, 而这些物体的空间分布状况比较复杂, 导致遮挡情况在目标检索中非常普遍。这些干扰直接影响目标检索方法的性能。因此, 检索方法需要对噪声和遮挡进行针对性处理, 以提高准确程度。

本文方法利用了 DOT 特征及其匹配函数降低计算复杂度。但是原始的 DOT 匹配函数存在一些不足, 该函数利用 SSE2 指令集对 128 位的数据同时进行操作。而在 DOT 特征中, 每个网格的特征被表示成一个 8 位的整数。也就是说, DOT 匹配函数会同时计算 16 个网格中的特征相似度。结果, 匹配函数仅返回一个搜索窗口的全局得分, 最后的判断决策也根据该全局得分进行。这样的处理方式能够提高匹配过程的速度, 但是局部网格的得分没有得到充分利用。虽然用户可以通过调整阈值使得 DOT 能容忍部分由于噪声和遮挡而导致全局得分较低的搜索窗口, 但是这种简单的处理方式也会使得一些错误窗口混入, 增加误检率。因此, 本文对匹配函数和决策过程进行了改进, 避免增加误检率, 获得适应性更强的方法。

1.2 基于金字塔的相似度映射图评分策略

根据以上的想法, 本文对 DOT 匹配函数的结果进行了改进。与原有实现方法不同, 本文方法保存了每个网格各自的相似度得分。这样不仅可以得知搜索窗口中有多少网格得到正确匹配, 还能获取正确匹配的网格列表。这一列表根据网格排列的顺序可以构造出该搜索窗口的相似度映射图, 即 DOT 相

似度映射图 (DOT likelihood map)。在理想的情况下, 相似度映射图上匹配的网格表示检索目标与搜索窗口的该位置相同, 即检索目标的该部分没有被遮挡, 相似度映射图反映了检索目标的可见部分。但是受到噪声等因素影响, 此时得到的映射图仍然比较粗糙, 还需要进一步处理, 以更好地对滑动窗口的内容进行判断。

首先需要排除噪声的影响。但是引言中提到的采用 Mean-shift 的分割方案无法适用于 DOT 相似度映射图的处理。与 HOG-LBP 情况不同, 使用 DOT 特征匹配结果构造的 DOT 相似度映射图是二值图像。在 DOT 特征中, 匹配函数只能判断两个网格是否相似, 而不能体现它们的相似程度。因此, 1 表示该网格匹配, 0 表示不匹配。如果存在噪声的影响, 两个本应相同的网格就可能无法匹配, 而两个不同的网格也可能偶然得到匹配。在这些情况下, 相似度会被直接判定为 0 或者 1, 而不会得到 0 和 1 之间的中间值。对梯度方向进行量化也导致了多义性, 使得一些本来并不相同的梯度方向被量化到同一个值上, 或者本应较为接近的梯度方向被量化到不同的值上, 提高了误匹配和错匹配的可能性。这些情况在 DOT 相似度映射图上显示, 则会呈现一些漏洞, 如图 1 中的初步映射图。为了填补这些漏洞, 从一定程度上减少噪声的干扰, 本方法采用中值滤波对 DOT 相似度映射图进行处理。中值滤波取每个像素一定范围内邻域上的像素值进行排序, 将中值赋予该像素, 即

$$m(x, y) = M(f(x+i, y+j)) \quad (1)$$

$$|i| < w, |j| < r$$

中值滤波的意义在于能够防止某像素上的值与其周围像素值的分布产生明显的区别, 从而对图像进行平滑。在实验中设置中值滤波的滤波器范围为 $w = r = 3$ 。

接着, 利用排除了噪声的 DOT 相似度映射图, 可以判断目标的可见性, 从而进行遮挡处理。大量实验表明, 如果搜索窗口中存在检索目标, 那么匹配的网格就在空间上会较为集中; 而如果不存在, 那么匹配的网格就较为分散。如果全局得分较低但是在 DOT 相似度映射图上体现出局部得分较高的情况, 往往意味着该窗口内存在被局部遮挡的物体。根据该观察, 本方法统计相似度映射图的多层金字塔得分。在金字塔的第 i 层, 本方法对 $i \times i$ 个网格组成的高层网格判断相似程度。该高层网格判断为匹配

的条件是当且仅当它所包含的所有最底层(第 1 层)网格都匹配,即

$$L_{x,y}^i = \begin{cases} 1 & \forall L_{x,y}^1 \in L_{x,y}^i, L_{x,y}^1 = 1 \\ 0 & \text{其他} \end{cases} \quad (2)$$

如果某个搜索窗口得到的低层次评分较高而高层次评分较低,说明该窗口内匹配的网格较为分散,则可以排除该窗口。实际上可以认为原始的 DOT 方法是本方法在金字塔层数为 1 时的特例。

2 检索方法的实现细节

利用前一节提出的改进过程,本文提出的目标检索方法的实现细节可以用图 1 的流程图进行表示。

- 1) 利用 T_b 过滤明显的错误窗口;
- 2) 通过步骤 1) 检验的搜索窗口被选定为候选窗口(candidate windows),为这些窗口构造 DOT 相似度映射图;
- 3) 对初步的映射图进行中值滤波,去除部分噪声;
- 4) 对滤波后的映射图计算多层金字塔评分,图

1 中灰色网格表示当前金字塔评分所覆盖的网格范围;

- 5) 排除错误的窗口,保留较为可能存在检索目标的窗口。该类窗口所表示的区域为候选区域(candidate regions);
- 6) 对候选区域进行排序。

为了缩短运算时间,本方法对金字塔最底层的得分独立进行检测阈值的设置,该阈值记为 T_b 。经过 T_b 的过滤,本方法可以排除大量明显不可能存在检索目标的窗口,而只对某些较可能存在检索目标的搜索窗口构造 DOT 相似度映射图。阈值 T_b 实际上与原始 DOT 中的检测阈值意义相同。但是为了避免将某些全局得分较低的正确检测窗口排除,本方法设置的 T_b 比原始 DOT 中的检测阈值小,而对于金字塔中高层次的得分,设置第 i 层的阈值为

$$T_i = (T_c^i) T_b \quad i \geq 2 \quad (3)$$

式中, T_c 为事先设定的常数,设定值根据实验经验获得。按照基于内容的目标检索问题定义,对特征匹配过程判断为正确的搜索窗口,使用金字塔的最底层得分进行排序,从而得到检索的结果。

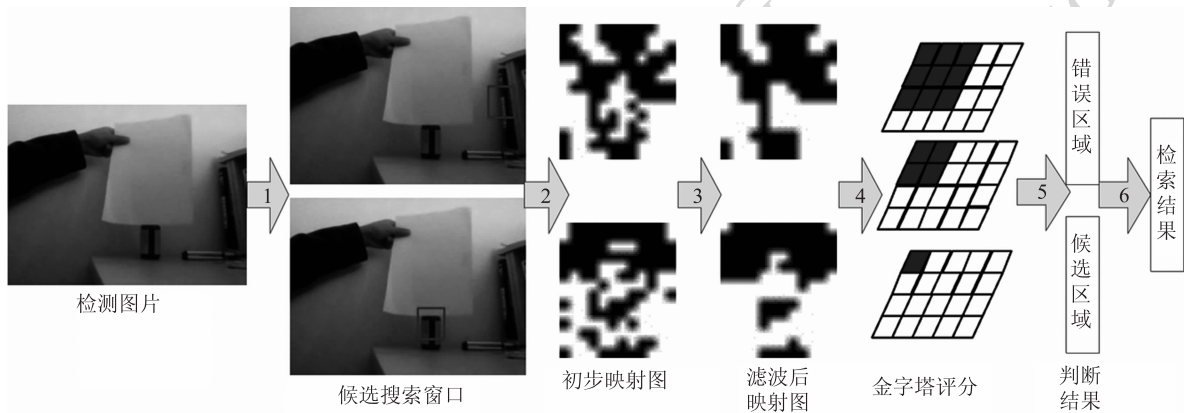


图 1 基于 Pyd-DOT 的目标检索过程
Fig. 1 The process of CBOR based on Pyd-DOT

3 实验结果

3.1 实验设定与数据集

为了验证 CBOR 方法的效果,本实验采用了最近流行的 PROST 数据集^[15]。这个数据集包含 4 段 VGA 分辨率(640 × 480)的视频。这些视频每隔 5 帧有 1 帧对检测目标的位置和大小进行了标记。为了更明确地展示和比较效果,实验数据只保留了视

频中有标记的帧组成被检测视频。实验中每次展示的数据结果都是平均结果,即从 4 段视频中各选择 3 帧,根据标记的矩形框位置提取检索目标,在 4 段视频的集合中进行检索,因此总共进行 4 × 3 = 12 次检索,取结果(包括检索效果、检索时间等)的平均值。

此外,本实验采用 PASCAL 分数作为性能衡量标准。这一评分广泛应用于物体检测与目标检索中,能够衡量检索结果与正确结果的矩形框重叠面

积。PASCAL 分数越高则表明效果越好。本实验对检索得到的前 n 个结果计算平均 PASCAL 得分,随着 n 的增大记录不同的得分情况,形成评价效果优劣的曲线。

3.2 参数的影响

在本方法中,有两个参数需要进行设定,即金字塔的层数和阈值 T_c 。本实验将针对这两个参数的不同设置进行检索性能、运行时间等的比较,从而对它们进行合理的设置。而 T_b 则设置为原始 DOT 实现中默认设置的 80%。

首先,实验在固定阈值 T_c 为 0.7 的情况下改变金字塔的层数,所得到的 PASCAL 得分曲线如图 2 所示。根据前面所述,金字塔层数为 1 的时候实际上就是原始的 DOT 方法,因此本实验并不展示其结果,而在 3.3 节的实验中进行比较。可以清楚地看到:在金字塔层数从 2 增加到 4 的过程中,检索性能只是略有提升。当金字塔层数为 5 的时候,性能最佳。而当金字塔层数继续增加的时候,性能又反而下降。其原因在于随着金字塔层数的增加,干扰区域会更多地被排除。而当层数过大的时候,有些正确的区域又被错误地过滤,导致了性能的下降。金字塔层数变化的影响还体现在候选区域的数量和处理时间上,表 1 所示为金字塔层数对速度的影响。随着层数的增加,候选区域的数量会减少。而由于金字塔层数增加,需要计算更多层次上的金字塔得分,从而导致了处理时间的增加。虽然候选区域数量的减少会降低排序的计算量,但是相比匹配与评

分过程,排序的计算量非常小。实验表明,排序的耗时不到 1ms,因此无法抵销计算更高的金字塔所带来的计算量增加。

表 1 金字塔层数对速度的影响

Table 1 The influence of layers of the pyramid on speed

金字塔层数	候选区域的数量	运算时间/s
2	987	22.479 33
3	907	23.610 675
4	789	25.289 197
5	712	25.535 529
6	435	26.602 323

其次,实验在固定金字塔层数为 4 的情况下改变 T_c ,所得到的 PASCAL 得分曲线如图 3 所示。 T_c 从 0.5 增加到 0.7 的过程中,检索性能变化很小。而当 T_c 增加到 0.8 甚至是 0.9 的时候,检索性能明显下降。这一现象的原因与前一段所述金字塔层数变化所带来的影响类似。随着 T_c 的增大,有些正确的区域被错误地认为是干扰区域而被排除,从而引起检索性能的降低。 T_c 数值变化的影响同样体现在候选区域的数量和运算时间上,如表 2 所示。候选区域的数量随着 T_c 的增大而减少,而运算时间会出现一些波动,且变化并不明显。因为 T_c 值的变化并不影响匹配过程的计算量,仅仅只是加快了最终排序的速度,而排序所需的运算时间在总体运算时间中所占比例很小。

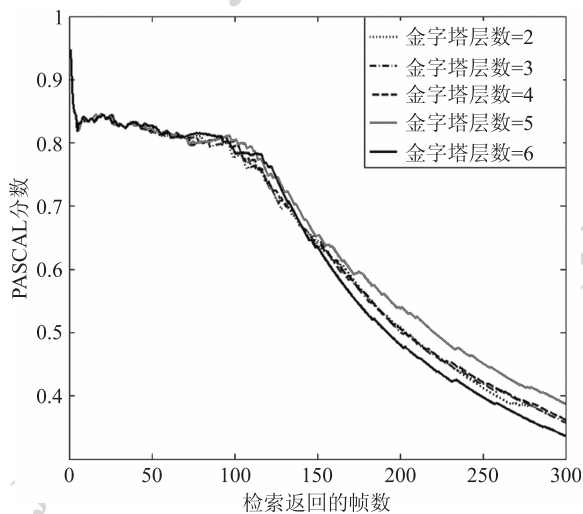


图 2 金字塔层数对性能的影响

Fig. 2 The influence of layers of the pyramid on performance

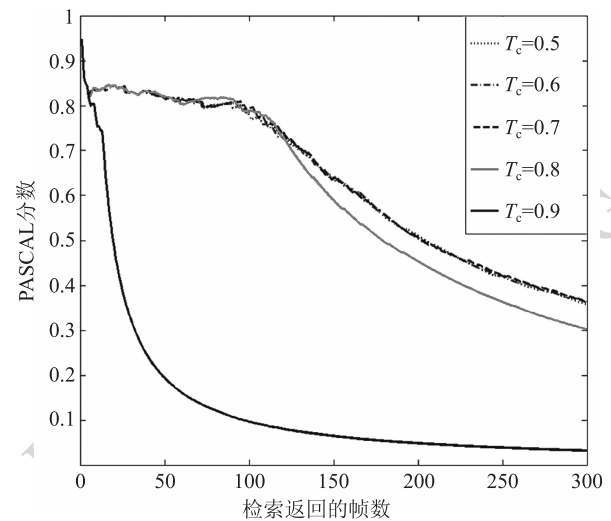


图 3 T_c 对性能的影响

Fig. 3 The influence of T_c on performance

表 2 T_c 对速度的影响
Table 2 The influence of T_c on speed

T_c	候选区域的数量	运算时间/s
0.5	990	24.815 180
0.6	957	25.041 197
0.7	789	25.289 197
0.8	270	24.451 018
0.9	13	26.412 325

3.3 与现有方法的比较

通过实验将本文提出的算法与 ESR 方法^[1]和利用原始 DOT^[17]实现的 CBOR 方法进行比较。根据 3.2 节的实验结果,本文方法在设置金字塔层数为 5, 阈值 T_c 为 0.7 的时候能达到最好的效果。因此,实验中,本文方法将按照此设置进行。

首先,实验通过 PASCAL 分数曲线对检索效果进行比较。3 种方法的曲线显示在图 4 中。从图中的曲线可以看出,原始的 DOT 方法在大部分情况下性能优于 ESR。而本文提出的在 DOT 方法基础上进行改进的 Pyd-DOT,其曲线始终保持其他两条曲线之上。因此,本方法(Pyd-DOT)的检索性能明显优于另外两种方法。

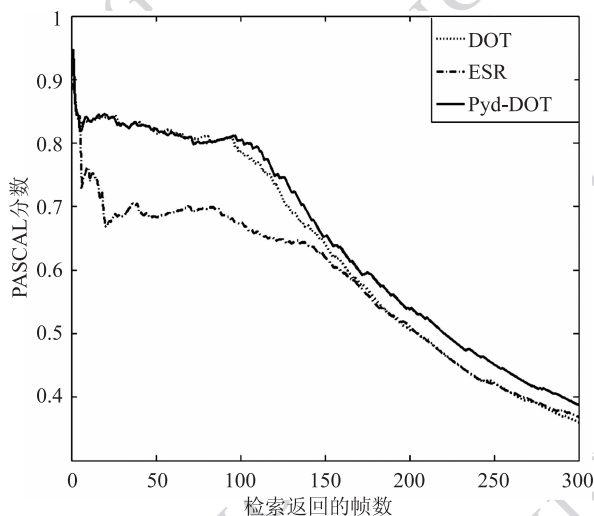


图 4 Pyd-DOT、ESR、DOT 的性能比较

Fig. 4 The comparison of performance among Pyd-DOT, ESR and DOT

其次,实验比较了 3 种方法平均每帧的运算时间,如表 3 所示。对于 ESR 而言,其运算时间与返回的结果数量相关。返回的结果数量越多,运算时间越长。由于分支定界方法的优越性,其检索速度

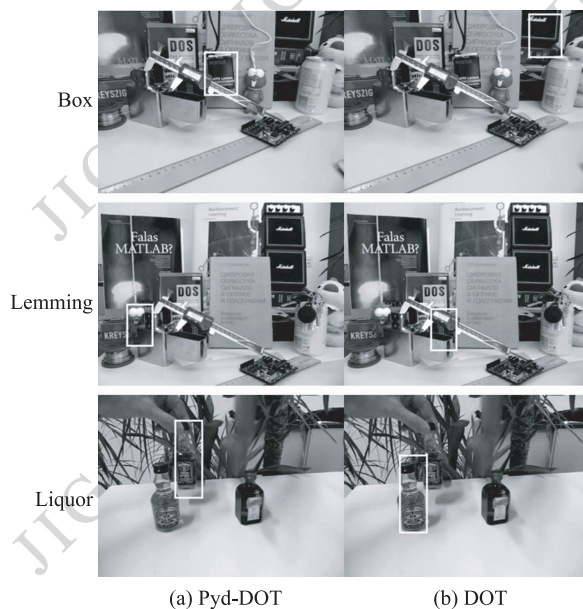
非常快。返回前 300 个检索结果平均只需要 0.9 s。但是该方法需要复杂的训练过程,包括提取局部特征、训练字典等。这一过程的时间消耗难以衡量。因此,与 ESR 进行检索时间消耗的对比没有太大的意义。而对于 DOT 与本文提出的 Pyd-DOT 而言,特征的提取过程非常迅速,真正具备了无需长时间对特征、词典、分类器的事先训练,直接从输入视频得到检索结果的能力。实验表明,引入了基于金字塔的评分改进之后,Pyd-DOT 的运算时间有所增加,但是仍然能够与 DOT 一样达到实时处理的效果。

表 3 运算时间比较

Table 3 Comparison of computational time

	/ms	
ESR	DOT	Pyd-DOT
3.0	17.9	25.8

最后,图 5 展示了一些典型的直观效果图。定位结果用白色矩形框画出。对于 Box 序列和 Liquor 序列,原始的 DOT 方法定位到与检索目标相似的物体上。这证明了原始 DOT 方法误匹配率较高的问题。对于 Lemming 序列,原始的 DOT 只能随机定位到一些错误的区域上。这证明原始 DOT 方法在存在遮挡的情况下难以正确定位目标。以上现象的原因在于检索目标的局部被遮挡,导致全局得分下降。而与检索目标相似的物体或者某些噪声区域在全局上得到了较高的分数。该现象也印证了引言中所述



(a) Pyd-DOT (b) DOT

图 5 定位结果比较

Fig. 5 The comparison of localization results

DOT 方法的缺陷。本文提出的 Pyd-DOT 方法通过构造与分析相似度映射图,能够利用局部得分情况对搜索窗口的内容进行更为准确的判断。利用局部得分进行判断的过程能够获取局部得分较为集中的区域,从而排除全局得分较高但局部得分较为分散的干扰区域,可以改善局部遮挡等干扰下的检索结果,提高检索性能。

4 结 论

本文提出了以主方向模板特征为基础进行改进的特征匹配方法,并且应用于基于内容的目标检测。首先,该方法保留了 DOT 特征匹配过程中各个网格的得分,构造 DOT 相似度映射图,从而便于对搜索窗口的局部匹配情况进行分析。其次,本方法利用中值滤波过滤 DOT 相似度映射图,从而排除部分误匹配和填补部分漏匹配,并且利用多层次金字塔模型计算评分,排除大量全局得分较高的错误窗口,保留可能产生局部遮挡的搜索窗口。最后,根据金字塔最底层得分对可能存在检索目标的窗口进行排序,从而得到基于目标的检索结果。实验结果表明,该方法在略微提高了运算时间的代价下,比现有方法获得了更好的检索效果。

参考文献 (References)

- [1] Yang F Y, Wang X Y. Color Image Retrieval Based on Multiple Features of Image Edges [J]. Computer Science, 2010, 37(2): 256-260. [杨芳宇,王向阳.一种基于边缘综合特征的彩色图像检索算法[J].计算机学报,2010,37(2):256-260.]
- [2] Lampert C H. Detecting objects in large image collections and videos by efficient subimage retrieval [C]//Proceedings of IEEE International Conference on Computer Vision. New York: IEEE, 2009:987-994.
- [3] Sivic J, Zisserman A. Efficient visual search of videos cast as text retrieval [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 31:591-606.
- [4] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [5] Bay H, Ess A, Tuytelaars T, et al. Surf: Speeded up robust features [J]. Computer Vision and Image Understanding, 2008, 110:346-359.
- [6] Olson C F, Huttenlocher D P. Automatic target recognition by matching oriented edge pixels [J]. IEEE Transactions on Image Processing, 1997, 6(1):103-113.
- [7] Wei Y, Tao L. Efficient histogram-based sliding window [C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2010:3003-3010.
- [8] Lampert C H, Blaschko M B, Hofmann T. Efficient subwindow search: A branch and bound framework for object localization [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31:2129-2142.
- [9] Gao K, Lin S X, Zhang Y D, et al. Object-based image retrieval using spatial context [J]. Journal of Computer-Aided Design & Computer Graphics, 2008, 20(11):1452-1458. [高科,林守勋,张勇东,等.基于空间上下文的目标图像检索[J].计算机辅助设计与图形学学报,2008,20(11):1452-1458.]
- [10] Gavrilin D, Philomin V. Real-time object detection for "smart" vehicles [C]//Proceedings of IEEE International Conference on Computer Vision. New York: IEEE, 1999: 87-93.
- [11] Steger C. Occlusion, clutter, and illumination invariant object recognition [C]//Proceedings of International Archives of Photogrammetry and Remote Sensing. Nottingham: ISPRS, 2002:345-350.
- [12] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2005:886-893.
- [13] Hinterstoisser S, Lepetit V, Ilic S, et al. Dominant orientation templates for real-time detection of texture-less objects [C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2010:2257-2264.
- [14] Wang X, Han T X, Yan S. An hog-lbp human detector with partial occlusion handling [C]//Proceedings of IEEE International Conference on Computer Vision. New York: IEEE, 2009:32-39.
- [15] Santner J, Leistner C, Saffari A, et al. Prost: Parallel robust online simple tracking [C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2010:723-730.