

非清晰区域抑制下的显著对象检测方法

郑 阳, 刘纯平, 柳 恭, 王朝晖

(苏州大学 计算机科学与技术学院, 江苏 苏州 215006)

摘 要: 基于上下文感知的显著区域检测模型(Context-Aware, CA)对于大目标和复杂背景图像中显著对象检测存在检测内容缺失和误检的问题. 在 CA 模型的基础上, 引入图像清晰度的视觉反差特性, 提出非清晰区域抑制下的图像显著对象检测方法. 该方法以离散度作为判断图像中是否存在清晰度差异的标准, 并对存在差异的图像进行抑制. 实验结果表明, 非清晰区域抑制的 CA 方法可以在较好的解决大目标检测和复杂背景误检问题, 提高了显著对象检测精度.

关键词: 上下文感知; 清晰度; 显著性; 显著对象检测

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2012)03-0084-05

Saliency Detection Based on Inhibition of Blur Regions

ZHENG Yang, LIU Chun-ping, LIU Gong, WANG Zhao-hui

(Department of Computer Science and Technology, Soochow University, Suzhou 215006, China)

Abstract: There are problems of content missing and false detection for Context-Aware(CA) Saliency detection model when detecting image with large scale object or complex background. Based on Context-Aware model, with the introduction of visual contrast characteristic of image clarity, a saliency detection method based on inhibition of blur regions was proposed. The dispersion was made as the criterion to judge whether there exist clarity difference in an image, and the image would be inhibited if there exist the difference. The experimental result shows that, the saliency detection method based on inhibition of blur regions can better solve the problem of content missing and false detection in some special images with higher detection precision.

Key words: context-aware; clarity; saliency; saliency detection

1 引言

人眼视觉注意机制类似于利用一个滤波器以过滤场景中非显著的信息的过程^[1]. 显著对象检测模型可分为自底向上、自顶向下以及混合模型三类. 特征集成理论(Feature Integration Theory, FIT)^[2]是自底向上模型的计算基础. 其中 Itti 等人^[3]以中心-周围强度、颜色和方向对比度特征提出了第一个也是最经典的视觉注意计算模型 Itti's98. 随后该模型通过引入运动和闪变的对比度特征而扩展到视频显著性计算中^[4]. Walther 和 Koch 在 Itti's98 模型的基础上进行了扩展以便得到一个显著典型(proto)目标^[5]. 利用图像分割获得区域显著性^[6]、利用提取

语义线索(人脸、语音等)^[7]、利用机器学习方法学习^[8]等提出了自顶向下的显著计算模型. 这些模型基本上是以比较简单的图像进行显著对象的检测, 对背景复杂和大目标的检测效果不佳.

本文以上下文感知的自顶向下模型(CA)^[8]为基础, 提出利用离散度指标作为图像清晰度判决准则, 在抑制非清晰区域的基础上计算图像的显著性, 以便有效改善 CA 模型中对大目标和复杂背景图像的显著对象检测性能.

2 图像清晰度判决规则

人眼按照自底向上的特征选择接收信号, 并将注意力转移到人眼视觉反差较大的区域. 图像清晰

收稿日期: 2011-07-06; 修回日期: 2011-08-10

基金项目: 国家自然科学基金项目(60873116, 61070223); 江苏省自然科学基金项目(BK2009593, BK2009116)

度也是影响视觉反差的一个重要因素. 图像清晰度指图像子区域细节轮廓边缘的清晰程度, 也是图像颜色亮度等特征在边界变化的剧烈程度, 变化越剧烈, 视觉反差越大, 图像越清晰. 人眼会首先迅速将注意力集中在相对清晰区域, 并在该区域检测目标显著与否. 如何判断图像是否存在清晰度差异是本文显著对象检测方法的一个关键.

2.1 离散度指标

由于图像聚焦点的不同, 一幅图像中存在不同子区域的清晰度差异. 从图像特征的角度而言, 清晰度衡量的是图像中颜色亮度等特征的剧烈变化, 在空域内表现为灰度突变, 在频域上表现为高频分量的丰富, 并且在频域上高频分量的突变比在空域内的灰度突变要剧烈得多. 因此离散度 λ 指标利用图像本身的高频信息来测量.

2.1.1 图像高频信息映射图

离散度 λ 是一个用来衡量图像高频信息聚集性的指标, 图像子区域是否存在清晰度差异则依据图像中高频信息的聚集程度来判断.

图像高频信息映射图 (High Frequency Map, HFM) 生成采用小波变换方法. 假设输入图像为 $f(i, j) = \{R, G, B\}^T$, 小波变换后的垂直、水平和对角线高频分量分别为 H_{vf} 、 H_{hf} 和 H_{df} , 则输入图像的高频信息区域 $H_{f(i, j)}$ 为:

$$H_{f(i, j)} = H_{vf(i, j)} + H_{hf(i, j)} + H_{df(i, j)} \quad (1)$$

其中 i, j 为像素坐标. 图 1 给出了有无图像清晰度差异的 HFM 结果. 对于前景和背景有明显清晰度差异的图像, HFM 图中高频信息点集中在原图像清晰区域位置, 高频信息点聚集性较好. 前景和背景没有明显清晰度差异的图像, 其 HFM 图中的高频信息点分散在整幅图像中, 聚集性差.



图1 有无清晰度差异图与 HFM 图 (第 1、3 行分别为有和无清晰度差异图像, 第 2、4 行分别为 1、3 行的 HFM 图)

2.1.2 HFM 聚集性度量

离散度 λ 反映的是图像数据分布以及聚集程度, 以此标准作为衡量 HFM 中高频信息点的聚集程度.

假设把 HFM 图以合适的大小分成 $m \times n$ 块, 统计每个块内的高频信息点分布, 并将其转换为一维向量 (图 2).

$$\vec{a} = (\sum H_{f1}, \sum H_{f2}, \dots, \sum H_{f(m \times n)}) \quad (2)$$

对 $\vec{a}$ 归一化:

$$\vec{b} = \frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}, \frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}, \dots, \frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}} \quad (3)$$

HFM 的离散度:

$$\lambda = \frac{\sigma_{\vec{b}}}{\mu_{\vec{b}}} = \frac{\sigma\left(\frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}, \frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}, \dots, \frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}\right)}{\mu\left(\frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}, \frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}, \dots, \frac{\sum_{i=1}^{m \times n} H_{fi}}{\sum_{i=1}^{m \times n} H_{fi}}\right)} \quad (4)$$

其中, $\sigma_{\vec{b}}$ 和 $\mu_{\vec{b}}$ 分别为 $\vec{b}$ 的标准方差和平均值, 离散度 λ 反映的是向量中数据的分布和离散程度, λ 值越大, 数据的离散性越差, 聚集性越好, 反之 λ 值越小, 数据的离散性越好, 聚集性越差.

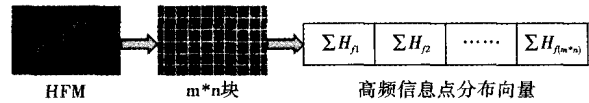


图2 提取高频信息点分布向量过程

2.2 基于离散度指标的判决规则

从式(4)离散度 λ 的计算可以看出, 离散度受 HFM 图分块尺度影响. 因此期望在某一种分块尺度下, 使有清晰度差异图像的离散度 λ_{Yes} 和无清晰度差异图像的离散度 λ_{No} 之间的差距最大, 即目标类别的类间距达到最大. 选取具有最大区分性的分类阈值, 可以减少错分概率.

为使 λ_{Yes} 和 λ_{No} 之间稳定有分界面, 即期望类间距最大的同时类内的波动较小, 定义某一尺度下有、无清晰度差异图像的类间距离:

$$dist_{Clarity} = (\mu_{\lambda_{Yes}} - \sigma_{\lambda_{Yes}}) - (\mu_{\lambda_{No}} + \sigma_{\lambda_{No}}) \quad (5)$$

其中 $\mu_{\lambda_{Yes}}$ 、 $\mu_{\lambda_{No}}$ 、 $\sigma_{\lambda_{Yes}}$ 和 $\sigma_{\lambda_{No}}$ 分别为有、无清晰度差异图像集合的平均值和方差. 由于离散度 λ 指标受分块的影响, 类间距离 $dist_{Clarity}$ 也随分块尺度变化而变化, 需要选取合适尺度的分块使得类间距离

最大. 经统计分析, 当选取块尺度为 $size/16$ 时, 类间距离 $dist_{Clarity}$ 达到最大. 其中 $Size = \max\{Height, Width\}$, 图像的大小为 $Height \times Width$.

在该尺度下, 由于类间距离较大, 对阈值的选取并不敏感, 为了获得最大的分类间隔, 选 $\lambda=1$ 为分类阈值, 即图像子区域清晰度差异的判决规则为:

$\lambda > 1$, 图像中存在子区域清晰度差异;

$\lambda \leq 1$, 图像中不存在子区域清晰度差异.

3 非清晰区域抑制下的显著对象检测

现有显著对象检测方法由于背景杂乱, 对显著对象的检测往往存在很多误检的情况, 通过对图像非清晰区域的抑制, 可以有效改善杂乱的背景, 从而使得清晰的或者显著的前景目标得到完整保留, 以提高显著对象检测精度. 由于清晰区域的大小、轮廓等特征的不确定性, 很难从图像的全局角度出发来检测整幅图像中的相对清晰区域. 本文采用多尺度局部块清晰度评价方法计算图像中的每个局部子块清晰度, 抑制清晰度值小的区域, 并结合 CA 模型^[8]进行显著对象的检测.

3.1 多尺度图像子区域清晰度

认知心理学研究表明, 人眼会从多尺度提取感兴趣的视觉信息, 先整体加工视觉信号, 然后再局部加工^[9]. 因此多尺度的图像分块、每个子块的局部清晰度计算, 符合视觉观察重要特征, 可以很好的保留显著对象全局和局部信息.

清晰区域特征在空域用梯度函数度量, 在频域用小波变换来凸显. 图像采取三种不同尺度 $\{Scale1, Scale2, Scale3\}$ 分块. 在以离散度指标计算中确定的一种合适分块尺度 $size/16$ 的基础上, 选取另外 $size/8$ 和 $size/32$ 两种尺度即 $scales = \{size/8, size/16, size/32\}$ 来综合评定块的清晰度值. 每个分块尺度下的第 i 个子块像素清晰度定义为:

$$D_{i(m,n)Scalej} = D_{iScalej} \quad (6)$$

其中 (m,n) 为第 i 子块像素坐标, $D_{iScalej}$ 为尺度 j 下第 i 子块清晰度, 其定义为:

$$D_{iScalej} = \frac{1}{n_{iScalej}} \left[\sum_{(x,y) \in D_{iScalej}} \left| \frac{H_{vf}(x,y)}{L_f(x,y)} \right| + \sum_{(x,y) \in D_{iScalej}} \left| \frac{H_{hf}(x,y)}{L_f(x,y)} \right| + \sum_{(x,y) \in D_{iScalej}} \left| \frac{H_{df}(x,y)}{L_f(x,y)} \right| \right] \quad (7)$$

其中 L_f 、 H_{vf} 、 H_{hf} 和 H_{df} 分别为子块图像 f 小波变换后子图像的低频、垂直高频、水平高频和对角

线高频分量. $n_{iScalej}$ 为尺度 j 下第 i 子块低频图像的像素数, (x,y) 为第 i 子块的行列数.

将多尺度像素清晰度进行线性融合, 可得到基于块的图像清晰度.

$$\bar{D}_{(m,n)} = \sum_{j=1,2,3} D_{(m,n)Scalej} \quad (8)$$

式(8)得到的图像清晰度块效应比较明显. 为消除块效应对后续显著对象检测的影响, 引入加权系数, 最终加权图像清晰度为:

$$D_{(m,n)} = \bar{D}_{(m,n)} \times Weight_{(m,n)} \quad (9)$$

$$Weight_{(m,n)} =$$

$$\frac{1}{\sqrt{[I(m+1,n) - I(m,n)]^2 + [I(m,n+1) - I(m,n)]^2}} \quad (10)$$

3.2 非清晰区域抑制

根据式(9)得到源图像清晰度值矩阵, 选用大津法对图像进行分类. 该方法能够从图像全局角度使不同类的类间方差达到最大, 使得对图像清晰度值分类可以达到较精确的分类效果, 抑制属于非清晰区域的像素点. 然而由于以上分类是基于像素点的, 没有考虑像素之间的关联性和区域像素的相似性, 分类结果存在空洞和噪声现象. 通过分类后去除小连通区域的图像后处理, 可以去除噪声和空洞, 提高分割精度. 图3给出了有清晰度差异图像的非清晰区域抑制结果.

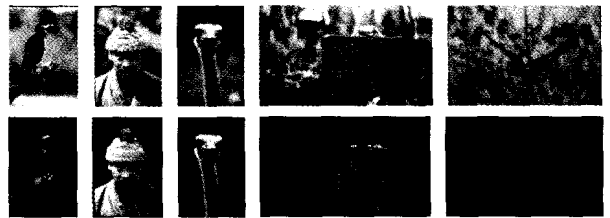


图3 有清晰度差异的图像抑制非清晰区域后的结果 (第1行为原图像, 第2行为抑制结果)

3.3 结合清晰度的 CA 显著对象检测方法

CA 模型^[8]对显著目标尺度较小且背景较简单的情况下能够得到较准确的结果, 如图4第2、4列所示. 然而对于某些背景较复杂且相对前景较模糊的图像, CA 得到的显著区域和非显著区域的显著值差异不大, 不利于最终的显著对象检测. 在图像存在明显的清晰度差别的情况下, 对图像进行非清晰区域抑制, 再对抑制后的图像用 CA 方法处理, 这样能够为 CA 方法提供一个理想的条件, 充分发挥其优势. 即从整体上为显著对象的检测增加了一个先验知识的控制, 抑制了非清晰区域中较显著点的影响, 从而解决了部分因背景复杂而造成的误检问题.

图5给出了存在清晰度差异的图像显著图结果,图5第2列可以看出,由于原始图像背景复杂,CA模型得到的显著图中显著对象并不明显.图5第4列是在第1列图像清晰度判断并抑制后得到的显著图,结合清晰度后的CA模型有效改善显著图效果,有利于后续的显著对象的提取.



图4 理想条件下CA方法结果
(第1、3列原图,第2、4列对应结果)



图5 CA与结合清晰度的CA方法对比(第1列原图像,第2列原CA模型,第3列抑制后图像,第4列结合清晰度的CA方法)

4 实验结果分析

为验证提出方法对于复杂背景图像中显著对象检测的有效性,在一个复杂且模糊背景下的图像库BlurDB以及国际公认的三个图像显著对象检测数据库MRSC21、BSD500和VOC2007中,先分析了提出模型的性能,然后对Itti's98模型^[3]、CA模型^[8]、SR模型、GBVS模型、频率调整的FT模型以及提出的显著检测模型进行了比较分析.

4.1 提出模型性能分析

提出显著对象检测是以子块划分为基础,而分块基础则是以高频信息点聚特性进行.因此为验证这一特点,首先在图像库BlurDB中进行测试,然后再在MRSC21、BSD500和VOC2007三个图像库上进行验证.并进行了高频信息点聚集特性分析和分块尺度确定分析,证明具有较好检测性能.

为验证结论普遍性,分别对MRSC21、BSD500、VOC2007三个库进行实验,得到三组有差异图像离散度平均值、无差异图像离散度平均值和类间距离,可以看出三个数据库中三种数据的分布走向基本一致.按图像子区域清晰度差异判决规则,MRSC21、BSD500和VOC2007三个数据库中图像分类准确

率分别为96.2%,98.1%,97.4%.

4.2 几种显著对象检测模型的对比分析

为客观评价提出模型显著对象检测精度,利用自构图像库SaliencyDB进行实验.该数据库选取人眼可分辨显著区域图像300幅,其中BSD500库120幅、MRSC21库80幅、VOC2007库80幅以及网上图像20幅. GroundTruth显著对象采用手工定位分割获得.对比方法分别为Itti's98^[3]、SR、GBVS、FT、CA五种.生成的显著图和显著对象提取结果分别见图6和图7,图像显著对象提取精度比较见表1.

表1 不同显著区域提取方法准确度比较

提取方法	Itti's98 ^[3]	SR	GBVS	FT	CA ^[8]	本文
准确度(%)	53.3	33.3	59.7	70.3	68.7	84.7



图6 各种方法显著图结果(从左到右为源图像, Itti's98, SR, GBVS, FT, CA和本文方法显著图)



图7 显著区域提取(从左到右为源图像, Itti's98, SR, GBVS, FT, CA, 本文方法和 GroundTruth 结果)

从图6可以发现,本文方法相对于其它5种方法而言,整个显著对象区域都具有较高的显著度,保持了更好的对象完整性.图7显著对象的提取结果表明,本文方法获得的显著对象分割结果更接近于GroundTruth的结果,较好的解决了大目标检测内容缺失以及背景复杂下显著对象误检的问题,提取显著目标的准确性也有明显的提高.

5 结束语

文中在视觉注意机制优先注意图像强烈视觉反差的事实上,提出利用离散度来衡量图像中高频信息点聚特性为基础的图像清晰度判别规则,并以此进行图像是否存在清晰度差异的判断,在多尺度下抑制非清晰区域,扩展 CA 模型建立了一种新的显著对象模型.实验结果表明,加入非清晰区域抑制的 CA 模型一定程度上解决了大目标和背景复杂图像显著对象检测中存在的内容缺失和误检问题,有效改善复杂背景下显著对象检测的准确性,提高了检测的精度.

参考文献:

- [1] Itti L, Baldi P. A principled approach to detecting surprising events in video [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA; San Diego, 2005: 631-637.
- [2] Treisman A M, Gelade G. A feature integration theory of attention [J]. Cognitive Psychology, 1980, 12 (1): 97-136.
- [3] Itti L, Koch C, Niebur E. 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.
- [4] Itti L, Koch C. Computational modeling of visual attention [J]. Nature Review Neuroscience, 2001, 2 (3): 194-203.
- [5] Walther D, Koch C. Modeling attention to salient pro-
- to objects [J]. Neural Networks, 2006, 19(9), 1395-1407.
- [6] Liu T, Yuan Z J, Sun J A. Learning to detect a salient object [J]. IEEE transactions on pattern analysis and machine intelligence, 2011 (33): 353-367.
- [7] Cerf M, Harel J, Einhauser W. Predicting human gaze using low-level saliency combined with face detection [C]//Advances in neural information processing systems. California, 2008: 241-248.
- [8] Stas Goferman, Lihi Zelnik-Manor, Ayellet Tal. Context-aware saliency detection [C]//Proc. IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, 2010:2376-2383.
- [9] 邵静. 基于尺度空间表示的视觉注意区域选择[J]. 微电子学与计算机, 2009, 26(10): 12-16.

作者简介:



郑 阳 女,(1987—),硕士研究生. 研究方向为图像处理、模式识别.

刘纯平 女,(1971—),硕士生导师. 研究方向为图像处理与分析、模式识别与计算机视觉、信息融合.

柳 恭 男,(1987—). 研究方向为图像处理、模式识别.

王朝晖 女,(1967—),副教授. 研究方向为多媒体技术.