

视觉注意与显著性计算综述

孙晓帅, 姚鸿勋

(哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001)

摘要: 人类视觉系统通过选择性视觉注意机制来对场景中位于重要位置的视觉内容进行动态的序列采样,进而获取必要的视觉信息。本文系统地总结了计算化注意模型和显著度计算领域的研究现状。通过在 YORK-120 和 MIT-1003 这两个国际标准数据库上进行的人眼视点预测实验,本文对 20 种计算模型的实际性能进行了详细的评估和对比。结果表明,基于统计的模型要比其它的方法更容易获取较好的预测结果。

关键词: 视觉注意; 显著度; 生物启发式方法; 人眼注视点预测

中图分类号: TP391.4

文献标识码: A

文章编号: 2095-2163(2014)05-0014-05

A Survey of Visual Attention and Saliency Computation

SUN Xiaoshuai, YAO Hongxun

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

Abstract: Human vision system acquires essential information from the environment by sequentially sampling visual contents at important locations under the control of selective visual attention mechanism. This paper systematically investigates the state-of-the-arts of computational visual attention modeling and saliency computation. And the paper compares the performance of 20 state-of-the-art models via dense eye-fixation prediction experiments conducted on YORK-120 and MIT-1003 datasets. The results show that statistical models tend to achieve better predictions against the rest approaches.

Key words: Visual Attention; Saliency; Biological Inspired Approach; Eye-Fixation Prediction

0 引言

人类的视觉系统在感知外界环境的过程中存在着极强的动态选择性,这同时体现在其视神经系统的生理结构及作用机理上。以眼睛的生理结构为例,在视网膜盘颞侧约 3.5mm 处,有一个黄色小区,称黄斑(macula lutea),其中央的凹陷叫做中央凹,此处的视神经细胞分布最为密集,感知到的视觉信息最精确。中央凹虽然只占整个视觉面的 0.01%,但是视神经里 10% 的信息是由连接在这里的轴突传递到大脑。当人们观察一个场景的时候,并不是一次性地感知场景中的所有信息,而是基于注意转移机制通过一系列的视点移动来对场景进行逐步采样。观察者能够利用视网膜中央凹的高分辨率感知机构,主动地探索环境中的重要视觉信息。得益于这种动态选择性行为,人类以及大多数的灵长类动物,能够在有限的时间内高效灵动地处理复杂环境中的信息。在过去的 25 年中,神经计算领域的科研工作者们通过理论推理和计算模型构建等手段,试图揭示指导眼动行为的基本原则和作用机理,取得的相关成果不但能够帮助人们更好地理解人类的动态认知行为,同时也为视觉相关的应用领域,如视频压缩^[1]、照片质量评价^[2]、场景理解^[3]、目标检测^[4]及识别^[5]、图像重构图^[6]等,提供了强大的信息分析和处理工具。

本文将系统介绍视觉注意建模及信号显著性计算领域的国际相关工作,并以研究思路为线索对领域发展的现状进

行梳理和分析,同时基于人眼注视点分类实验来对多种模型的实际性能进行量化的比较。

1 视觉注意与显著度计算研究概述

注意是一个心理学概念,属于认知过程的一部分,具体来说,指的是选择性地将对处理资源集中到环境中的某个部分而将其余部分忽略的过程。注意能够将知觉选择性地集中,并导致局部刺激的意识水平的提高。例如,侧耳倾听某人的说话,而忽略房间内其他人的交谈。注意并不是一种独立的心理过程,而是心理过程的一种共同特征。人在同一时间内不能感知很多对象,只能感知环境中的少数对象。而要获得对事物的清晰、深刻和完整的反映,就需要使心理活动有选择地指向有关的对象。人在清醒的时候,每一瞬间总是注意着某种事物。通常所谓“没有注意”,只不过是当前所应当指向的事物没有注意,而注意了其他无关的事物。注意有两个基本特征:指向性和集中性。指向性表现为对出现在同一时间的许多刺激的选择;集中性表现为对干扰刺激的抑制,其产生和范围以及持续时间取决于外部刺激的特点和人的主观因素。学界普遍认为眼动行为是由自底向上(Bottom-Up)和自顶向下(Top-Down)的两种注意因素所驱动^[7-11]。自顶向下的注意因受到具体任务和人的主观意识的影响,其研究结果往往呈现出较大的差异性^[12-19];相比较而言,自底向下的研究工作探索的是人类视觉系统中不受特定任务和特定人物个性约束的共性机理,具有更可控的研究环境和更

收稿日期: 2014-07-18

基金项目: 国家自然科学基金(61133003); 国家重点基础研究发展计划(973)(2009CB320906)。

作者简介: 孙晓帅(1984-),男,辽宁鞍山人,博士研究生,主要研究方向: 计算机视觉与模式识别;

姚鸿勋(1964-),女,黑龙江哈尔滨人,博士,教授,博士生导师,主要研究方向: 多媒体内容分析与检索、计算机视觉、人工智能等。

广泛的应用场景,因此,本章着重从自底向上的角度对注意建模相关工作展开分析和讨论。表 1 按照模型对注意和显著性的不同度量标准,包括对近年来提出的有代表性的注意和显著度计算模型进行了梳理及呈现。

表 1 近年提出的有代表性的注意及显著度计算模型
Tab.1 The representative computational models for visual attention and saliency

模型	作者	年份	注意及显著性度量	时间开销
ITTI ^[7]	Itti and Koch	1998	颜色、方向等生物启发式特征融合	0.55s
CG ^[48]	Torralba et al.	2006	上下文特征指引	1.56s
GBVS ^[50]	Harel et al.	2006	基于图模型的显著度	1.22s
SR ^[40]	Hou and Zhang	2007	频域谱残差	0.02s
QPFT ^[27]	Guo et al.	2008	相位保持的傅里叶变换	0.02s
SUN ^[25]	Zhang et al.	2008	基于自然图像统计的显著度	1.35s
ICL ^[26]	Hou et al.	2009	稀疏表示的增量编码距离	0.12s
AIM ^[23]	Bruce and Tsotsos	2009	稀疏表示的信息最大化	4.87s
JUDD ^[44]	Judd et al.	2009	基于底层和高层特征的有监督学习	14s
FT ^[5]	Achanta et al.	2009	频率修剪	0.49s
STSR ^[51]	Seo et al.	2009	时空自相似性	1.43s
SER ^[27]	Wang et al.	2010	稀疏表示的点熵率	21.4s
SIM ^[45]	Murray et al.	2011	非参数的多尺度整合	4.26s
ELG ^[42]	Borji and Itti	2012	图像块的局部和全局稀少性	N/A
CAS ^[6]	Goferman et al.	2012	上下文敏感的显著性	50.5s
QSA ^[41]	Schauerte et al.	2012	QPFT 的特征角和特征轴方法	0.03s
SIG ^[31]	Hou et al.	2013	图像签名	0.16s
HFT ^[52]	Li et al.	2013	超复数傅里叶变换	0.26s
RARE ^[33]	Riche et al.	2013	多尺度的特征稀少性	1.21s
SGP ^[43]	Sun et al.	2012	超高斯成分分解与信息最大化	0.03s~6.00s

1.1 生物启发式建模方法

视觉注意建模的理论研究工作始于 20 世纪 80 年代。1980 年,美国普林斯顿大学的 Treisman 和 Gelade 提出了注意的特征集成理论 (Feature Integration Theory)^[20],该理论给出了视觉注意计算的重要视觉特征,同时也解释了这些特征是如何在视觉搜索任务中被集成起来并进一步指引人类注意的分配。1985 年,加州理工大学的 Koch 和 Ullman^[21]提出了“显著图”(Saliency Map)的概念,即一张与输入信号相同大小的二维图像,其中每个像素的值代表了图像源中对应位置视觉信息的显著性。基于显著图的概念,Koch 和 Ullman 随之构建了首个有可信生理依据的计算机化视觉注意模型(后文简称 KU 模型)。KU 模型主要包括四个功能模块:(1)底层视觉特征提取;(2)基于中央-外围差分的特征图 (Feature Map) 计算;(3)多通道、多尺度特征图融合;(4)注意选择 (Attentional selection) 和返回抑制 (Inhibition of Return)。KU 模型的建立为视觉注意建模研究的发展奠定了稳固且良好的基础。1998 年,Itti 和 Koch^[7]通过可运行的计算机程序对 KU 模型进行仿真模拟和验证,并在显著图计算的基础上,利用 WTA (Winner-Takes-All) 和 IOR (Inhibition of Return) 来

扫描视觉信号,生成仿真的动态注视点序列,同时就人类和该程序对视觉场景,心理模式图像等视觉刺激的响应进行了分析比较,这种创新的做法使得本项工作成为 KU 模型提出后的另一个里程碑,目前该文献被他人引用次数已达到 5 062 次(2014 年 3 月)。

1.2 数学及信息论建模方法

从神经计算科学的角度来看,KU 模型只描述了视觉注意的初级阶段,模拟的神经元仅局限于视网膜,侧膝体和视觉初级皮层的简单细胞。由于 KU 模型本身具有的局限性,以及稀疏编码^[22]等新理论的提出,推动了视觉注意研究的新发展。此时研究者们已经不再关注于对人类视觉系统的相关神经处理过程作一一对应的模拟,而是从数学上入手,尝试建立符合注意问题实质并且具有生理上可信性的理论模型或方法。期间则诞生了很多有代表性的工作。

加拿大约克大学的 Bruce 等认为人们倾向于注意那些场景中具有较高不可预测性的区域,而这种不可预测性恰恰可以用该区域相对于场景整体的自信息 (Self-Information) 来度量,即显著度正比于自信息,由此而提出了注意的信息最大化模型^[23-24];与 Bruce 的工作不同,加州大学圣迭亚哥

分校的 Zhang 等^[25]认为信号的显著度可以由其自身的不可预测性(自底向上)和其对目标的预测能力(自顶向下)来共同衡量,并可以由贝叶斯推理计算得出。根据 Zhang 的理论,自底向上的注意可以用输入信号片相对于一个较大的信号片集合的自信息来度量,而自顶向下的注意则可以用该信号片与目标信号片集合的互信息(互信息)来度量。从稀疏编码理论出发,上海交通大学的 Hou 等,认为视觉信息的显著性是以一种动态的方式计算得出的,而且又提出了一种增量编码长度(Incremental Coding Length^[26])的显著性度量准则。同样基于稀疏编码,北京大学的 Wang 等借鉴了人脑认知机理和神经元活动的相关生理事实,提出了点熵率(Site Entropy Rate^[27])的显著性度量,并在之后的工作中,将 SER 成功应用于眼动行为模拟^[28]。

除传统的信息论角度,近期的一些工作还从分类判别力(Discriminability^[29-30])、频域签名(Signature^[31])、空间不相似性(Spatial Dissimilarity^[32])、稀缺性(Rareness^[33])、惊讶度(Surprise^[34-39])等信号的频率及空间属性方面展开探索。加州大学圣迭亚哥分校的 Gao 等^[29-30]认为显著度的计算问题可以转换为对像素进行“显著—不显著”的二分类问题。较强区分能力的特征更适合于参与显著—不显著分类。由此可以将自顶向下的注意建模成一个以目标为训练集的二分类器,该分类器通过训练集选取有判别力的特征进行分类,而自底向上的注意则可建模成一个以周边区域为训练集的二分类器。基于以上想法,Gao 提出了有判别力的中央—外围差分算子(Discriminative Center-Surround Difference)来对视觉信号的显著性进行度量。基于频率域相位分析,加州理工大学的 Hou 等提出了谱残差(Spectral Residual)^[40]和图像签名(Image Signature)^[31]方法,利用傅里叶变换估计出图像中的前景,即引人注意的显著区域。卡尔斯鲁厄理工学院的 Schauerte 等^[41]进一步提出了四元傅里叶变换,对 Hou 的工作进行了理论扩展。不同于以往工作的复杂理论背景,法国蒙斯大学的 Riche^[28]和美国南加州大学的 Borji 等^[42]直接采用底层特征的稀缺性作为自底向上的显著性度量,并在注视点预测实验中取得了极好的效果。Itti 等在 1998 年的针对 KU 模型的工作以后,重新考察了生理心理学中的各种有价值的概念,发现注意的起源很可能是“惊讶”(Surprise)^[35]。Itti 的新理论认为,信息量大(熵值较大)的视觉输入不一定是容易引起注意的区域,相比而言,人们更倾向于关注那些能使其惊讶的地方。进一步地,Itti 提出用先验知识与后验预测之间的 KL 散度来计算“惊讶”,并预测人类在观看视频过程中的注视点。

与上述方法基于启发式的思路不同,本文通过直接对人眼注视点的分布进行统计分析,归纳得到了显著度的超高斯先验,并结合投影追踪技术,发展且提出了能够同时进行眼动行为模拟和信号显著度估计的统计注意模型^[43]。

2 模型评估及对比分析

目前国际主流的评测策略是考察探求模型在预测人眼真实注视点时的 ROC 响应曲线,并通过 ROC 线下区域的大小来对模型的性能进行量化评估。这一评测方法主要考察的是待测模

型生成的显著图与真实的人眼注视点之间的一致性。

2.1 测试数据集与评估指标

本节用于实验的数据集包括:两个图像数据集,分别是加拿大约克大学的 YORK-120^[24]和美国麻省理工大学的 MIT-1003^[44]。由 Bruce 和 Tsotsos^[24]提供的注视点数据库(YORK-120)包含了采集自 20 个志愿者在观察 120 张室内外自然图像的过程中(4 秒一张)产生的 11 999 个注视点。为了减少志愿者个性差异所造成的影响,本文在实验中还通过滤除空间分布上孤立的注视点创建了两个子数据库:YORK-120-SUB-1 和 YORK-120-SUB-2。具体地,首先将数据库中附有的视点分布密度图正则化至 $[0,1]$ 区间来量化描述注视点的空间聚集度,而后通过阈值化操作来构建具有不同聚集度的子数据库。YORK-120-SUB-1 包含了 8 190 个密度值大于 0.2 的注视点,而 YORK-120-SUB-2 则包含了 4 339 个密度值大于 0.5 的注视点。由 Judd 提供的 MIT-1003^[44]数据库包含了采集自 LabelMe 的 1 003 张风景和人物图像。为了和有监督的方法进行公正的比较,本节使用从 MIT-1003 随机抽取的子集作为测试数据库,每个子集包含了不重叠的 100 张图像。与对 YORK-120 的处理类似,本测试仅使用 MIT-1003 中每个志愿者在各张图像上的前 6 个注视点,以保证视点在空间分布上的一致性。

本实验采用 ROC 线下区域(Area Under ROC Curve,简称 AUC)和 KL 散度(Kullback-Leibler divergence,简称 KL^[24-25,35])作为模型的量化性能度量。正如以往的工作所指出的那样^[7, 23, 25-27, 40, 44-45],不同的人的观察同一场景的过程中可能会产生完全不同的注视点序列,而同一个人在不同的上下文环境下观察同一场景时产生的序列也可能会有不同,这种序列模式在人与人之间、不同上下文之间的差异性使得直接对眼动序列进行对比评测变得极其困难。因此,为了保证评测的通用性和公正性,本文采用了被学界广泛接受和使用的 AUC 和 KL 评价体系对提出的模型以及基线方法进行评估。原始的 AUC 和 KL 评测策略依据注视点的位置在显著图上采集正例样本,同时在随机位置上采集负例样本,进而计算得到 AUC 和 KL 指标。传统的评测策略在很大程度上受到“边缘效应”的影响,因为在人工拍摄的大多数自然图像中,显著的目标往往安置在画面的中心区域(即所谓的中心偏置,Center Bias),这就导致评测规则偏好于中心显著性强而边缘显著性弱的显著图。加州大学圣迭亚哥分校的 Zhang 等人指出,一个简单的置于图像中心的高斯团块,可以在传统的 AUC 评测中得到 0.80 的高分(基于 YORK-120 数据库)。虽然这种“显著图”完全不是根据图像内容所计算得出,但却可以在传统 AUC 评价体系中击败领域中提出的大多数模型。为了消除由“边缘效应”所带来的干扰,本节采用了 Zhang^[25]设计的一种改进的基于随机排列的评测策略来计算 AUC 和 KL 指标,该评价策略使用非同场景的注视点作为抽样集来构建负例样本集合,可以从根本上消除中心偏置的影响。具体地,随机排列的迭代次数设为 100,用于 KL 散度计算的概率密度分布函数则采选 16 个 Bin 的直方图来表示。

2.2 性能对比分析

图 1 和图 2 展示了本文作者发表在文献[46]中的基于 YORK-120 和 MIT-1003 上的大规模对比实验结果。

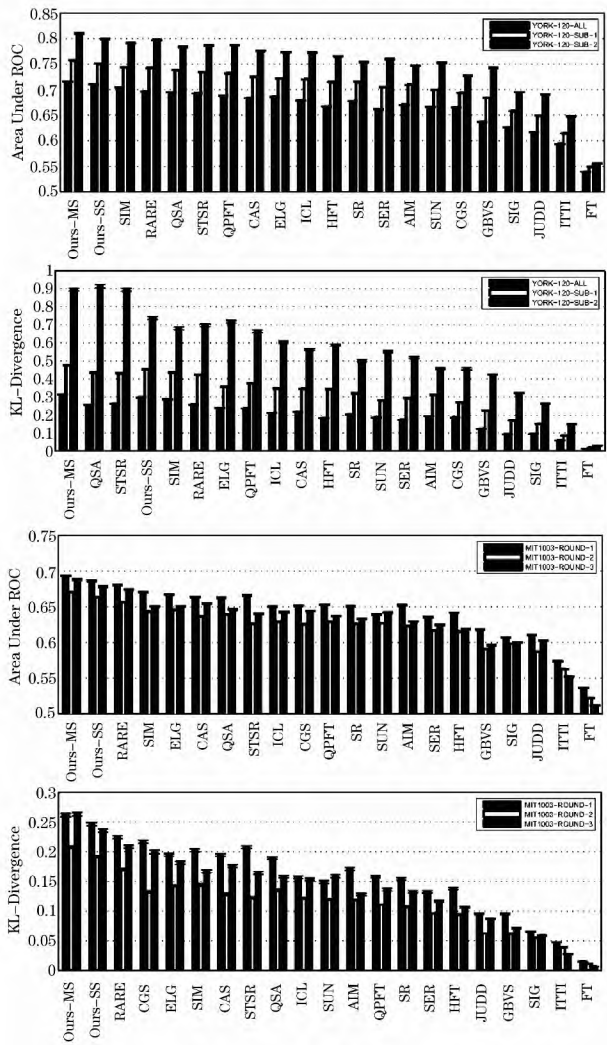


图 1 模型的 AUC 和 KL 性能排序

Fig.1 AUC and KL performance of the tested models

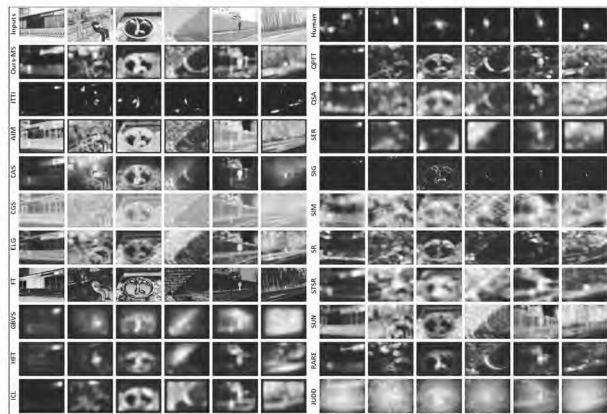


图 2 模型的显著度检测结果视觉对比

Fig.2 Visual comparisons of saliency

检测 results from the tested models

图 1 以从高到低的顺序直观展示了表 1 中所有测试模型的 SL-AUC 和 SL-KL 指标。其中本文作者提出的 SGP 模型分为单尺度 (Ours-SS) 和多尺度 (Ours-MS) 两种方案。图 2 展示一些视觉对比结果,包括测试数据库的示例图像以及各个模型所生成的显著图。可以看出 SGP^[43] 和 RARE^[33] 这两种基于统计的模型在测试数据库上均能取得非常好的效果。

3 结束语

视觉“显著性”是人类视觉智能中一个十分重要的认知属性。本文通过人眼视点分类实验对 20 种国际主流模型进行了量化的综合对比。结果表明,相对于传统方法,基于统计的模型在性能表现上有着精度高、稳定性强等明确的优越性。

参考文献:

[1] ITTI L. Automatic foveation for video compression using a neurobiological model of visual attention[C]//Image Processing, IEEE Transactions on, 2004, 13(10):1304 - 1318.

[2] SUN X, YAO H, JI R, et al. Photo assessment based on computational visual attention model[C]// ACM Multimedia, 2009:541 - 544.

[3] SIAGIAN C, ITTI L. Rapid biologically - inspired scene classification using features shared with visual attention[C]//IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007:300 - 312.

[4] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency - tuned salient region detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2009:1597 - 1604.

[5] ELAZARY L, ITTI L. A bayesian model of visual search and recognition[C]//Proc. Vision Science Society Annual Meeting (VSS08), 2008.

[6] GOFEMAN S, ZELNIK - MANOR L, TAL A. Context - aware saliency detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10):1915 - 1926.

[7] 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.

[8] ITTI L, KOCH C. Computational modelling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2(3):194 - 203.

[9] TSOTSOS J, CULHANE S, KEI W W, et al. Modeling visual attention via selective tuning[J]. Artificial intelligence, 1995, 78(1):507 - 545.

[10] OLIVA T A C M, HENDERSON J A. Top - down control of visual attention in object detection[C]// Proceedings of International Conference on Image Processing (ICIP2003). Barcelona, Catalonia, 2003: 253 - 256.

[11] BORJI A, ITTI L. State - of - the - art in visual attention modeling [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013, 35(1):185 - 207.

[12] Yarbus A. Eye movements and Vision[M]. Plenum Press, 1967.

[13] FOULSHAM T, UNDERWOOD G. What can saliency models predict about eye movements? spatial and sequential aspects of fixations during encoding and recognition[J]. Journal of Vision, 2008, 8(2).

[14] HAYHOE M, BALLARD D. Eye movements in natural behavior [J]. Trends in cognitive sciences, 2005, 9(4):188 - 194.

[15] HAYHOE M, SHRIVASTAVA A, MRUCZEK R, et al. Visual memory and motor planning in a natural task[J]. Journal of Vision,

- 2003, 3(1).
- [16] LAND M, MENNIE N, RUSTED J, et al. The roles of vision and eye movements in the control of activities of daily living[J]. PERCEPTION – LONDON, 1999, 8(11):1311 – 1328.
- [17] TATLER B, HAYHOE M, LAND M, et al. Eye guidance in natural vision: reinterpreting salience[J]. Journal of vision, 2011, 11(5).
- [18] FOERSTER R, CARBONE E, KOESLING H, et al. Saccadic eye movements in a high – speed bimanual stacking task: Changes of attentional control during learning and automatization[J]. Journal of Vision, 2011, 11(7).
- [19] RENNINGER L, VERGHESE P, COUGHLAN J. Where to look next? eye movements reduce local uncertainty[J]. Journal of Vision, 2007, 7(3).
- [20] M. T A, GARRY G. A feature – integration theory of attention[J]. Cognitive Psychology, 1980, 12(1):97 – 136.
- [21] KOCH C, ULLMAN S. Shifts in selective visual attention: towards the underlying neural circuitry[J]. Human neurobiology, 1985, 4(4):219.
- [22] BELL A, SEJNOWSKI T. The “independent components” of natural scenes are edge filters[J]. Vision research, 1997, 37(23):3327 – 3338.
- [23] BRUCE N, TSOTSOS J. Saliency, attention, and visual search: an information theoretic approach[J]. Journal of Vision, 2009, 9(3):1 – 24.
- [24] BRUCE N, TSOTSOS J. Saliency based on information maximization[J]. Advances in Neural Information Processing Systems, (NIPS), 2006: 155 – 162.
- [25] ZHANG L, TONG M, MARKS T, et al. SUN: a bayesian framework for saliency using natural statistics[J]. Journal of Vision, 2008, 8(7):1 – 20.
- [26] HOU X, ZHANG L. Dynamic visual attention: searching for coding length increments[J]. Advances in Neural Information Processing Systems, (NIPS), 2008:681 – 688.
- [27] WANG W, WANG Y, HUANG Q, et al. Measuring visual saliency by site entropy rate[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010:2368 – 2375.
- [28] WANG W, CHEN C, WANG Y, et al. Simulating human saccadic scanpaths on natural images[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011:441 – 448.
- [29] GAO D, VASCONCELOS N. Discriminant saliency for visual recognition from cluttered scenes[J]. Advances in neural information processing systems, 2005, 17:481 – 488.
- [30] GAO D, MAHADEVAN V, VASCONCELOS N. The discriminant center – surround hypothesis for bottom – up saliency[J]. Neural Information Processing Systems (NIPS), 2007:1 – 8.
- [31] HOU X, HAREL J, KOCH C. Image signature: highlighting sparse salient regions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(1):194 – 201.
- [32] DUAN L, WU C, MIAO J, et al. Visual saliency detection by spatially weighted dissimilarity[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011: 473 – 480.
- [33] RICHEL N, MANCAS M, DUVINAGE M, et al. RARE2012: A multi – scale rarity – based saliency detection with its comparative statistical analysis[J]. Signal Processing: Image Communication, 2013.
- [34] ITTI L, BALDI P F. A surprise theory of attention[C]//Proc. Vision Science Society Annual Meeting (VSS05), 2005.
- [35] ITTI L, BALDI P F. Bayesian surprise attracts human attention [C]//Advances in Neural Information Processing Systems, (NIPS), 2006:547 – 554.
- [36] ITTI L, BALDI P F. Bayesian surprise attracts human attention[J]. Vision Research, 2008.
- [37] BERG D J, BOEHNKE S E, MARINO R A, et al. Characterizing Surprise in Humans and Monkeys[C]// Proc. Vision Science Society Annual Meeting (VSS06), 2006.
- [38] Byrne E. Surprise moves eyes. Primary Visual Cortex, 2008.
- [39] MUNDHENK T N, EINHAUSER W, ITTI L. Automatic computation of an image’s statistical surprise predicts performance of human observers on a natural image detection task[J]. Vision Research, 2009.
- [40] HOU X, ZHANG L. Saliency detection: A spectral residual approach[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2007:1 – 8.
- [41] SCHAUERTE B, STIEFELHAGEN R. Quaternion – based spectral saliency detection for eye fixation prediction[C]// ECCV, 2012: 116 – 129.
- [42] BORJI A, ITTI L. Exploiting local and global patch rarities for saliency detection[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012:478 – 485.
- [43] SUN Xiaoshuai, YAO Hongxun, JI Rongrong. What are we looking for: towards statistical modeling of saccadic eye – movement and visual saliency[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012:1552 – 1559.
- [44] JUDD T, EHINGER K, DURAND F, et al. Learning to predict where humans look [C]//IEEE 12th International Conference on Computer Vision, 2009: 106 – 2113.
- [45] MURRAY N, VANRELL M, OTAZU X, et al. Saliency estimation using a non – parametric low – level vision model[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011: 433 – 440.
- [46] SUN Xiaoshuai, YAO Hongxun, JI Rongrong Ji, et al. Towards statistical modeling of saccadic eye – movement and visual saliency [C]//IEEE Transactions on Image Processing, July 5, 2014. In Press.
- [47] ITTI L, KOCH C. Computational modelling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2(3):194 – 203.
- [48] TORRALBA A, CASTELHANO M S, OLIVA A, et al. Contextual guidance of eye movements and attention in real – world scenes: the role of global features in object search[J]. Psychological Review, 2006, 113:766 – 786.
- [49] HAREL J, KOCH C, PERONA P. Graph – based visual saliency [C]// Advances in neural information processing systems, 2006: 545 – 552.
- [50] GUO C, MA Q, ZHANG L. Spatio – temporal saliency detection using phase spectrum of quaternion fourier transform[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008: 1 – 8.
- [51] SEO H, MILANFAR P. Nonparametric bottom – up saliency detection by self – resemblance[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops, 2009: 45 – 52.
- [52] LI J, LEVINE M D, AN X, et al. Visual saliency based on scale – space analysis in the frequency domain, 2013.

作者: 孙晓帅, 姚鸿勋, [SUN Xiaoshuai](#), [YAO Hongxun](#)
作者单位: [哈尔滨工业大学 计算机科学与技术学院, 哈尔滨, 150001](#)
刊名: [智能计算机与应用](#)
英文刊名: [Intelligent Computer and Applications](#)
年, 卷(期): 2014, 4(5)

参考文献(52条)

1. ITTI L [Automatic foveation for video compression using a neurobiological model of visual attention](#) 2004
2. SUN X;YAO H;JI R [Photo assessment based on computational visual attention model](#) 2009
3. SIAGIAN C;ITTI L [Rapid biologically-inspired scene classification using features shared with visual attention](#) 2007
4. ACHANTA R;HEMAMI S;ESTRADA F [Frequency-tuned salient region detection](#) 2009
5. ELAZARY L;ITTI L [A bayesian model of visual search and recognition](#) 2008
6. GOFEMAN S;ZELNIK-MANOR L;TAL A [Context-aware saliency detection](#) 2012(10)
7. ITTI L;KOCH C;NIEBUR E [A model of saliency-based visual attention for rapid scene analysis](#) 1998(11)
8. ITTI L;KOCH C [Computational modelling of visual attention](#) 2001(03)
9. TSOTSOS J;CULHANE S;KEI W W [Modeling visual attention via selective tuning](#) 1995(01)
10. OLIVA T A C M;ENDERSON J A [Top-down control of visual attention in object detection](#) 2003
11. BORJI A;ITTI L [State-of-the-art in visual attention modeling](#) 2013(01)
12. Yabus A;Yabus A [Eye movements and Vision](#) 1967
13. FOULSHAM T;UNDERWOOD G [What can saliency models predict about eye movements? spatial and sequential aspects of fixations during encoding and recognition](#) 2008(02)
14. HAYHOE M;BALLARD D [Eye movements in natural behavior](#) 2005(04)
15. HAYHOE M;SHRIVASTAVA A;MRUCZEK R [Visual memory and motor planning in a natural task](#) 2003(01)
16. LAND M;MENNIE N;RUSTED J [The roles of vision and eye movements in the control of activities of daily living](#) 1999(11)
17. TATLER B;HAYHOE M;LAND M [Eye guidance in natural vision:reinterpreting salience](#) 2011(05)
18. FOERSTER R;CARBONE E;KOEHLING H [Saccadic eye movements in a high-speed bimanual stacking task:Changes of attentional control during learning and automatization](#) 2011(07)
19. RENNINGER L;VERGHESE P;COUGHLAN J [Where to look next? eye movements reduce local uncertainty](#) 2007(03)
20. M. T A, GARRY G [A feature-integration theory of attention](#) 1980(01)
21. KOCH C;ULLMAN S [Shifts in selective visual attention:towards the underlying neural circuitry](#) 1985(04)
22. BELL A;SEJNOWSKI T [The "independent components" of natural scenes are edge filters](#) 1997(23)
23. BRUCE N;TSOTSOS J [Saliency, attention, and visual search:an information theoretic approach](#) 2009(03)
24. BRUCE N;TSOTSOS J [Saliency based on information maximization](#) 2006
25. ZHANG L;TONG M;MARKS T SUN;a bayesian framework for saliency using natural statistics 2008(07)

26. [HOU X;ZHANG L Dynamic visual attention:searching for coding length increments](#) 2008
27. [WANG W;WANG Y;HUANG Q Measuring visual saliency by site entropy rate](#) 2010
28. [WANG W;CHEN C;WANG Y Simulating human saccadic scanpaths on natural images](#) 2011
29. [GAO D;VASCONCELOS N Discriminant saliency for visual recognition from cluttered scenes](#) 2005
30. [GAO D;MAHADEVAN V;VASCONCELOS N The discriminant center-surround hypothesis for bottom-up saliency](#) 2007
31. [HOU X;HAREL J;KOCH C Image signature:highlighting sparse salient regions](#) 2012(01)
32. [DUAN L;WU C;MIAO J Visual saliency detection by spatially weighted dissimilarity](#) 2011
33. [RICHE N;MANCAS M;DUVINAGE M RARE2012:A multi-scale rarity-based saliency detection with its comparative statistical analysis](#) 2013
34. [ITTI L;BALDI P F A surprise theory of attention](#) 2005
35. [ITTI L;BALDI P F Bayesian surprise attracts human attention](#) 2006
36. [ITTI L;BALDI P F Bayesian surprise attracts human attention](#) 2008
37. [BERG D J;BOEHNKE S E;MARINO R A Characterizing Surprise in Humans and Monkeys](#) 2006
38. [Byrne E Surprise moves eyes](#) 2008
39. [MUNDHENK T N;EINHAEUSER W;ITTI L Automatic computation of an image' s statistical surprise predicts performance of human observers on a natural image detection task](#) 2009
40. [HOU X;ZHANG L Saliency detection:A spectral residual approach](#) 2007
41. [SCHAUERTE B;STIEFELHAGEN R Quaternion-based spectral saliency detection for eye fixation prediction](#) 2012
42. [BORJI A;ITTI L Exploiting local and global patch rarities for saliency detection](#) 2012
43. [SUN Xiaoshuai;YAO Hongxun;JI Rongrong What are we looking for:towards statistical modeling of saccadic eye-movement and visual saliency](#) 2012
44. [JUDD T;EHINGER K;DURAND F Learning to predict where humans look](#) 2009
45. [MURRAY N;VANRELL M;OTAZU X Saliency estimation using a non-parametric low-level vision model](#) 2011
46. [SUN Xiaoshuai;YAO Hongxun;JI Rongrong Ji Towards statistical modeling of saccadic eye-movement and visual saliency](#) 2014
47. [ITTI L;KOCH C Computational modelling of visual attention](#) 2001(03)
48. [TORRALBA A;CASTELHANO M S;OLIVA A Contextual guidance of eye movements and attention in real-world scenes:the role of global features in object search](#) 2006
49. [HAREL J;KOCH C;PERONA P Graph-based visual saliency](#) 2006
50. [GUO C;MA Q;ZHANG L Spatio-temporal saliency detection using phase spectrum of quatemion fourier transform](#) 2008
51. [SEO H;MILANFAR P Nonparametric bottom-up saliency detection by self-resemblance](#) 2009
52. [LI J;LEVINE M D;AN X Visual saliency based on scale space analysis in the frequency domain](#) 2013