

# 协同视觉显著性检测方法综述

钱晓亮<sup>1</sup>, 白臻<sup>1</sup>, 陈渊<sup>1</sup>, 张鼎文<sup>2</sup>, 史坤峰<sup>1</sup>,  
王芳<sup>1</sup>, 吴青娥<sup>1</sup>, 毋媛媛<sup>1</sup>, 王慰<sup>1</sup>

(1. 郑州轻工业学院电气信息工程学院, 河南郑州 450002; 2. 西安电子科技大学机电工程学院, 陕西西安 710071)

**摘要:** 协同视觉显著性检测是视觉注意力计算领域中一个快速发展的新兴分支, 致力于检测多幅相关场景图像中的公共显著目标, 在各种计算机视觉任务中有广泛应用. 考虑到特征提取策略的设计是协同视觉显著性检测当前研究的重点, 本文首先根据特征提取策略的不同对现有的协同视觉显著检测方法进行了分类介绍和定性分析. 其次, 通过在5个公开数据库上的主观和定量对比, 对各流行算法的性能进行了评估, 分析了特征提取策略对算法性能的影响以及各数据库的复杂度, 并验证了协同视觉显著性检测和视觉显著性检测的区别. 最后, 对本文工作进行了总结, 并对当前研究中存在的问题和未来的研究工作进行了讨论.

**关键词:** 视觉注意力; 协同视觉显著性; 特征提取策略; 手工特征; 浅层学习特征; 深度学习特征

**中图分类号:** TP391.4      **文献标识码:** A      **文章编号:** 0372-2112 (2019)06-1352-14

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2019.06.024

## A Review of Co-saliency Detection

QIAN Xiao-liang<sup>1</sup>, BAI Zhen<sup>1</sup>, CHEN Yuan<sup>1</sup>, ZHANG Ding-wen<sup>2</sup>, SHI Kun-feng<sup>1</sup>,  
WANG Fang<sup>1</sup>, WU Qing-e<sup>1</sup>, WU Yuan-yuan<sup>1</sup>, WANG Wei<sup>1</sup>

(1. School of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan 450002, China;

2. School of Mechano-Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China)

**Abstract:** Co-saliency detection is a new branch with the rapid development in the field of visual attention, which concerns the detection of the common salient objects from multiple relevant scene images, and can be widely used in various computer vision tasks. Considering the key point of current research is the design of feature extraction strategy, the existing co-saliency detection methods are firstly summarized and qualitatively analyzed according to the different feature extraction strategies in this paper. Subsequently, based on the subjective and quantitative comparisons in the five open datasets, the performance of the state-of-the-art algorithms is evaluated, the influence of the feature extraction strategy on the performance of algorithms and the complexity of the datasets is analyzed, and the difference of co-saliency detection and saliency detection is also verified. Finally, the conclusion of this paper are presented, the problems of current research and the future development are also discussed.

**Key words:** visual attention; co-saliency; feature extraction strategy; hand-designed features; shallow learning features; deep learning feature

## 1 引言

随着大数据时代的到来, 大量图像和视频类的数字化信息充斥在网络和我们的生活中, 在众多计算机视觉任务中, 往往需要在大量图像中快速获取其中的有效信息<sup>[1]</sup>. 协同视觉显著性检测 (Co-saliency Detec-

tion) 致力于检测多幅相关场景图像中的公共显著目标, 一次性过滤多幅图像中的大量无效数据, 应用在各种计算机视觉任务中, 可以显著降低后续处理的计算量. 随着协同视觉显著性检测研究的不断深入, 其应用拓展的研究也是当前的热门, 如: 协同分割<sup>[2,3]</sup>、协同识别<sup>[4]</sup>和图像检索<sup>[5,6]</sup>等.

收稿日期: 2018-05-02; 修回日期: 2018-08-02; 责任编辑: 孙瑶

基金项目: 国家自然科学基金 (No. 61501407, No. 61603350, No. 61703373); 河南省高校科技创新团队 (No. 19IRTSTHN013); 河南省高等学校重点科研项目 (No. 19A413014, No. 16A413017); 郑州轻工业学院博士基金 (No. 2014BSJJ016, No. 2015BSJJ004, No. 2017BSJJ008)

协同视觉显著性检测是视觉注意力计算领域中一个快速发展的新兴分支,经过近几年的研究发展,定义和应用场合也发生了一定变化.协同视觉显著性检测最早由 Jacobs 等在文献[7]中提出,早期的协同视觉显著性检测是在几乎相同的场景中,利用两幅图像之间的显著对象引起的局部结构变化检测同一目标,适应性比较差.之后,Chen 等<sup>[8]</sup>将协同视觉显著性检测定义为检测两幅图像中相似的显著性目标,Chen, Li 等<sup>[9]</sup>利用两幅图像之间的公共信息来检测协同显著对象.2013 年, Li, Fu 等<sup>[10,11]</sup>提出在多幅图像中发现协同显著性目标, Li 等在文献[11]中利用多尺度分割和图像块相似性排序等方法,按照图像匹配顺序获得协同显著图.2016 年, Jerripothula 等<sup>[12]</sup>把协同视觉显著性检测应用在视频(连续多帧的图像)中,对协同显著性目标进行协同定位.之前的大多数方法都是在直接给定相似目标的图像组上进行协同显著性目标的检测,而 Yao 等<sup>[13]</sup>提出一种如何在一堆图像中分离出具有相似目标的图像组的检测方法.

从上述介绍中可以看出,检测出的协同显著区域必须满足以下两个要求:(1)多幅图像的协同显著区域在每幅图像中必须是显著的;(2)来自不同图像的协同显著区域必须有相似的特点<sup>[14]</sup>.也就是说,协同视觉显著性检测的目的是寻找多幅图像的显著目标区域的交集<sup>[15]</sup>.

满足以上两个要求的协同视觉显著性检测方法可分为两类:一类是根据协同显著目标间的特征相似性和一致性学习得到协同显著图<sup>[8,16,17]</sup>,例如 Chen 等<sup>[8]</sup>基于分布式表示塑造两幅图像的稀疏特征,利用 K-L 散度(交叉熵)来计算图像块间相似性,从而表示图像的协同显著性.另一类是采用多特征和多显著图融合的方式得到协同显著图<sup>[9-11,18-20]</sup>.Liu 等<sup>[20]</sup>通过融合局部相似度、全局相似度和先验目标来检测协同显著对象.Li 等<sup>[9]</sup>通过线性融合多种特征描述符得到内部显著图(Intra-saliency Map)和交叉显著图(Inter-saliency Map),两者融合获得最终的协同显著图(Co-saliency Map).

以上两类协同视觉显著性检测方法都涉及两个环节:单幅图像的特征提取(单幅图像内部显著图的计算也依赖特征提取)和多幅图像特征之间的相似性度量;相似性度量算法的研究目前已相对成熟,当前研究的重点是特征提取策略的设计.

本文从特征提取策略的角度出发,对现有的协同视觉显著性检测方法进行了归类分析和性能评估,主要贡献如下:

(1)依据特征提取策略的不同,对现有的协同视觉显著性检测方法进行分类总结和定性分析;

(2)对流行的协同视觉显著性检测算法进行主观、

检测精度和时间复杂度的综合实验对比,并分析了特征提取策略对算法性能的影响;

(3)将协同视觉显著性检测和普通视觉显著性检测算法进行了主观、精测精度和算法复杂度的综合实验对比,验证了协同视觉显著性检测与普通显著性检测的区别;

(4)对参与实验对比的 5 个数据库进行复杂度分析,并对当前数据库规模不足的问题提出了两个可能的解决方案.

## 2 相关工作

视觉显著性检测和协同视觉显著性检测是视觉注意力计算领域的两个分支,都以视觉显著图作为输出结果,其中,视觉显著性检测包括人眼注意点预测和显著目标检测两种.协同视觉显著性的研究起步相对较晚(2010 年),人眼注意点预测和显著目标检测的研究方法对其有很强的借鉴意义.

### 2.1 人眼注意点预测

人眼注意点预测的任务是寻找图像中能引起人眼注意的像素点,输出的显著图以显著程度的量化值来表示<sup>[21]</sup>.

ITTI 等提出的基于特征融合的人眼注意点预测模型,在该领域具有里程碑的意义,ITTI 等将图像用高斯金字塔进行表示,计算每层各底层特征的中心外围对比度,将这些特征各自正则化后叠加得到各层的特征图,最后将各层的特征图也分别正则化后叠加从而得到最终的显著图. Judd 等<sup>[22]</sup>利用眼动仪建立了一个带标签的数据库,并利用它训练得到一个基于 SVM 分类器的显著性预测模型. Li 等<sup>[23-25]</sup>提出了基于信息论的检测方法,将信息量的大小来衡量各像素点的显著性大小.近几年,基于深度学习的方法<sup>[26-29]</sup>取得了不错的效果,例如, Marcella 等<sup>[29]</sup>结合卷积神经网络和堆叠自动编码器来进行人眼注意点预测.此外,还有基于频域分析<sup>[30,31]</sup>和基于图论<sup>[32,33]</sup>的方法.

### 2.2 显著目标检测

显著目标检测的任务是寻找图像中能引起人眼注意的前景目标区域,输出的显著图要能均匀的渲染整个目标体并准确的界定目标的边界<sup>[34]</sup>.

Liu 等<sup>[35]</sup>提出的基于条件随机场(Conditional Random Field, CRF)的方法是该方向的开创性工作,文献[35]通过 CRF 将局部和全局区域进行融合得到显著图. Cheng 等<sup>[36]</sup>提出的基于全局对比度的显著目标检测方法也是该方向里程碑工作.采用机器学习的方法进行显著目标检测是该方向的主流<sup>[37-41]</sup>, Li 等<sup>[37]</sup>利用稀疏编码来进行显著目标检测, Jiang 等<sup>[38]</sup>采用超像素分割算法对图像进行多尺度分割并对各区域进行特征提取,通过随机

森林将特征融合得出显著图. Tong 等<sup>[40]</sup>构造一个基于图像先验知识的弱显著图,生成训练样本来训练分类器,利用分类器来检测显著性. Han 等<sup>[41]</sup>将图像四个边框区域视作背景,采用堆叠去噪自动编码器( Stacked Denoising Autoencoders, SDAE)来学习背景结构,利用各区域在 SDAE 上的重构误差来衡量其显著程度.

人眼注意点预测、显著目标检测和协同视觉显著性检测的检测结果主观对比如图 1 所示. 原始图像是两幅场景类似的图像,人眼关注点预测算法将图中所有的球员(显著目标)都检测出来,但检测结果比较模糊,只是一种大致的判断,显著目标检测算法则对球员的整体区域界定较为精确,协同显著性检测只将两幅图像中穿红色球衣的利物浦球员(公共显著目标)检测出来.

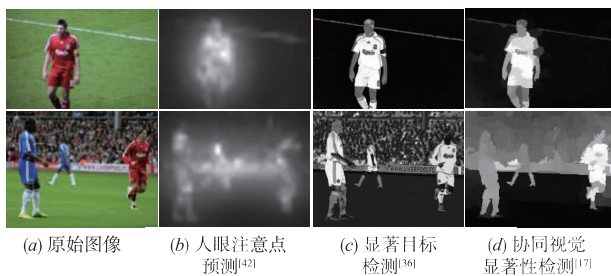


图1 视觉注意力计算3个分支领域的检测结果主观对比

### 3 基于特征提取策略的协同视觉显著性检测方法分析

本节依据特征提取策略的不同将现有协同视觉显著性检测方法分为三类:①基于手工特征的检测方法;②基于浅层学习特征的检测方法;③基于深度学习特征的检测方法. 下面对三类方法中的经典工作进行简要介绍.

#### 3.1 基于手工特征的协同视觉显著性检测方法

现有的协同视觉显著性检测方法大部分都采用手工特征. 常用的手工特征有:颜色直方图( Color Histogram, CH)、梯度直方图( Histograms Of Gradients, HOG)、尺度不变特征转换描述符 SIFT( Scale-Invariant Feature Transform)和深度描述符 DSP( Depth Shape Prior)等.

采用 CH 的协同视觉显著性检测方法包括文献[9, 10, 20, 43]等. 其中, Fu 等<sup>[9]</sup>将提取到的 Lab 颜色空间的 CH 特征和纹理特征进行聚类, Fu 等认为多幅图像协同区域的聚类中心距离较小,因此利用特征聚类中心的距离进行相似性度量,从而得到协同显著图. Li 等<sup>[9]</sup>利用 ITTI<sup>[42]</sup>、FT<sup>[44]</sup>和 SR<sup>[45]</sup>的算法分别获取显著图,并通过线性融合得到各图像的内部显著图,再利用 SimRank 相似性计算各幅图像的 CH 特征和纹理直方图特征的相似度,将相似性度量结果和各图像的单幅显著图做线性融合得到协同显著图.

采用 HOG 特征的协同视觉显著性检测方法包括<sup>[46, 47]</sup>等. Song 等<sup>[46]</sup>先利用文献[48]的算法获取各幅图像的内部显著图,将内部显著图上分割得到的候选目标区域作为聚类单元<sup>[49]</sup>, Song 等认为不同场景起关键作用的特征不同,所以采用 HOG 特征和颜色空间特征等构建特征集后,基于多个任意组合的子特征集对候选目标区域进行聚类,并根据聚类评价准则自适应计算特征的权重,将基于多聚类结果生成的弱协同显著图进行线性融合获取检测结果. Huang 等<sup>[47]</sup>在 HOG 和其他手工提取特征的基础上,构建多尺度超像素金字塔,结合高斯混合模型 GMM 生成协同显著图.

Ye 等<sup>[50]</sup>采用 SIFT 特征进行协同视觉显著性检测方法. Ye 等利用 SIFT 和颜色空间特征对整个图像组进行聚类分析,获取在各幅图像中多次出现的显著目标并将其作为候选目标区域,最后在图像组上进行基于特征的区域匹配获取最终检测结果.

此外, Du 等<sup>[51]</sup>提出一种利用 GIST 特征进行对比度分析的协同视觉显著性检测方法. Cong 等<sup>[52]</sup>提出基于 DSP 特征的协同视觉显著性检测方法,并结合图像相似性线索检测 RGBD 图像组的协同显著区域. Zhang 等<sup>[53]</sup>提出基于结构特征的协同视觉显著检测方法,并结合细胞自动机获得区域相似性,从而检测协同显著区域.

#### 3.2 基于浅层学习特征的协同视觉显著性检测方法

浅层特征提取是相对深度特征提取而言,可以视作一种单隐层的神经网络特征提取方法. 协同视觉显著性检测常采用的浅层学习特征提取算法包括:稀疏主成分分析( Sparse Principal Component Analysis, SP-CA)、独立成分分析( Independent Component Analysis, ICA)和稀疏编码( Sparse Coding)等.

Shen 等<sup>[54]</sup>提出一种基于 SPCA<sup>[55]</sup>的协同视觉显著性检测方法. 该方法将均匀分割后的图像块当作计算单元,使用截断幂( Truncated Power) SPCA 方法对底层特征进行稀疏主特征提取. 在稀疏主特征聚类结果的基础上进行 3 种(对比特征、位置优先特征、全局分布特征)显著特征权值的计算. 最后,将通过特征融合生成的单幅图像显著图和多幅图像显著图进行组合,生成协同显著图.

Chen 等<sup>[8]</sup>提出一种基于 ICA 的协同视觉显著性检测方法. 文献[8]采用 ICA 从随机选取的大量自然场景图像块中学习出词典,再利用词典对图像进行滤波得到图像的特征矩阵,将特征矩阵各列(对应各特征)进行归一化得到相应的特征概率分布,然后逐列计算图像间的 K-L 散度并进行求导,散度下降明显的行(特征)属于协同显著性检测区域的特征,基于协同显著性检测区域特征对图像滤波响应得到的值即为协同显著性值.

Chen 等<sup>[56]</sup>提出一种基于稀疏编码<sup>[57]</sup>的协同视觉显著性检测方法.文献[56]先使用文献[8]的方法预测协同显著目标区域的大致位置,把显著值较大区域覆盖的图像块作为训练样本,利用 K-SVD 算法学习出稀疏词典并对图像进行稀疏编码,最后根据编码的重构误差来量化协同显著性.

### 3.3 基于深度学习特征的协同视觉显著性检测方法

2006 年,Hinton 等在文献[58]中提出了深度学习的概念,2015 年,Zhang 等<sup>[59]</sup>首次把深度学习应用在协同视觉显著性检测领域.现有的基于深度学习特征的协同视觉显著性检测方法主要采用受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)、堆叠降噪自动编码器(Stacked Denoising Autoencoders, SDAE)和卷积神经网络(Convolutional Neural Network, CNN)等深度网络.

Zhang 等<sup>[60]</sup>提出基于 SDAE 的协同视觉显著性检测方法,Zhang 等在辅助的数据库中训练 SDAE,通过 SDAE 提取的特征计算各图像的内部显著图,并作为计算交叉显著性的先验知识,再使用深度残差网络进行多图像显著性挖掘,得到多幅图像的交叉显著图,最后与各图像的内部显著图相融合,获得协同显著图.

Zhang 等<sup>[17]</sup>采用 CNN 特征对任意组合的图像组进行协同视觉显著性检测.首先,采用多样例学习算法(Multiple Instance Learning, MIL)判别输入的任意图像是否为协同显著性检测的相关场景图像,去除不相关图像,并进一步判别图像中的超像素块是否属于协同显著目标;随后,采用自步学习算法(Self-Paced Learning, SPL),把多样例-自步学习模型表示转化为标准加权 SVM 问题,使用 CNN 提取出的特征训练协同目标检测器,检测出协同显著目标.此外,Jeong 等<sup>[61]</sup>使用 VGG16<sup>[62]</sup> CNN 模型来提取高层次特征,并形成像素级显著图.

Han 等在文献[63]中提出基于度量学习的协同显著性检测方法.Han 等先采用在 ImageNet 上预训练好的 CNN 提取特征,然后将度量学习正则项嵌入到 SVM 训练中,通过迭代训练同时完成判别特征性学习和协同显著性判别.

Zhang 等<sup>[59]</sup>提出基于 RBM 和 CNN 相结合的协同视觉显著性检测方法.Zhang 等在预训练好的 CNN 层间加入全连接的受限玻尔兹曼机,利用协同显著性数据库对其训练以提取适合当前任务的深度特征,最后将聚类算法与贝叶斯框架结合,计算得到像素级的协同显著值.

### 3.4 基于特征提取策略的定性分析

手工特征一般只针对某几类特定应用,泛化能力较弱.此外,手工特征提取策略的设计对设计者的专业知识

要求较高,而且很难利用大数据的优势.因此,基于手工特征的协同视觉显著性检测方法的性能上限不高.

浅层学习特征和深度学习特征均属于数据驱动特征,可以从大量数据中自动学习到具有代表性的特征表示,相比手工特征具有更强的泛化能力,基于数据驱动特征的协同视觉显著性检测方法也具有更好的性能<sup>[53]</sup>.

深度学习模型可以从原始数据中提取出层次化的特征表达<sup>[64,65]</sup>,层次越深,对数据本质刻画越深刻.而浅层学习模型则相当于单隐层的深度学习模型,例如,稀疏编码学习得到的字典与堆叠自动编码器(Stacked Autoencoders, SAE)第一层权重或 CNN 第一层卷积核的可视化图像比较类似,与人类视觉皮层 V1 区的视觉感受野形状相似<sup>[66]</sup>.因此,相比深度学习特征,浅层学习特征难以提取图像中的高级语义特征,相应的,基于浅层学习特征的协同视觉显著性检测方法的性能也弱于基于深度学习特征的方法.

深度神经网络(Deep Neural Network, DNN)可分为“一维”DNN 和“二维”DNN,前者的输入为一维向量,例如:SAE,后者的输入为二维图像(或是三通道的彩色图像),例如:CNN.对于计算机视觉的各种任务,“二维”DNN 相比“一维”DNN 具有较为明显的两点优势.

(1)“一维”DNN 需要将图像/图像块拉成列向量再输入 DNN 中,损失了图像的空间结构信息,“二维”DNN 则不存在此问题.

(2)“二维”DNN 通过卷积和池化操作来分层提取特征的方式与人类视觉系统分级处理视觉信息的机理更相似,都是从边缘到局部到整体,最后进行分类判断,从低层到高层的特征表达也越来越抽象和概念化,“一维”DNN 不具备这样的特性.

因此,在基于深度学习特征的协同视觉显著性检测方法中,基于“二维”DNN 的方法普遍优于基于“一维”DNN 的方法.

综上所述,从特征提取策略的角度来看,基于“二维”深度学习特征的协同视觉显著性方法应该具有更好的性能.

## 4 实验对比与分析

本节在 5 个公开数据库上对协同视觉显著性检测的流行算法进行了主观和定量对比,并依据实验对比结果分析了特征提取策略对算法性能的影响以及各数据库的复杂度.

### 4.1 数据库

参与实验对比的 5 个公开数据库分别是:Image pair<sup>[9]</sup>、Cosal2015<sup>[59]</sup>、iCoseg<sup>[67]</sup>、MSRC-A<sup>[68]</sup>和 MSRC-B<sup>[68]</sup>,均以人工标注的二值图作为 Ground Truth,其中,

协同显著区域标记为 1,其他区域为 0.

Image pair 数据库是协同视觉显著性检测领域的第一个数据库,包括 210 幅测试图像,分为 105 对,目标包括人类、花朵、汽车、轮船和各种动物等,每对图像都有相似的目标和不同的背景.大多数图像的宽、高尺寸不超过 200 像素,图像间目标相同,没有太多颜色和形状变化,图像背景较为单一.

iCoseg 数据库目前应用最为广泛,原本是针对协同分割任务而设计,共包含 25 个场景,涵盖体育、动物、地标等常见场景.数据库包含 643 幅测试图像,分为 38 组,每组图像中都至少包含 5 幅测试图像,每组图像包含相同的目标,形状姿态虽不同,但变化较小,同组图像的背景也较为相似.

MSRC-A 库和 MSRC-B 库是 John 等<sup>[68]</sup>在目标分类问题中提出来的数据库. MSRC-A 库按照目标种类将 240 幅图像分为 7 组图像, MSRC-B 库将 591 幅图像分为 20 组.图像包括飞机、汽车、人脸等,每组图像间的目标具有一定颜色和形状变化.为了适应协同显著性检测的实验验证,Zhang 等<sup>[17,59]</sup>去除了 MSRC-A 中 7 幅不含目标的图像,Liu 等<sup>[20]</sup>去除了 MSRC-B 中 6 个不适合做协同视觉显著性检测的图像组.

Cosal2015 数据库是 Zhang 等<sup>[59]</sup>最近提出的数据库,测试图像从 ILSVRC2014<sup>[69]</sup>和 YouTube video set<sup>[70]</sup>上收集得到.数据库共包含 2015 幅测试图像,分为 50 组,每组都至少有 30 幅图像,同组图像公共目标的颜色和形状变化程度较大,单幅图像中,显著目标多样,背景复杂.从数据库的规模和内容来看,Cosal2015 数据库最具挑战性.

除上述 5 个数据库之外,还有 SIVAL 数据库<sup>[71]</sup>、RGBD Cosal2015 数据库<sup>[52]</sup>和 RGBD Coseg183 数据库<sup>[72]</sup>等. SIVAL 数据库最初用于图像检索,由于各图像组内的语义相似性很高,所以也可用于协同视觉显著性检测. RGBD Cosal2015 和 RGBD Coseg183 数据库的测试图像均是包含深度信息的 RGBD 图像,主要用于 RGBD 图像组的协同视觉显著性检测.

## 4.2 评价指标

### 4.2.1 检测精度

衡量协同视觉显著性检测算法检测精度的常用评价指标是 P-R (Precision-Recall) 曲线,和它的综合指标 AP (Average Precision) 值和  $F$ -measure 值,其定义如式 (1) 所示.

$$\begin{aligned} P &= \frac{M \cap G}{M} \\ R &= \frac{M \cap G}{G} \\ AP &= \int P(R) dR \\ F\text{-measure} &= \frac{(1 + \beta) PR}{\beta P + R} \end{aligned} \quad (1)$$

其中, $P$  (Precision) 代表准确率, $R$  (Recall) 代表召回率, $M$  代表协同显著图阈值分割后的二值图, $G$  代表人工标注的 Ground Truth. 取遍协同显著图中的所有灰度值作为阈值,可得到一系列  $P$  和  $R$  值,得到  $P$ - $R$  曲线, $P$ - $R$  曲线下的面积即为 AP,也称为平均准确率.  $F$ -measure 是  $P$  和  $R$  的加权平均, $\beta$  设为 0.3<sup>[15]</sup>. AP 和  $F$ -measure 都是  $P$  和  $R$  的综合指标,数值越大说明算法性能越好.

### 4.2.2 时间复杂度

时间复杂度是衡量算法实用性的重要指标,本文在同一平台下对各算法的时间复杂度进行了对比,以某数据库单幅图像的平均运行时间来代表时间复杂度.

## 4.3 实验对比算法

本文参与对比的算法共 17 种,分为两类,一类为协同视觉显著性检测算法,共 13 种,分别是: ASPM<sup>[17]</sup>、CBCS<sup>[10]</sup>、CBCS-S<sup>[10]</sup>、CSHS<sup>[20]</sup>、DARM<sup>[50]</sup>、ESMG<sup>[73]</sup>、IPTD<sup>[60]</sup>、IPCS<sup>[9]</sup>、HCN<sup>[74]</sup>、LDAW<sup>[59]</sup>、RFPR<sup>[43]</sup>、SACS<sup>[19]</sup>和 SACS-R<sup>[19]</sup>;另一类为视觉显著性检测算法,共 4 种,分别是: RC<sup>[75]</sup>、MILP<sup>[76]</sup>、HCAF<sup>[77]</sup>、LPS<sup>[78]</sup> (用于验证协同视觉显著性和普通视觉显著性检测的区别).

### 4.3.1 协同视觉显著性检测算法

一方面,13 种协同显著性检测算法按照有无代码可分为两类,一类为开放了代码的 7 种算法,包括: CBCS、CBCS-S (CBCS 的单幅图像显著目标检测算法)、ESMG、SACS、SACS-R (SACS 的优化算法)、HCN 和 IPCS. 其中,CBCS-S、CBCS、ESMG、SACS-R 和 SACS 这 5 种算法适用于所有数据库. HCN 和 IPCS 算法只适用于两幅图像的协同视觉显著性检测,因而只在 Image pair 数据库上进行了实验对比. 上述 7 种算法的参数设置均采用默认设定. 另一类为未开放代码的 6 种算法,只提供部分数据库的实验结果,无法在所有数据库上都进行实验对比,因此,在各数据库上进行实验对比的算法个数不同.

另一方面,13 种算法可按照特征提取策略的不同分为两类:一类为采用手工特征的算法,包括 CBCS (CH 特征)、CBCS-S (CH 特征)、CSHS (CH 特征)、DARM (CH 和 SIFT 特征)、ESMG (颜色特征)、IPCS (颜色和纹理特征)、HCN (颜色特征)、RFPR (CH 特征)、SACS (CH 特征) 和 SACS-R (CH 特征);另一类为采用深度学习特征的算法,包括 ASPM (CNN 特征)、IPTD (SDAE 特征) 和 LDAW (CNN 特征).

### 4.3.2 视觉显著性检测算法

为了验证协同视觉显著性与普通视觉显著性检测的区别,添加 RC、MILP、HCAF 和 LPS 这 4 种显著性检测算法进行实验对比,4 种算法均有公开的代码. 与协同视觉显著性检测算法的分类方式一样,4 种算法也分为两类:一类为采用手工特征的算法,包括 RC (颜色特

征)和 MILP(颜色特征),另一类为采用深度学习特征的算法,包括 HCAF(CNN 特征)和 LPS(CNN 特征),均是 2018 年的最新工作。

考虑到 iCoseg 数据库在协同视觉显著性检测领域应用最为广泛且可对比算法最多(13 种算法中,除了只适用于两幅图像协同视觉显著性检测的 HCN 和 IPCS 算法,其余 11 种算法均可参与对比),本文选择在 iCoseg 数据库上进行协同视觉显著性检测算法和显著性检测算法的实验对比。

#### 4.4 检测精度对比

本节主观对比和定量对比图中算法顺序按照各算法 AP 指标的升序排列。

##### 4.4.1 在 Image pair 数据库上的对比

在 Image pair 数据库上,可以进行 8 种算法的对比,包括:CBCS-S<sup>[10]</sup>,CBCS<sup>[10]</sup>,ESMG<sup>[73]</sup>,HCN<sup>[74]</sup>,IPCS<sup>[9]</sup>,IPTD<sup>[60]</sup>,SACS-R<sup>[19]</sup>和 SACS<sup>[19]</sup>。

##### (1) 主观对比

本文选取 Image pair 数据库中的一个图像对进行主观对比,各算法的检测结果和 Ground Truth(GT)的对比如图 2 所示。由于篇幅有限,没有展示 CBCS-S 和 SACS 两个算法的结果。

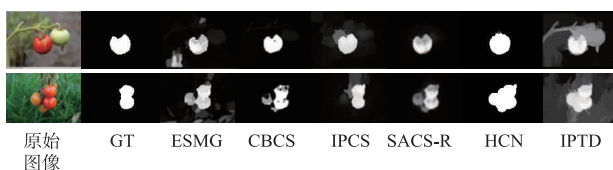


图2 Image pair数据库上6种算法的主观对比

##### (2) 定量对比

Image pair 数据库的定量对比如图 3 所示。综合主观对比(图 2)和定量对比(图 3),HCN 和 IPTD 的综合性能较好,HCN 的 Precision 值最高,IPTD 算法的 P-R 曲线和 AP 指标最高。

##### 4.4.2 MSRC-A 和 MSRC-B 数据库上的对比

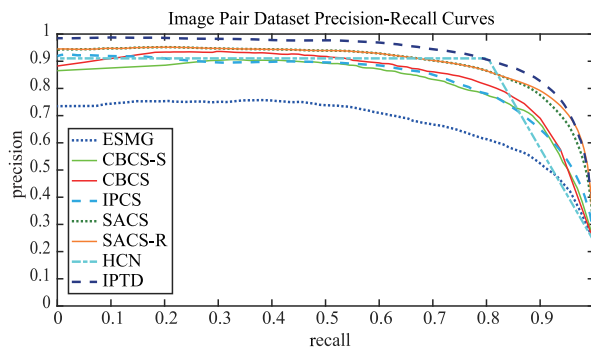
##### (1) 主观对比

MSRC-A 库与 MSRC-B 存在少量相同的图像组,为了使更多算法进行统一的主观对比,这里选取 MSRC-A 和 MSRC-B 共有图像组中的 6 幅图像进行测试,使所有开放代码或是在两个数据库提供测试结果的算法都能进行统一的主观对比。如图 4 所示。

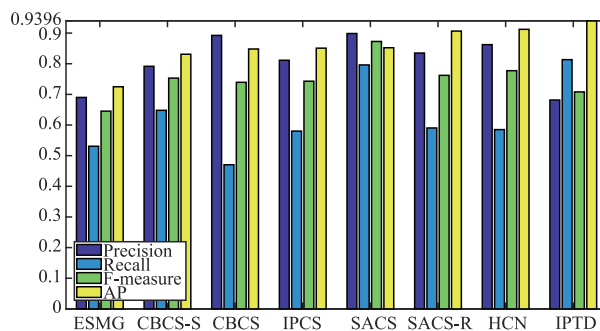
##### (2) MSRC-A 数据库上的定量对比

在 MSRC-A 数据库,可以进行 7 种算法的定量对比,包括:ASPM<sup>[17]</sup>,CBCS-S<sup>[10]</sup>,CBCS<sup>[10]</sup>,ESMG<sup>[73]</sup>,LDAW<sup>[59]</sup>,SACS-R<sup>[19]</sup>和 SACS<sup>[19]</sup>。如图 5 所示。

综合图 4 和图 5,在 MSRC-A 库上,ASPM 的 P-R 曲线、F-measure 和 AP 这 3 项指标均是最优,LDAW 的表现也较为优异。



(a) Image pair数据库上8种算法的P-R曲线对比



(b) Image pair数据库上8种算法的P, R, F-measure和AP值对比

图3 Image pair数据库上8种算法的定量对比

##### (3) MSRC-B 数据库上的定量对比

在 MSRC-B 数据库上,可以进行 8 种算法的定量对比,包括:CBCS-S<sup>[10]</sup>,CBCS<sup>[10]</sup>,CSHS<sup>[20]</sup>,DARM<sup>[50]</sup>,ESMG<sup>[73]</sup>,RFPR<sup>[43]</sup>,SACS-R<sup>[19]</sup>和 SACS<sup>[19]</sup>。如图 6 所示。

由图 4 和图 6 可知,DARM 在 P-R 曲线、F-measure 和 AP 这 3 项指标上的表现均是最优,SACS 和 SACS-R 算法除 Recall 值相对较低外,其他指标的表现也较为优异。

##### 4.4.3 在 Cosal2015 数据库上的对比

在 Cosal2015 数据库上,可以进行 6 种算法的定量对比,包括:CBCS-S<sup>[10]</sup>,CBCS<sup>[10]</sup>,ESMG<sup>[73]</sup>,SACS-R<sup>[19]</sup>,SACS<sup>[19]</sup>和 LDAW<sup>[59]</sup>。

##### (1) 主观对比

Cosal2015 数据库上的主观对比如图 7 所示。

##### (2) 定量对比

Cosal2015 数据库上的定量对比如图 8 所示。

综合图 7 和图 8,LDAW 在 P-R 曲线、F-measure 和 AP 这 3 项指标上的表现均是最优,SACS/SACS-R 算法的表现也较为优异。

##### 4.4.4 在 iCoseg 数据库上的对比

iCoseg 是协同视觉显著性检测领域最常用的数据库,在该库上共有 15 种算法参与对比,包括 11 种协同视觉显著性检测算法:ASPM<sup>[17]</sup>,CBCS-S<sup>[10]</sup>,CBCS<sup>[10]</sup>,CSHS<sup>[20]</sup>,DARM<sup>[50]</sup>,ESMG<sup>[73]</sup>,IPTD<sup>[60]</sup>,LDAW<sup>[59]</sup>,RFPR<sup>[43]</sup>,SACS-R<sup>[19]</sup>,SACS<sup>[19]</sup>,和 4 种显著性检测算法:

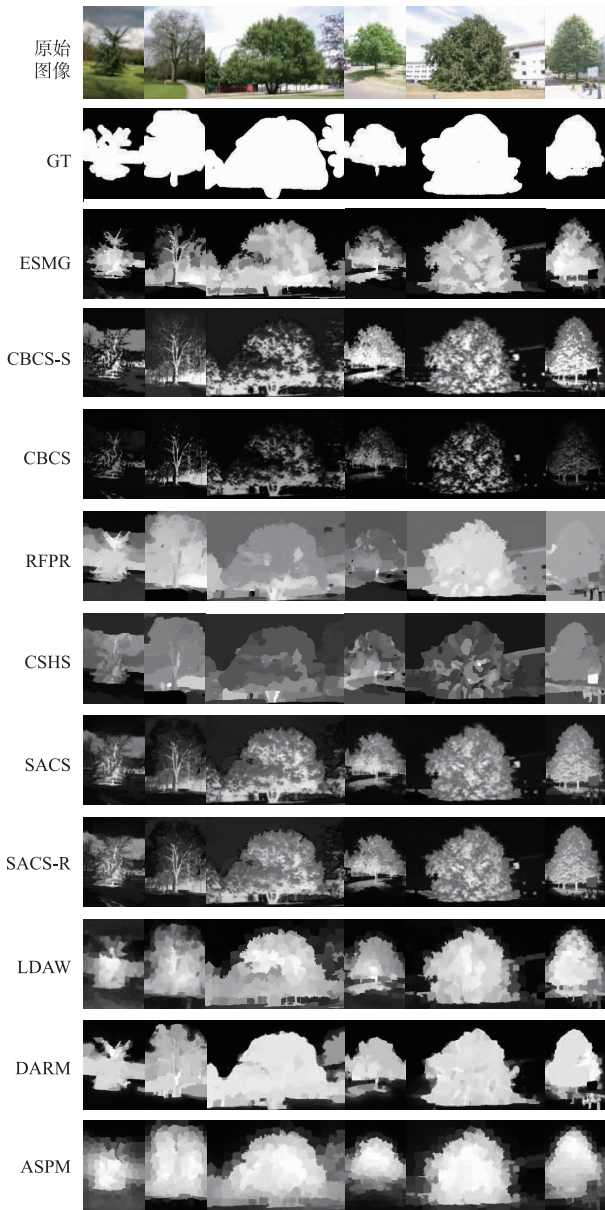


图4 MSRC数据库上10种算法的主观对比

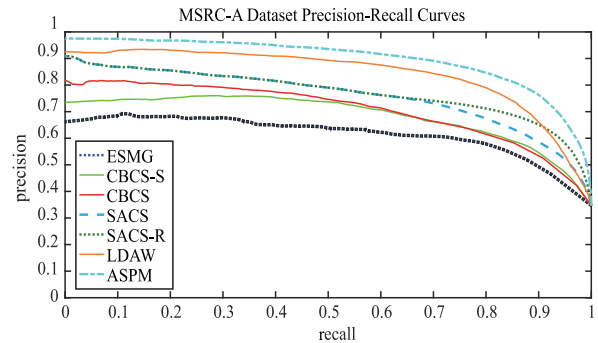
RC<sup>[75]</sup>, MILP<sup>[76]</sup>, HCAF<sup>[77]</sup>, LPS<sup>[78]</sup>.

#### (1) 主观对比

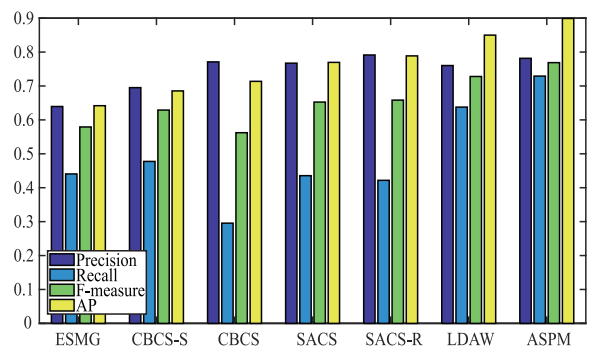
iCoseg 数据库上 15 种算法的主观对比如图 9 所示. 图 9 直观的展示出协同视觉显著性检测和普通视觉显著性检测的区别, 协同视觉显著性检测的结果只凸显 5 幅图像中的公共显著目标, 即: 红衣球员, 显著性检测的结果凸显被检测图像中所有显著目标, 除了红衣球员, 第二幅图像中的蓝衣球员、第四幅图像中的荧光绿衣球员, 第五幅图像中的浅褐色守门员等非公共显著目标也被凸显.

#### (2) 定量对比

iCoseg 数据库上 15 种算法的定量对比如图 10 所示. 图 10(b)、图 10(c) (考虑到曲线太多不利于观察,



(a) MSRC-A数据库上7种算法的P-R曲线对比



(b) MSRC-A数据库上7种算法的P, R, F-measure和AP值对比

图5 MSRC-A数据库上7种算法的定量对比

图 10(b) 只选择了 6 种协同视觉显著性检测算法) 展示了协同视觉显著性检测和显著性检测算法的定量对比结果, 总体来看, 4 种显著性检测算法的排名均位于后 50% 内, 说明显著性检测算法不能直接用于协同视觉显著性检测.

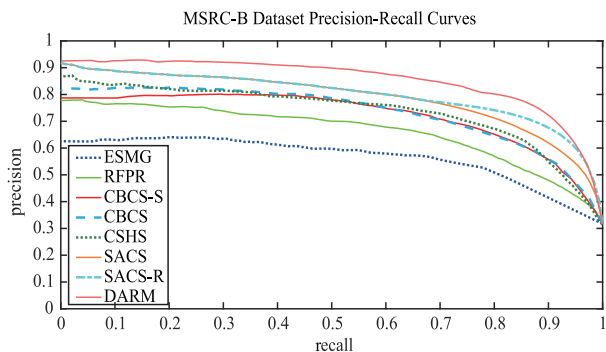
综合图 9 和图 10(a)、图 10(c) 的对比结果, 在 11 种协同视觉显著性检测算法中, DARM 的 P-R 曲线、F-measure 和 AP 这 3 项指标均是最优, ASPM、IPTD、LDAW 和 RFPR 的表现也较为优异, 其中 ASPM 算法的 Recall 指标最高.

#### 4.5 时间复杂度对比

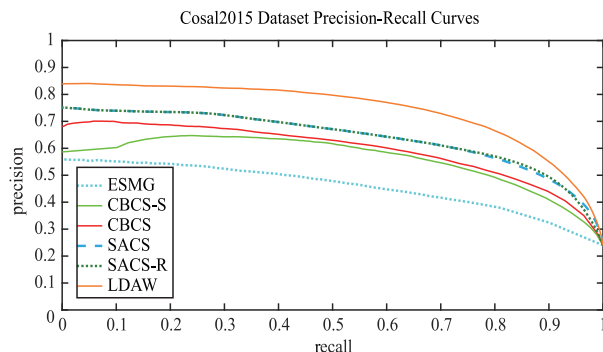
为方便时间复杂度和算法检测精度的综合对比和分析, 本文继续选择在 iCoseg 数据库上进行 5 种协同视觉显著性检测算法 (只有 5 种算法提供了可在所有数据库上运行的代码) 和 4 种显著性检测算法的时间复杂度对比.

本文所有实验的运行平台为: CPU i5-7400 (3.0G 主频), 内存 16G (DDR4 2400), 显卡 GTX1070 (8G 显存).

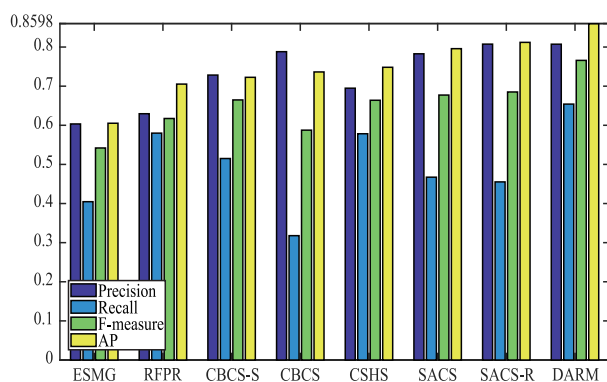
如表 1 所示, 9 种算法中, 运行速度较快的为 RC 和 LPS 算法, 其余依次 CBCS-S、CBCS、ESMG、SACS、HCAF、SACS-R 和 MILP. 然而, RC 和 LPS 算法都是普通视觉显著性检测算法, 在协同视觉显著性数据库上的



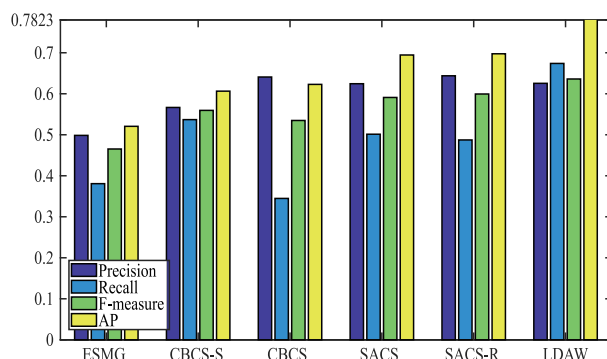
(a) MSRC-B数据库上 8 种算法的  $P-R$



(a) Cosal2015 数据库上 6 种算法的  $P-R$  曲线对比



(b) MSRC-B数据库上 8 种算法的  $P, R, F$ -measure 和 AP 值对比



(b) Cosal2015 数据库上 6 种算法的  $P, R, F$ -measure 和 AP 值对比

图6 MSRC-B数据库上8种算法的定量对比

图8 Cosal2015数据库上6种算法的定量对比

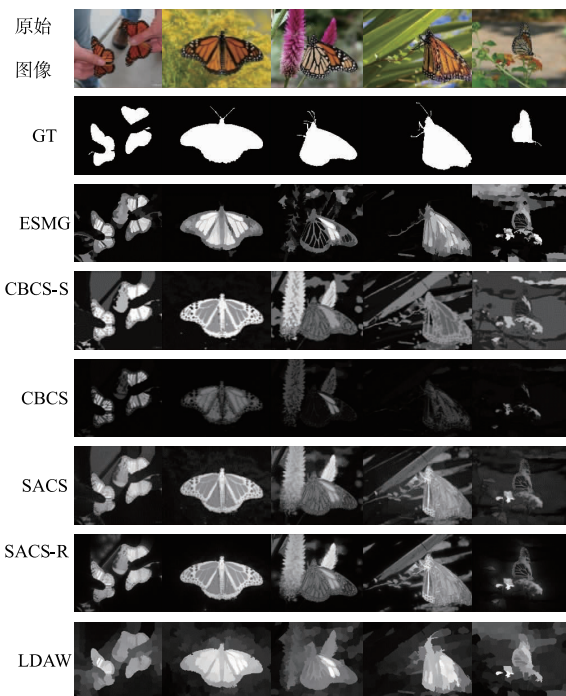


图7 Cosal2015数据库上6种算法的主观对比

比的 9 种算法中, SACS 的综合性最佳, 它在检测精度和时间复杂度两方面的排名都是中上.

表 1 在 iCoseg 数据库上时间复杂度对比

协同视觉显著性检测算法	CBCS-S	CBCS	ESMG	SACS	SACS-R
时间/秒	0.351	1.688	1.723	2.652	8.873
代码类型	Matlab	Matlab	Matlab	Matlab	Matlab
视觉显著性检测算法	RC	LPS	HCAF	MILP	
时间/s	0.023	0.112	7.893	77.691	
代码类型	C++	Python	Matlab	Matlab	

#### 4.6 特征提取策略对算法性能影响的分析

参与实验对比的算法中, 基于手工特征的协同视觉显著性检测方法包括: CBCS、CSHS、ESMG、DARM、SACS、RFPR、IPCS 和 HCN. 分析实验对比结果可得到如下结论:

(1) CBCS 和 ESMG 这类直接采用颜色或纹理等传统特征的方法整体性能一般;

(2) CSHS、SACS 和 RFPR 在使用手工特征的基础上加入了低秩矩阵恢复的算法, 检测性能相比 CBCS 和

检测精度较低, 综合表 1 和图 10, 在参与时间复杂度对

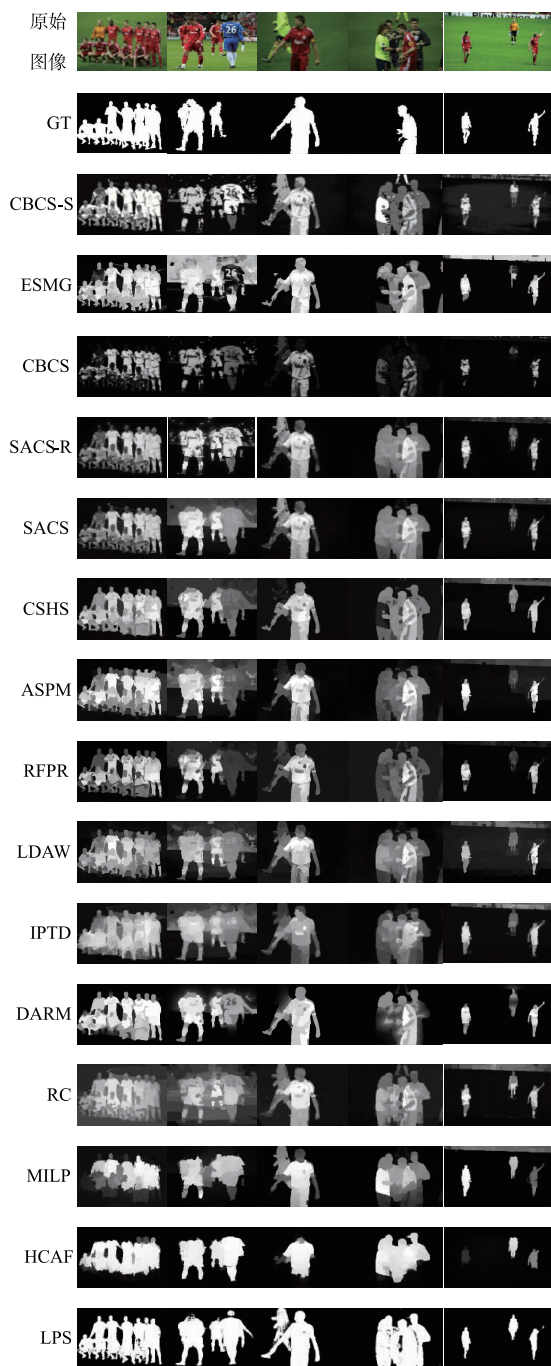


图9 iCoseg数据库上15种算法的主观对比

ESMG 算法有所提升,其中,CSHS 算法对显著非相似性区域的误检较少,而 RFPR 算法的背景干扰较为严重,SACS 算法对显著非相似性区域误检较多;

(3)DARM 算法虽然采用的是手工特征(SIFT 特征),但是凭借优秀的整体算法设计,在所有对比算法中性能优异,与基于深度学习特征的算法性能不相上下。

基于深度学习特征的算法包括:LDAW、IPTD 和 ASPM. 分析实验对比结果可得到如下结论:

(1)基于深度学习特征的协同视觉显著性检测算法的性能整体优于基于手工特征的算法;

(2)从定量对比来看,3 种算法的性能较为接近,但是主观对比结果反映,IPTD 算法存在背景干扰的问题。

综上所述,特征提取策略对协同视觉显著性检测的性能有明显的影 响,基于深度学习特征的算法整体表现更优。

#### 4.7 数据库的复杂度分析

CBCS、CBCS-S、ESMG、SACS 和 SACS-R 这 5 种算法在 5 个数据库上都参与了实验对比,通过它们在 5 个数据库上的定量指标对比,可以对 5 个数据库的复杂度进行直观的分析.表 2 和表 3 中分别给出了上述 5 种算法在 5 个数据库上的  $F$ -measure 和 AP 值对比。

表 2 5 个数据库上的  $F$ -measure 对比

算法	Image pair	iCoseg	MSRC-A	MSRC-B	Cosal2015
CBCS	0.7393	0.6860	0.5622	0.2875	0.5349
CBCS-S	0.7531	0.6861	0.6854	0.6648	0.5595
ESMG	0.6453	0.7208	0.5791	0.5420	0.4653
SACS	0.7621	0.7146	0.6525	0.6773	0.5910
SACS-R	0.7773	0.7250	0.6583	0.6851	0.5994
均值	0.7354	0.7065	0.6275	0.6313	0.5500

表 3 5 个数据库上的 AP 对比

算法	Image pair	iCoseg	MSRC-A	MSRC-B	Cosal2015
CBCS	0.8480	0.8172	0.7137	0.7363	0.6229
CBCS-S	0.8310	0.7776	0.6290	0.7226	0.6064
ESMG	0.7248	0.7948	0.6418	0.6052	0.5206
SACS	0.9065	0.8479	0.7696	0.7959	0.6945
SACS-R	0.9123	0.8481	0.7887	0.8120	0.6975
均值	0.84452	0.8171	0.7086	0.7344	0.6284

如表 2 和表 3 所示,5 个数据库的  $F$ -measure 和 AP 均值从高到低的排序均为 Image pair、iCoseg、MSRC-B、MSRC-A 和 Cosal 2015. 这说明 Cosal2015 数据库的复杂度最高,对算法的挑战性最强,Image pair 则相反.事实上,Cosal2015 数据库的测试图像数量和图像种类在 5 个数据库中都是最高;单幅测试图像中大多包含多个前景目标,背景也较为复杂;同组图像中,公共目标的颜色和形状变化程度较大。

综上所述,Cosal2015 数据库的复杂度最高,测试图像更符合实际场景,在该数据库上的检测结果更具说服力。

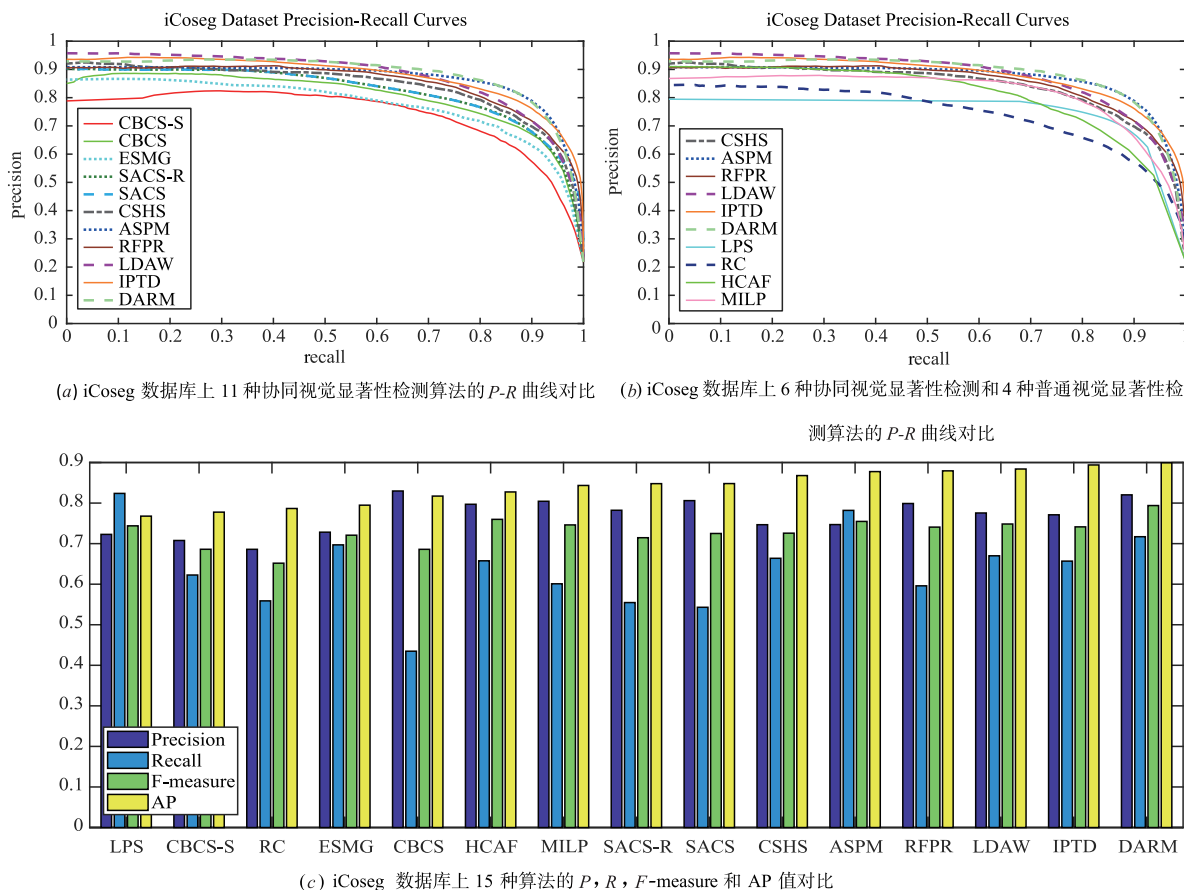


图10 iCoseg数据库上15种算法的定量对比

## 5 总结与展望

本文从特征提取策略的角度出发,对现有的协同视觉显著性检测方法进行分类总结和定性分析,并在 5 个数据库上对流行算法进行了实验对比,由实验结果可得到以下两个结论:①基于深度学习特征的协同视觉显著性检测方法性能较好;②5 个数据库中 Co-sal2015 数据库的复杂度最高,对算法最具挑战性。

协同视觉显著性检测目前还存在一些问题需要在后续研究中解决。①测试数据库的规模仍稍显不足:一方面,限制了对算法的评估能力;另一方面,人工标注样本数量的不足会导致训练中的过拟合等问题,尤其是深度学习特征的训练。②基于协同视觉显著性的应用相对较少。

针对测试数据库规模不足的问题,一方面,要加强无监督或弱监督的协同视觉显著性检测算法的研究,减少对人工标注样本的依赖;另一方面,可以尝试设计一种从现有的显著性数据集中生成训练数据的算法。本文给出一种可能的设计思路,仅供参考:利用视觉显著性检测数据集中的 Ground Truth 准确提取测试图像中的显著目标并对其进行特征提取,对所有显著目标

的特征向量进行聚类分析,将特征向量属于同一类的测试图像归为一组,则该组的测试图像具有同类的显著目标和不同的背景,符合协同视觉显著性检测的要求。

针对协同视觉显著性检测应用较少的问题,可以拓展协同视觉显著性在视频上的应用<sup>[79]</sup>,如,视频监控、运动目标检测及跟踪等。

### 参考文献

- [1] QIAN Xiao-liang, ZHANG Huan-long, LIU Yu-cui, et al. Visual Significance Detection Method Based on Self-Learning Characteristics and Matrix Low-Rank Recovery: China ZL 201610926623.0[P]. 2017.
- [2] WANG Wen-guan, SHEN Jian-bing, LI Xue-long, et al. Robust video object cosegmentation[J]. IEEE Transactions on Image Processing, 2015, 24(10): 3137-3148.
- [3] ZHU Hong-yuan, MENG Fan-man, CAI Jian-fei, et al. Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation[J]. Journal of Visual Communication & Image Representation, 2016, 34(2): 12-27.
- [4] CHO Min-su, SHIN Young-min, LEE Kyoung-mu. Co-rec-

- ognition of image pairs by data-driven Monte Carlo image exploration [A]. Proceedings of the 10th European Conference on Computer Vision [C]. Berlin; Springer, 2008. 144 – 157.
- [5] GOFERMAN S, TAL A, ZELNIK M L. Puzzle-like collage [J]. Computer Graphics Forum, 2010, 29(2): 459 – 468.
- [6] 柯圣财, 赵永威, 李弼程, 等. 基于卷积神经网络和监督核哈希的图像检索方法 [J]. 电子学报, 2017, 45(1): 157 – 163.
- KE Sheng-cai, ZHAO Yong-wei, LI Bi-cheng, et al. Image retrieval based on convolutional neural network and kernel-based supervised hashing [J]. Acta Electronica Sinica, 2017, 45(1): 157 – 163. (in Chinese)
- [7] JACOBS D E, DAN B G, SHECHTMAN Eli. Cosaliency: where people look when comparing images [A]. Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology [C]. New York: ACM, 2010. 219 – 228.
- [8] CHEN H T. Preattentive co-saliency detection [A]. Proceedings of 2010 IEEE International Conference on Image Processing [C]. Piscataway, NJ: IEEE, 2010. 1117 – 1120.
- [9] LI Hong-liang, NGAN K N. A co-saliency model of image pairs [J]. IEEE Transactions on Image Processing, 2011, 20(12): 3365 – 3375.
- [10] FU Hua-zhu, CAO Xiao-chun, TU Zhuo-wen. Cluster-based co-saliency detection [J]. IEEE Transactions on Image Processing, 2013, 22(10): 3766 – 3778.
- [11] LI Hong-liang, MENG Fan-man, NGAN K N. Co-salient object detection from multiple images [J]. IEEE Transactions on Multimedia, 2013, 15(8): 1896 – 1909.
- [12] JERRIPOTHULA K R, CAI Jian-fei, YUAN Jun-song. CATS: co-saliency activated tracklet selection for video co-localization [A]. Proceedings of 18th European Conference on Computer Vision [C]. Berlin; Springer, 2016. 187 – 202.
- [13] YAO Xi-wen, HAN Jun-wei, ZHANG Ding-wen, et al. Revisiting co-saliency detection: a novel approach based on two-stage multi-view spectral rotation co-clustering [J]. IEEE Transactions on Image Processing, 2017, 26(7): 3196 – 3209.
- [14] GE Chen-jie, FU Ke-ren, LIU Fang-hui, et al. co-saliency detection via inter and intra saliency propagation [J]. Signal Processing-image Communication, 2016, 44: 69 – 83.
- [15] ZHANG Ding-wen, FU Hua-zhu, HAN Jun-wei, et al. A Review of co-saliency detection technique: fundamentals, applications, and challenges [A]. LNCS: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition [C]. New York: IEEE, 2016. arXiv: 1604.07090.
- [16] ZHANG Ding-wen, MENG Deyu, Han Junwei. co-saliency detection via a self-paced multiple-instance learning framework [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(5): 865 – 878.
- [17] ZHANG Ding-wen, MENG De-yu, LI Chao, et al. A self-paced multiple-instance learning framework for co-saliency detection [A]. Proceedings of the 2015 IEEE International Conference on Computer Vision [C]. Washington: IEEE, 2015. 594 – 602.
- [18] CAO Xiao-chun, CHENG Yu-peng, TAO Zhi-qiang, et al. co-saliency detection via base reconstruction [A]. Proceedings of the 22nd ACM International Conference on Multimedia [C]. New York: ACM, 2014. 997 – 1000.
- [19] CAO Xiao-chun, TAO Zhi-qiang, ZHANG Bao, et al. Self-adaptively weighted co-saliency detection via rank constraint [J]. IEEE Transactions on Image Processing, 2014, 23(9): 4175 – 4186.
- [20] LIU Zhi, ZOU Wen-bin, LI Li-na, et al. Co-saliency detection based on hierarchical segmentation [J]. IEEE Signal Processing Letters, 2014, 21(1): 88 – 92.
- [21] 孙晓, 潘汀. 基于兴趣区域深度神经网络的静态面部表情识别 [J]. 电子学报, 2017, 45(5): 1189 – 1197.
- SUN Xiao, PAN Ting. Static facial expression recognition system using ROI deep neural network [J]. Acta Electronica Sinica, 2017, 45(5): 1189 – 1197. (in Chinese)
- [22] JUDD Tilke, EHINGER K A, DURAND Fredo, et al. Learning to predict where humans look [A]. Proceedings of 2009 IEEE 12th International Conference on Computer Vision [C]. Piscataway, NJ: IEEE, 2009. 2106 – 2113.
- [23] LI Yin, ZHOU Yue, YAN Jun-chi, et al. Visual saliency based on conditional entropy [A]. Proceedings of the 9th Asian Conference on Computer Vision [C]. Berlin: ACM, 2009. 246 – 257.
- [24] YAN Jun-chi, LIU Jian, LI Yin, et al. Visual saliency detection via rank-sparsity decomposition [A]. Proceedings of 2010 IEEE International Conference on Image Processing [C]. Piscataway, NJ: IEEE, 2010. 1089 – 1092.
- [25] WANG Wei, CHEN Cheng, WANG Yi-zhou, et al. Simulating human saccadic scanpaths on natural images [A]. Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition [C]. DC: IEEE, 2011. 441 – 448.
- [26] KUMMERER M, THEIS L, BETHGE M. Deepgaze I: boosting saliency prediction with feature maps trained on imagenet [A]. LNCS: Proceedings of Poster Presented at International Conference on Learning Representations [C]. Berlin; Springer, 2014. arXiv: 1411.1045.
- [27] KUMMERER M, WALLIS T S, BETHGE M. Deep gaze II: reading fixations from deep features trained on object

- recognition[A]. LNCS:Proceedings of the 2016 Computer Vision and Pattern Recognition [C]. Berlin; Springer, 2016. arXiv:1610.01563.
- [28] KRUTHIVENTI S S, GUDISA V, DHOLAKIYA J H, et al. Saliency unified; a deep architecture for simultaneous eye fixation prediction and salient object segmentation [A]. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition [C]. Piscataway, NJ: IEEE, 2016. 5781 – 5790.
- [29] CORNIA M, BARALDI L, SERRA G, et al. A deep multi-level network for saliency prediction[A]. Proceedings of 23rd International Conference on Pattern Recognition [C]. New York: IEEE, 2016. 3488 – 3493.
- [30] HOU Xiao-di, ZHANG Liqing. Saliency detection; a spectral residual approach [A]. Proceedings of the 2007 Computer Vision and Pattern Recognition [C]. Berlin; Springer, 2007. 1 – 8.
- [31] HOU Xiao-di, HAREL J, KOCH C. Imagesignature: highlighting sparse salient regions [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(1): 194 – 201.
- [32] AVRAHAM T, LINDENBAUM M. Esaliency (extended saliency): meaningful attention using stochastic image modeling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(4): 693 – 708.
- [33] HAREL J, KOCH C, PERONA P. Graph-based visual saliency [A]. Proceedings of the 19th International Conference on Neural Information Processing Systems [C]. Cambridge; MIT, 2006. 545 – 552.
- [34] 郑云飞, 张雄伟, 曹铁勇, 等. 基于全卷积网络的语义显著性区域检测方法研究 [J]. 电子学报, 2017, 45(11): 2593 – 2601.  
ZHENG Yun-fei, ZHANG Xiong-wei, CAO Tie-yong, et al. The semantic salient region detection algorithm based on the fully convolutional networks [J]. Acta Electronica Sinica, 2017, 45(11): 2593 – 2601. (in Chinese)
- [35] LIU Tie, YUAN Ze-jian, SUN Jian, et al. Learning to detect a salient object [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(2): 353 – 367.
- [36] CHENG Ming-ming, MITRA N J, HUANG Xiao-lei, et al. Global contrast based salient region detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 569 – 582.
- [37] LI Yin, ZHOU Yue, XU Lei, et al. Incremental sparse saliency detection [A]. Proceedings of the 16th IEEE International Conference on Image Processing [C]. Piscataway, NJ: IEEE, 2009. 3093 – 3096.
- [38] JIANG Huai-zu, WANG Jing-dong, YUAN Ze-jian, et al. Salient object detection; a discriminative regional feature integration approach [A]. Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition [C]. Washington: IEEE, 2013. 2083 – 2090.
- [39] SHEN Xiao-hui, WU Ying. A unified approach to salient object detection via low rank matrix recovery [A]. Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition [C]. Washington: IEEE, 2012. 853 – 860.
- [40] TONG Na, LU Hu-chuan, RUAN Xiang, et al. Salient object detection via bootstrap learning [A]. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition [C]. Washington: IEEE, 2015. 1884 – 1892.
- [41] HAN Jun-wei, ZHANG Ding-wen, HU Xin-tao, et al. Background prior-based salient object detection via deep reconstruction residual [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(8): 1309 – 1321.
- [42] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1998, 20(11): 1254 – 1259.
- [43] LI Li-na, LIU Zhi, ZOU Wen-bin, et al. Co-saliency detection based on region-level fusion and pixel-level refinement [A]. Proceedings of 2014 IEEE International Conference on Multimedia and Expo [C]. New York: IEEE, 2014. 1 – 6.
- [44] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection [A]. Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition [C]. Piscataway: IEEE, 2009. 1597 – 1604.
- [45] HOU Xiao-di, ZHANG Li-qing. Saliency detection; a spectral residual approach [A]. Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition [C]. New York: IEEE, 2007. 1 – 8.
- [46] SONG Hang-ke, LIU Zhi, XIE Yu-feng, et al. RGBD Co-saliency detection via bagging-based clustering [J]. IEEE Signal Processing Letters, 2016, 23(12): 1722 – 1726.
- [47] HUANG Rui, FENG Wei, SUN Ji-zhou. Saliency and co-saliency detection by low-rank multiscale fusion [A]. Proceedings of 2015 IEEE International Conference on Multimedia and Expo [C]. New York: IEEE, 2015. 1 – 6.
- [48] SONG Hang-ke, LIU Zhi, DU Huan, et al. Depth-aware saliency detection using discriminative saliency fusion [A]. Proceedings of 2016 IEEE International Conference on Acoustics, Speech and Signal Processing [C]. New York: IEEE, 2016. 1626 – 1630.
- [49] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,

- 2011, 33(5):898–916.
- [50] YE Lin-wei, LIU Zhi, LI Jun-hao, et al. Co-saliency detection via Co-salient object discovery and recovery [J]. *IEEE Signal Processing Letters*, 2015, 22(11):2073–2077.
- [51] DU Shu-ze, CHEN Shi-feng. Detecting Co-salient objects in large image sets [J]. *IEEE Signal Processing Letters*, 2014, 22(2):145–148.
- [52] CONG Run-min, LEI Jian-jun, FU Hua-zhu, et al. An iterative Co-saliency framework for RGBD images [J]. *IEEE Transactions on Cybernetics*, 2017, PP(99):1–14.
- [53] ZHANG Zhao-feng, WU Ze-min, JIANG Qing-zhu, et al. Co-saliency detection based on superpixel matching and cellular automata [J]. *Ksii Transactions on Internet & Information Systems*, 2017, 11(5):2576–2589.
- [54] SHEN Ning-min, LI Jing, ZHOU Pei-yun, et al. BSFCoS: block and sparse principal component analysis-based fast Co-saliency detection method [J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2016, 30(1):1655003.
- [55] 胡正平, 陈俊岭. 多层融合深度局部 PCA 子空间稀疏优化特征提取模型 [J]. *电子学报*, 2017, 45(10):2383–2389.
- HU Zheng-ping, CHEN Juan-ling. Feature extraction model based on multi-layered deep local subspace sparse optimization [J]. *Acta Electronica Sinica*, 2017, 45(10):2383–2389. (in Chinese)
- [56] CHEN Duan-yu, LIN Chuan-yu, YANG NienTzu, et al. Sparse coding-based co-salient object detection with application to video abstraction [A]. *Proceedings of 2013 International Conference on Machine Learning and Cybernetics* [C]. New York: IEEE, 2013. 1474–1479.
- [57] 钱晓亮, 郭雷, 韩军伟, 等. 一种基于加权稀疏编码的频域视觉显著性检测算法 [J]. *电子学报*, 2013, 41(6):1159–1165.
- QIAN Xiao-liang, GUO Lei, HAN Jun-wei, et al. Aspectral algorithm based on weighted sparse coding for visual saliency detection [J]. *Acta Electronica Sinica*, 2013, 41(6):1159–1165. (in Chinese)
- [58] HINTON G E, SALAKHUTDINOV R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786):504–507.
- [59] ZHANG Ding-wen, HAN Jun-wei, LI Chao, et al. Co-saliency detection via looking deep and wide [A]. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition* [C]. New York: IEEE, 2015. 2994–3002.
- [60] ZHANG Ding-wen, HAN Jun-wei, HAN Jun-gong, et al. Cosaliency detection based on intrasaliency prior transfer and deep intersaliency mining [J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2016, 27(6):1163–1176.
- [61] JEONG D J, HWANG I, CHO N I. Co-salient object detection based on deep saliency networks and seed propagation over an integrated graph [A]. *LNCS: Proceedings of 2017 Computer Vision and Pattern Recognition* [C]. New York: IEEE, 2017. arXiv:1706.09650.
- [62] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [A]. *LNCS: Proceedings of 2014 Computer Vision and Pattern Recognition* [C]. New York: IEEE, 2014. arXiv:1409.1556.
- [63] HAN Jun-wei, CHENG Gong, LI Zhen-peng, et al. A unified metric learning-based framework for co-saliency detection [J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2017, PP(99):1–1.
- [64] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. *Nature*, 2015, 521(7553):436–444.
- [65] BRAHMA P P, WU Da-peng, SHE Yi-yuan. Why deep learning works: a manifold disentanglement perspective [J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2016, 27(10):1997–2008.
- [66] OLSHUAUSEN B A, FIELD D J. Sparse coding with an overcomplete basis set: a strategy employed by V1? [J]. *Vision Research*, 1997, 37(23):3311–3325.
- [67] BATRA D, KOWDEL A, PARIKH D, et al. iCoseg: interactive co-segmentation with intelligent scribble guidance [A]. *Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* [C]. New York: IEEE, 2010. 3169–3176.
- [68] WINN J M, CRIMINISI A, MINKA T. Object categorization by learned universal visual dictionary [A]. *Proceedings of Tenth IEEE International Conference on Computer Vision* [C]. New York: IEEE, 2005. 1800–1807.
- [69] RUSSAKOVSKY O, DENG Jia, SU Hao, et al. Image net large scale visual recognition challenge [J]. *International Journal of Computer Vision*, 2015, 115(3):211–252.
- [70] PREST A, LEISTNER C, CIVERA J, et al. Learning object class detectors from weakly annotated video [A]. *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition* [C]. Washington: IEEE, 2012. 3282–3289.
- [71] MARON O, EREZ L P. A framework for multiple-instance learning [J]. *Advances in Neural Information Processing Systems*, 1998, 200(2):570–576.
- [72] FU Hua-zhu, XU Dong, LIN Stephen, et al. Object-based RGBD image co-segmentation with mutex constraint [A]. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition* [C]. New York: IEEE,

2015. 4428 – 4436.
- [73] LI Yi-jun, FU Ke-ren, LIU Zhi, et al. Efficient saliency-model-guided visual co-saliency detection [J]. IEEE Signal Processing Letters, 2015, 22(5): 588 – 592.
- [74] LOU Jing, XU Feng-lei, XIA Qing-yuan, et al. Hierarchical co-salient object detection via color names [A]. Proceedings of the Asian Conference on Pattern Recognition [C]. New York: IEEE, 2017. 718 – 724.
- [75] CHENG Ming-ming, MITRA N J, HUANG Xiao-lei, et al. Global contrast based salient region detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 569 – 582.
- [76] HUANG Fang, QI Jin-qing, LU Hu-chuan, et al. Salient object detection via multiple instance learning [J]. IEEE Transactions on Image Processing, 2017, 26(4): 1911 – 1922.
- [77] QIN Yao, FENG Meng-yang, LU Hu-chuan, et al. Hierarchical cellular automata for visual saliency [J]. International Journal of Computer Vision, 2018, 126(7): 751 – 770.
- [78] ZENG Yu, LU Hu-chuan, ZHANG Li-he, et al. Learning to promote saliency detectors [A]. Proceedings of 2018 IEEE Conference on Computer Vision and Pattern Recognition [C]. New York: IEEE, 2018. 1644 – 1653.
- [79] WANG Wen-guan, SHEN Jian-bing, SUN Han-qiu, et al. ViCoS2: video co-saliency guided co-segmentation [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2018, 28(8): 1727 – 1736.

### 作者简介



**钱晓亮** 男, 1982 年 10 月生于河南孟州。2013 年毕业于西北工业大学, 获工学博士学位。现为郑州轻工业学院电气信息工程学校聘副教授, 硕士生导师。主要研究方向为计算机视觉与人工智能, 视觉注意力计算, 遥感图像分析与处理。发表 SCI/EI 论文 20 余篇; 授权国家发明专利 9 项; 主持或参与国家自然科学基金、河南省高等学校重点科研项目、河南省科技攻关项目、河南省高校科技创新团队等项目 10 余项。

E-mail: qxl\_sunshine@163.com



**白臻** 女, 1992 年 11 月生于河南巩义。2016 年毕业于武汉华夏理工学院, 获学士学位。现为郑州轻工业学院电气信息工程学院硕士研究生。主要研究方向为计算机视觉与机器学习, 视觉显著性检测。

E-mail: bz532476@163.com



**毋媛媛 (通讯作者)** 女, 1982 年 8 月生于河南商丘。2010 年毕业于东南大学, 获工学博士学位。现为郑州轻工业学院电气信息工程学校聘副教授, 硕士生导师。主要研究方向为计算机视觉、非线性控制和复杂网络。发表学术论文 20 余篇; 主持国家自然科学基金、中国博士后基金和河南省科技攻关项目各 1 项。

E-mail: wuyanyuan82@163.com



**王慰 (通讯作者)** 男, 1969 年 11 月生于河北保定。2002 年毕业于加拿大康科迪亚大学, 获工学博士学位。现为国家千人计划特聘专家, 郑州轻工业学院电气信息工程学院教授。主要研究方向为模式识别与人工智能, 计算机视觉, 触觉感知与柔性电子。曾任美国纽约州立大学纳米学院的教授/博导, 美国国家纳米科研机构 SRC、NRI 和 Sematech 的主要领导成员, 国际工

程师总会 IEEE 的纳米电子协会主席, 中科院物联网中心的传感器研究室主任和北京印刷学院首席科学家。曾获 2004 年加拿大纳米新技术 CFI Award, 2005 年美国普渡大学科技创新奖, 2008 年、2010 年 IBM 教授奖。共发表研究 SCI 论文 130 余篇, 出版专著 2 部, 授权 30 多项专利和软件著作权。曾主持美国数个重大纳米电子和人工智能类脑芯片的科技专项, 北京市类脑专项课题, 参与十一五重大专项和中科院创新基金。 E-mail: 18900616029@189.cn