

# 立体视觉感知中的深度敏感信息提取综述

牟海维<sup>1,2</sup>, 王妙妙<sup>1,2</sup>, 韩建<sup>1,2</sup>, 曹志民<sup>1,2\*</sup>

<sup>1</sup>东北石油大学电子科学学院, 黑龙江 大庆 163318;

<sup>2</sup>东北石油大学黑龙江省高校共建测试计量技术及仪器仪表研发中心, 黑龙江 大庆 163318

**摘要** 图像显著性是实现视觉信息感知的最主要信息源,图像显著性信息的检测是计算机视觉研究中的热点。随着数据获取能力的极大提升,三维或立体视觉感知的需求越来越迫切,深度敏感信息的有效提取是影响当前立体视觉体验的重要元素,已经得到了很多研究人员的注意并取得了一定的研究成果。本文对图像显著性检测及深度敏感信息提取研究的相关检测模板及技术进行广泛研究,对深度敏感信息提取的研究现状进行了分类和概括,并通过对比实验,分析了几种典型深度敏感信息提取算法。最后,对深度敏感信息提取技术存在的问题及发展趋势进行了讨论。

**关键词** 图像处理; 深度敏感; 显著性; 立体视觉; 视觉注意力

中图分类号 O235; TP391

文献标志码 A

doi: 10.3788/LOP57.120003

## Survey of Depth Sensitive Information Extraction in Stereo Vision Perception

Mu Haiwei<sup>1,2</sup>, Wang Miaomiao<sup>1,2</sup>, Han Jian<sup>1,2</sup>, Cao Zhimin<sup>1,2\*</sup>

<sup>1</sup>School of Electronic Science, Northeast Petroleum University, Daqing, Heilongjiang 163318, China;

<sup>2</sup>Research and Development Center for Testing and Measurement Technology and Instrumentation, Heilongjiang University, Northeast Petroleum University, Daqing, Heilongjiang 163318, China

**Abstract** Image saliency is the most important information source to realize visual information perception. The detection of image saliency information has always been a hot topic in computer vision research. With the great improvement of data acquisition capability, the demand of three-dimensional or stereoscopic vision is becoming more and more urgent. The effective extraction of deep sensitive information is an important element that affects the current stereoscopic vision experience. Therefore, it has been paid attention by many researchers and obtained some research achievements. In this paper, through extensive research on the relevant detection templates and technologies of image saliency detection and depth sensitive information extraction research, the current status of depth sensitive information extraction research is classified and summarized, and several typical depth sensitive information extraction algorithms are analyzed through comparative experiments. Finally, the problems and development trends of depth-sensitive information extraction techniques are discussed.

**Key words** image processing; deep sensitivity; saliency; stereo vision; visual attention

**OCIS codes** 100.2960; 100.3008; 110.2960; 280.4750

## 1 引言

在未知环境中确定感兴趣的对象是计算机立体视觉中的重要任务,图像显著性检测就是在在一幅图像中获取人类视觉关注的部分,已经应用到图像处理与视觉相关的各个领域,如图像分类和图像分

割<sup>[1-2]</sup>、视频增强、物体检测和模式识别<sup>[3-4]</sup>、图像质量评估<sup>[5-8]</sup>等。显著性检测通常被分为自下而上或自上而下两种方法<sup>[9]</sup>,自下而上的方法是在没有先验知识指导下处理的数据驱动预注意机制,通常利用一些低级线索(例如特征,颜色和空间距离)来获取显著性特征<sup>[10]</sup>,自上而下的方法是目标驱动的,会受到

收稿日期: 2019-08-28; 修回日期: 2019-10-22; 录用日期: 2019-11-06

基金项目: 国家自然科学基金(51574087)、东北石油大学研究生创新科研项目(JYCX\_CX09\_2018)

\* E-mail: 1162814997@qq.com

具体任务和人的主观意识的影响,其研究结果通常呈现出较大的差异性。另一方面,显著性检测方法还可以分为两种类型:基于固定预测和基于突出物体预测<sup>[11-13]</sup>的显著性模型。固定预测的方法是预测人类眼睛观察的注视点,包括最早提出的 Itti 模型<sup>[14]</sup>,模仿人类视觉皮层神经元产生多个生物特征图,再由多尺度的图像特征因素组合成一个显著图。基于突出物体预测的方法会突出整个感兴趣的对象,在充分抑制背景区域的同时提取显著性前景。

在过去几年中,已经提出各个方面的二维(2D)显著性模型,但是在现有的显著性模型中没有很好地引入深度信息,在真实的三维世界中实际视觉显著性的潜在机制仍然很少被探索。目前,立体视觉感知已经成为三维信息提取的重要内容,三维(3D)数据的获取也越来越容易,成为计算机视觉研究中感兴趣的热点<sup>[15-16]</sup>。但是,对于3D显著性检测,如何同时考虑各种二维因素(例如颜色、方向和纹理等)和3D因素(例如深度感知、视觉舒适度等),并将深度信息结合到显著性模型中一直是目前研究领域中亟待解决的问题。总之,立体视觉中的深度敏感信息获取是二维图像显著性检测的一个扩展,目前提出的一些3D视觉注意力模型大多可以分为深度加权模型,深度显著模型和立体视觉模型<sup>[17-18]</sup>。深度加权模型就是将深度视为另一个维度,通过乘法或加法运算与2D显著图融合。深度显著模型将深度显著性视为附加信息<sup>[19-20]</sup>,首先从深度图提取诸如相对深度,弹出效果和深度梯度的深度特征以创建附加特征图,然后将其用于生成深度显著图,最后,将这些深度显著图与2D显著图组合以获得最终结果。

## 2 显著性物体检测及深度信息提取发展趋势

第一阶段:1998年,Itti等<sup>[14]</sup>提出了一个视觉注意系统,由多尺度图像特征因素组合成一个显著图,引起了计算机视觉等学科的关注。

第二阶段:第二波热潮是由Harel等<sup>[21]</sup>掀起。2007年,Harel提出一种新的自下而上的视觉显著性模型:基于图形的视觉显著性。2008年,Liu等<sup>[22]</sup>提出图像区域检测和分割的显著性检测算法,将显著性检测定义为二元分割问题,该模型简单合理,自此涌现了大量的图形显著性检测模型。

第三阶段:近几年显著性检测开始引入卷积神经网络的算法,与传统的基于对比度和各种先验的经典方法不同,基于卷积神经网络的方法可以智能

的从大量的图像中提取得到高级特征和多尺度特征,有效地消除了人工建立图像特征的弊端。因此,基于传统方法与卷积神经网络相结合的显著性检测逐渐成为主流方向。

第四阶段:目前,三维深度信息已经得到了很多研究人员的注意。Wang等<sup>[23]</sup>通过扩展传统的二维显著性检测方法,提出了三维图像视觉注意的计算模型。Fang等<sup>[24]</sup>分析了同一场景的2D和3D人体固定之间的差异,并通过扩展现有的2D显著性模型提出了一种3D显著性检测方法。在今后的显著性发展中,深度感知将在很大程度上影响人类的观看行为,在3D图像的显著性计算中已经充分考虑深度因子。

图1列出了显著性物体检测及深度信息提取发展趋势。

## 3 显著性模型介绍

经过十多年的发展,图像显著性检测取得了长足的进步,算法层出不穷,性能稳步提升。本节详细介绍了近几年一些经典显著性模型算法。

### 3.1 基于变换域分析的显著性模型

#### 3.1.1 基于傅里叶变换的算法

光谱残留(SR)算法<sup>[25]</sup>的显著性检测模型通过对输入图像的对数谱进行分析,提取出图像在谱域内的谱残差,并将光谱残差变换为图像空间域<sup>[26]</sup>,从而构造相应的显著图。2013年,Imamoglu等<sup>[27]</sup>提出了一种基于2Dmel-cepstrum的频谱残余显著性检测模型(MCSR),产生全分辨率显著图,无需下采样。

MCSR算法通过傅里叶变换(FT)获得每个颜色通道的幅度和相位谱,

$$A^c(f) = R\{\mathcal{F}[I^c(x)]\}, \quad (1)$$

$$P^c(f) = I\{\mathcal{F}[I^c(x)]\}, \quad (2)$$

式中: $c = \{l, a, b\}$ ,是CIE Lab颜色空间通道,然后,通过寻找振幅谱的不规则性,可以获得所提出模型的残差。

$$R^c(f) = \log_2\{\text{abs}[A^c(f) - G^c(f)]\}, \quad (3)$$

$$G_{m,n}^c(k,l) = \sum_{k,l \in A_{m,n}^c} \omega(m,n)A_{m,n}^c(k,l). \quad (4)$$

利用2Dmel-cepstrum的计算方法在幅度谱上应用网格, $G_{m,n}^c(k,l)$ 是基于网格的频率分量 $(k,l)$ 在一个网格单元 $(m,n)$ 处的平均幅度谱,对每个网格单元 $(m,n)$ 处应用非均匀加权 $\omega(m,n)$ 。

然后,利用(3)式的残差和傅里叶逆变换(IFT)生成显著图。

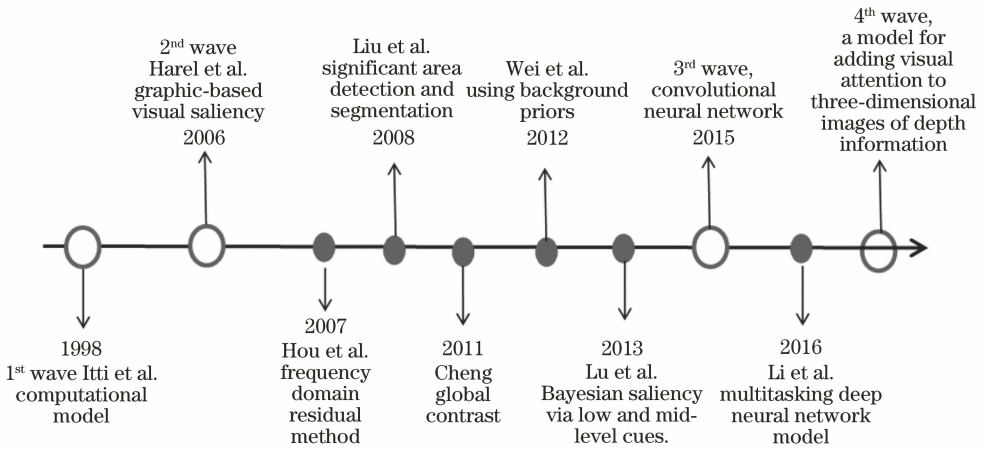


图 1 显著性物体检测及深度信息提取发展趋势

Fig. 1 Development trend of significant object detection and depth information extraction

$$S(x) = g(x) * \left\{ m(x) * \sum_c \mathcal{F}^{-1} \{ \exp[R^c(f) + p^c(f)] \} \right\}^2, \quad (5)$$

式中： $m(x)$ 和  $g(x)$ 分别是中值和高斯平滑滤波器。

以光谱残余显著性方法和 2Dmel-cepstrum 计算为基础的算法可以生成具有均匀区域的显著图，但是频域分析得到的结果中会丢失许多图像信息，导致显著图不能被明亮清晰地表现出来。

### 3.1.2 基于小波变换的算法

在文献[25]和[27]中，傅里叶变换是最终显著图的关键，因此，场景中的全局不规则性比局部不规则性更占优势。2012年，研究人员发现小波变换(WT)算法在视觉注意建模方面的应用，WT算法的优势在于它能够同时提供多尺度的环境空间信息和频率分析。Imamoglu等<sup>[27]</sup>消除了图像的一些高频成分，提出一种基于小波分解高通系数的显著性检测模型。

WT模型<sup>[28]</sup>通过整数小波变换(IWT)在多级小波分解上创建特征映射，建立多个频域带宽不断增加的特征图。

$$f_s^c(x, y) = \frac{[IWT_s(H_s^c V_s^c D_s^c)]^2}{\eta}, \quad (6)$$

式中： $f_s^c(x, y)$ 每个图像子带  $c$  的第  $s$  级分解生成的特征映射； $\eta$ 是缩放因子； $H$ 、 $V$ 和  $D$ 分别是给定  $c$ 和  $s$ 的水平、垂直和对角线细节的小波系数。特征图中给定位置的特征用具有正态分布的概率密度函数来表示，多维空间的高斯概率密度函数为

$$p[f(x, y)] = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \times \exp\{-1/2[f(x, y) - \mu]^T \Sigma^{-1} [f(x, y) - \mu]\}, \quad (7)$$

$$\Sigma = E\{[f(x, y) - \mu][f(x, y) - \mu]^T\}, \quad (8)$$

式中： $\mu$ 是每个特征图的平均值，即  $\mu = E[f]$ ； $T$ 是转置操作； $\Sigma$ 是  $n \times n$ 的协方差矩阵。用  $k \times k$  2D高斯低通滤波器进行滤波，以获得平滑映射，根据(7)式中的概率密度函数可以得到全局显著性图，

$$S(x, y) = \{\log_2 \{p[f(x, y)]^{-1}\}\}^{1/2} * \mathbf{I}_{k \times k}. \quad (9)$$

## 3.2 基于图形的显著性模型

### 3.2.1 基于流行排序的算法

Yang等<sup>[29]</sup>将图像表示为具有超像素作为节点的闭环图，提出了基于图的流形排序的显著性检测(MR)算法，在MR算法中，使用的是非正规拉普拉斯矩阵。给定数据集  $X = \{x_1, \dots, x_l, x_{l+1}, \dots, x_n\}$ ，一些数据点被标记为查询，其余数据点需要根据它们与查询的相关性进行排序， $f$ 表示为每个点  $x_i$ 分配排名值  $y_i$ 的排名函数，令  $\mathbf{y} = \{y_1, y_2, \dots, y_n\}^T$ 表示指示向量，其中，如果  $x_i$ 是查询，则  $y_i = 1$ ，否则  $y_i = 0$ ，得到一个排序函数如下得到一个排序函数如下

$$f^* = (\mathbf{D} - \alpha \mathbf{W})^{-1} \mathbf{y}, \quad (10)$$

式中： $\mathbf{W}$ 为权重矩阵； $\alpha = 1/(1 + \mu)$ ； $\mathbf{D}$ 为度矩阵。以顶部图像边界为例，将此侧的节点用作查询，其他节点用作未标记的数据，顶部边界先验的基于背景的显著性为

$$S_t(i) = 1 - \bar{f}^*(i), \quad i = 1, 2, \dots, N. \quad (11)$$

式中： $i$ 是索引图上的超像素节点；用归一化排名分数  $\bar{f}^*$ 来测量节点的显著性。类似地，使用底部、左、右图像边界来计算其他三个图  $S_b$ 、 $S_l$ 和  $S_r$ ，表示为

$$S_{bq}(i) = S_t(i) \times S_b(i) \times S_l(i) \times S_r(i). \quad (12)$$

通过(11)、(12)式分别计算他们对应的侧面特定背景显著图，接下来基于背景估计的互补值生成

前景显著性估计,

$$S_{\text{fg}}(i) = \bar{f}^*(i), i = 1, 2, \dots, N. \quad (13)$$

MR 算法计算效率高,得到的显著图效果也优于当时很多算法,但是用作背景查询的四个边界对于背景显著性检测不完全准确,并且完全依赖于超像素分割(SLIC)算法,使整个算法存在缺陷。为了克服这些缺点,Li 等<sup>[30]</sup>提出一种正规化随机游走排序,其权重矩阵和度矩阵的构造与 MR 算法类似,但进一步将  $n \times n$  拉普拉斯矩阵定义为

$$L = \begin{cases} \mathbf{d}_u, & \text{if } u = v \\ -\mathbf{w}_{uv}, & \text{if } x_u \text{ and } x_v \text{ are adjacent node.} \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

在逐像素图中使用  $u$  和  $v$  作为元素下标来区分超像素图中使用的  $i$  和  $j$ 。令  $\mathbf{p}^k = [p_1^k, \dots, p_n^k]^T$  表示标签  $k$  的概率向量,该方法提出如下方程的拟合约束,提供了对所有像素的先验显著性估计,而不是单独的种子像素。

$$\text{Dir}[\mathbf{p}^k] = \frac{1}{2} (\mathbf{p}^k)^T L (\mathbf{p}^k) + \frac{\mu}{2} (\mathbf{p}^k - \mathbf{Y})^T (\mathbf{p}^k - \mathbf{Y}), \quad (15)$$

式中: $\mu$  为控制参数; $\mathbf{Y}$  为逐个像素的特征向量。

正则化随机游走排序(WR)算法不仅大大改善了先前显著性估计步骤的显著性图,而且显著性优于随机游走,但是并非所有的邻域节点都适用于表示空间关系,尤其是对于与标记节点不均匀或不相干的邻域节点。

### 3.2.2 基于像素扩散过程的算法

2017 年,Zhou 等<sup>[31]</sup>利用扩散过程对两层稀疏图进行显著区域检测(DP)算法。该算法以超像素为节点,通过将每个节点与相邻节点连接,以及与相邻节点共享一个共同边界的最相似的节点连接而生成,构造了一个 2 层稀疏图。首先将图像抽象为一个超像素图像并运用  $K$  均值算法对图像进行分组聚类,利用超像素聚类的紧致性得到前景和背景种子,然后进一步计算基于前景和背景种子的两种互补显著性映射,生成最终的结果。

DP 算法使用下式计算集群的空间方差,

$$X_{\text{sv}}(j) = \frac{\sum_{i=1}^N h_{ij} \times n_i \times \|b_i - \mu_j\|}{\sum_{i=1}^N h_{ij} n_i}, \quad j = 1, 2, \dots, K, \quad (16)$$

式中: $n_i$  为超像素的像素数; $b_i$  为超像素的质心; $h_{ij}$  为扩散过程之后的相似性矩阵。

使用节点分别和前景种子、背景种子之间的加权相似性  $a_{ij}$  计算显著性图,分别表示为

$$S_{\text{fg\_p}}(i) = \sum_{C_j \in \text{FG}} p(j) \cdot a_{ij}, \quad (17)$$

$$S_{\text{bg\_p}}(i) = \prod_{C_j \in \text{BG}} (1 - a_{ij}), \quad (18)$$

式中: $S_{\text{fg\_p}}$  和  $S_{\text{bg\_p}}$  被归一化到 0 和 1 之间的范围。

DP 算法可以更完整地突出显著区域,更好地抑制背景区域。

### 3.2.3 基于稀疏残差的算法

与使用紧凑型,连续性和背景先验的上述方法不同,Ye 等<sup>[32]</sup>认为对于输入图像,具有较高预测误差的像素比其他像素更突出,将显著性视为预测误差,提出通过稀疏残差和离群点检测(SROD)显著性模型。

将图像  $I$  表示为超完备字典  $D$  的稀疏线性组合,近似地, $I = DX$ , $Y$  表示 Lab 图像块。

$$\min_{x_i} |y_i D_{x_i}|_2, \text{ s.t. } |x_i|_0 \leq \epsilon, \quad \text{for } i = 1, 2, \dots, N. \quad (19)$$

设  $F$  是字典  $D$  的倒数,并将  $F$  视为每个图像补丁的滤波器组,从而得到响应映射  $R = F \times I$ 。

由于规则像素的滤波响应高于显著像素的滤波响应,将输入图像和响应图像之间的差异定义为稀疏残差(SR),由 SR 计算的误预测误差生成关注图,

$$S_{\text{SR}} = \|I - R\|_2^2. \quad (20)$$

假设不可预测误差为异常值得分(OS)的度量,对于输入图像  $I$ ,将像素的不可预测误差定义为 OS,并生成 OS 的视觉注意图,

$$S_{\text{OS}}(p_i) = \min_{p_j \in \tau(p)} \|p_i p_j\|_1. \quad (21)$$

最后将  $S_{\text{SR}}$  和  $S_{\text{OS}}$  采用导向滤波器进行融合,得到最终显著性图。

由于稀疏编码很难求解系数,在学习字典上也比较困难,没有足够快的方法来平衡显著性模型的性能和效率,所以 SROD 算法要致力于提高稀疏残差编码的时效性。

## 3.3 基于多个视觉线索的显著性模型

大多数的检测显著性模型都采用单一的特征,如强度、方向、颜色等来确定图像像素或区域与其周围环境之间的对比度。尽管已经取得了良好的效果,但也有自身的局限性,为了突破限制,还有一些方法结合了多个视觉线索。

### 3.3.1 基于紧凑性和局部对比度的算法

Zhou 等<sup>[33]</sup>提出了基于扩散的紧凑性和局部对

比度的显著性区域检测 (IDCL) 算法。首先将图像抽象为超像素并构造图形  $G=(V,E)$ , 接下来, 使用基于扩散的紧凑性检测显著性图  $S_{\text{com}}$  和局部对比度检测显著性图  $S_{\text{loc}}$ , 这两个图互为补充, 通过直接整合两个不同的显著图来定义最终的显著性图,

$$S = \text{Norm}(S_{\text{com}} + S_{\text{loc}}). \quad (22)$$

### 3.3.2 基于对比度, 空间和对应的算法

Fu 等<sup>[34]</sup> 提出一种新的基于簇的协同显著性检测 (CB) 算法, 集合了三个视觉注意线索, 即对比度, 空间和对应, 有效地测量群集显著性<sup>[35]</sup>。

给定  $M$  张图像  $\{I^j\}_{j=1}^M$ , 对数据点聚类后得到  $K$  个簇群  $C^k$ ,  $\omega^c(k)$  是其与其他群集的特征对比,  $\omega^s(k)$  为  $C^k$  的空间线索,

$$\omega^s(k) = \sum_{i=1, i \neq k}^K \left( \frac{n^i}{N} \|\mu^k - \mu^i\|_2 \right), \quad (23)$$

$$\omega^c(k) = \frac{1}{n^k} \sum_{j=1}^M \sum_{i=1}^{N_j} \{N(\|z_i^j - o^j\|_2) \times \delta[b(p_j^i) - C^k]\}, \quad (24)$$

式中:  $n_i$  表示聚类  $C_i$  的像素数;  $N$  表示所有图像的像素数。

使用所有聚类中的得分分布将每个提示图标准化为标准高斯。簇级共显著性概率定义为

$$p(C^k) = \prod_i \omega_i(k), \quad (25)$$

式中:  $\omega_i(k)$  表示显著性线索。

平滑每个像素的共显著值, 属于簇  $C^k$  的像素  $x$  的显著性似然满足高斯分布  $N$ ,

$$p(x|C^k) = N(\|v_x, \mu^k\|_2 | 0, \delta_k^2). \quad (26)$$

通过对所有聚类上的联合显著性  $p(x|C^k)$  求和得到边际显著性概率为

$$p(x) = \sum_{k=1}^K p(x, C^k) = \sum_{k=1}^K p(x|C^k) p(C^k). \quad (27)$$

共显著性是检测多个图像的共同显著性, 目前并未引起太多的关注。单个图像算法提取每个图像的显著性特征, 容易与复杂背景混淆, 缺少关于多个图像的相关性信息。

## 4 实验结果对比分析

通过对近几年的算法做相关实验, 采用多种不同视觉效果测试图像, 进行了显著性检测性能对比。为了进一步展示各种显著性检测算法的性能, 采用了 WR, IDCL, SR, WT, DP, MCSR, SROD, MR, CB 等几种经典算法, 通过对比分析发现, 基于变换域的显著

性方法处理速度快, 但是频域分析得到的结果会丢失部分图像信息, 导致显著图不能明亮清晰的表现出来。正则化随机游走排序算法大大改善了显著图的效果, 但是并非所有的邻域节点都适用于表示空间关系, 不能适用于所有图像。基于图形的方法用单一的特征检测显著性信息, 有自身的局限性, 容易与复杂背景混淆, 缺少关于多个图像的相关性信息。结合深度特征的显著性模型, 可以提供重复出现的共同显著性对象, 感知深度信息, 具有高效性和有效性。

为了更好的评估各种算法检测性能, 除了给出  $S$  测度 ( $S$ -measure)、标准化扫描路径显著性 (NSS)、曲线下面积 (AUC) 等典型显著性检测测度外, 结合多尺度深度信息, 给出了一种全新的评价指标, 称之为多尺度深度边缘 (MDE)。

图 2 显示了评估方法的一些显著性图, 可以看到, SR, WT, SROD 算法没有很好地引入深度信息, 实际视觉显著性的潜在机制仍然没有被探索, WR, DP, MR, CB 的算法相对可以更统一突出显著区域并保留比其他地方更精细的对象边界, 简而言之, 在现有的深度显著性检测算法研究中, 虽然结合了局部分组线索和边界先验, 但是依然需要更好的结合图像深度信息, 显示出更精准的预测人类固定的能力。

### 4.1 基于多个视觉线索的显著性模型评估

#### 1) 显著性检测模型性能度量之 $S$ -measure

$S$ -measure 值基于区域视角及对象视角两个观察结果提出结构相似性度量, 区域感知度量通过组合所有对象部分的结构信息来捕获全局结构信息。对象感知相似性度量通过比较显著图 (SM) 和地面实况图 (GT) 中前景和背景区域的全局分布来获得结构信息。

对具有区域感知和对象感知的结构相似性评估进行定义, 可以将最终度量表示为,

$$S = \alpha \times S_o + (1 - \alpha) \times S_r, \quad (28)$$

式中:  $\alpha \in [0, 1]$ , 实际设置  $\alpha = 0.5$ 。

#### 2) 显著性检测模型性能度量之 NSS

NSS 是模型的显著性图中人眼位置的响应值的平均值, 其已被归一化为具有零均值和单位标准偏差。

#### 3) 显著性检测模型性能度量之 AUC

通过改变显著性图的分类阈值, 可以得到接受者操作特性曲线 (ROC)。显著性检测指标 AUC 是一种用来度量分类模型好坏的标准, 被定义为 ROC 曲线下的面积。对应 AUC 值越大, 则效果越好。表 1 记录了几种算法的  $S$ -measure 值、NSS 值、AUC 值以及运行时间。

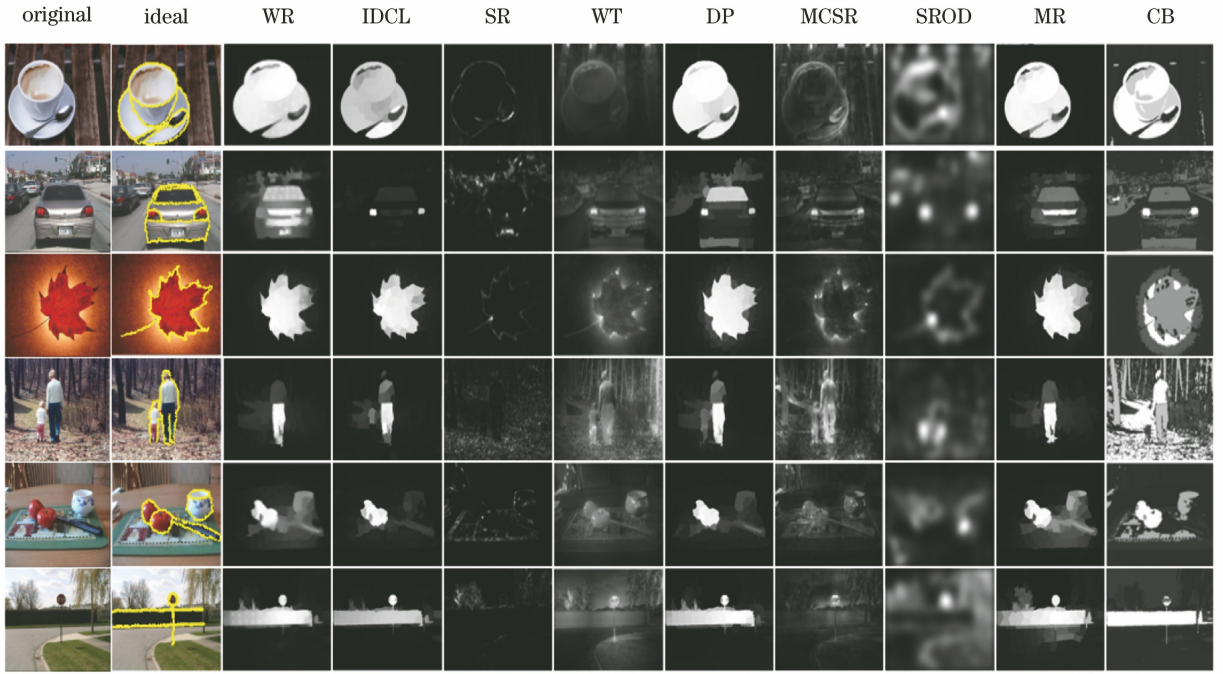


图 2 不同深度敏感信息检测结果对比图

Fig. 2 Comparison of different depth sensitive information detection results

4) 多尺度深度边缘评价 MDE

本文提出了一种基于多尺度的深度边缘评价模型,对原图分别进行 Canny、Sobel、Roberts、Log、Prewitt 五种方法的边缘检测,进行融合后可以得到图像的深度权重值,如图 3(f) 所示, [0.4, 0.4, 0.4] 是在图像像素坐标 [171, 210] 处的 RGB 值,融合后的图像在每一个像素坐标上都有 RGB 深度

信息。

该方法先对原图及几种经典边缘检测算法进行了多尺度评价,在三个不同尺度上记录了原图与边缘图的均方误差(MSE),如表 2 所示。

$$E_{MSE} = \frac{\sum_{i=1}^r (n_i - 1) S_i^2}{N - r} \quad (29)$$

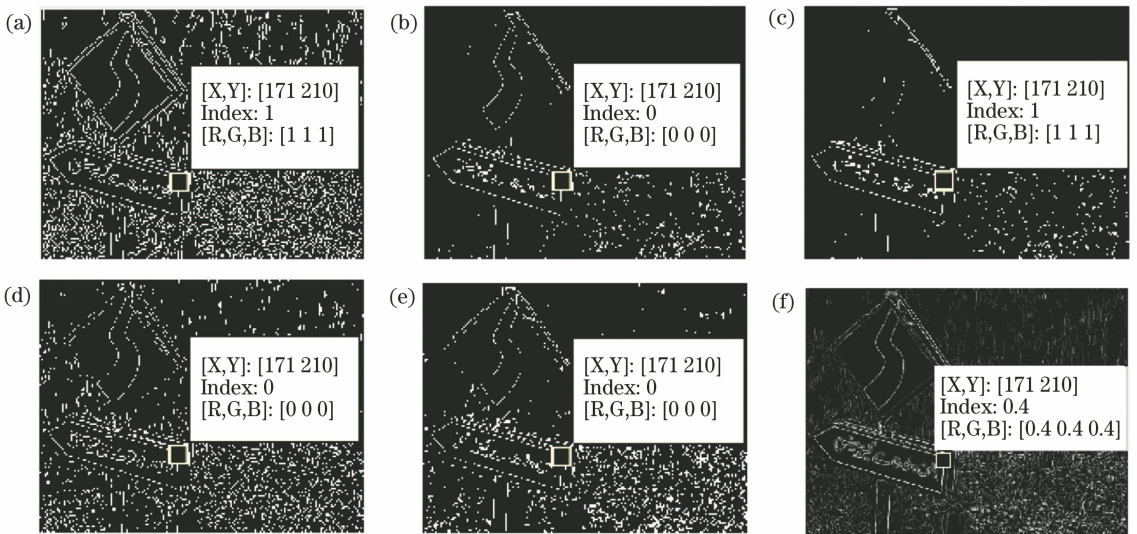


图 3 各算法边缘检测图及深度融合。(a) Canny;(b) Sobel;(c) Roberts;(d)Log;(e) Prewitt;(f)融合图像

Fig. 3 Edge detection map and depth fusion map. (a) Canny; (b) Sobel; (c) Roberts; (d) Log; (e) Prewitt; (f) fusion image

表1 不同检测结果的评估指标

Table 1 Evaluation indicators for different test results

Model	S-measure	NSS	AUC	Time /s
WR	0.5478	0.9479	0.6042	1.3941
IDCL	0.6364	1.4657	0.6477	1.9358
SR	0.4040	0.1233	0.5086	0.3396
WT	0.7292	1.6100	0.7149	4.2347
DP	0.6730	1.4622	0.6085	2.2354
MCSR	0.7053	1.4959	0.6775	2.4578
SROD	0.6167	1.0879	0.6712	4.7010
MR	0.7162	1.2566	0.6484	2.0106
CB	0.7430	1.4844	0.6920	1.9903

将多个不同尺度的图像结合,充分利用小尺度

表3 显著图与深度边缘融合图在不同尺度上的均方误差( $\times 0.001$ )Table 3 Mean square error of saliency and depth edge fusion graphs at different scales ( $\times 0.001$ )

Scale spatial	WR	IDCL	SR	WT	DP	MCSR	SROD	MR	CB
One	0.0665	0.2898	0.4381	0.4010	0.2820	0.6476	0.4454	0.5519	0.6433
Two	0.2597	0.4319	0.4594	0.4235	0.4009	0.5758	0.9015	0.4059	0.6566
Three	1.4454	2.1541	1.6154	1.5154	2.1759	2.0144	1.5476	2.0012	2.6487

## 4 结 论

本文对图像显著性检测及深度敏感信息提取技术进行了分类和概括,并通过对比实验分析了几种典型深度敏感信息提取算法,结果证明,适当引入图像深度感知信息,则显著性性能可以得到改善。例如,基于深度测量的方法通常具有更好的效果,因为它可以进一步从图像中捕获有效的深度属性。结合图像像素对应关系和深度提示,深度图像(RGBD)共显性检测模型可以实现更好的性能。在大多数方法中,深度信息或者用作补充颜色特征的附加特征,或者被设计为进一步表示深度属性(如形状)的度量。虽然基于深度测量的方法可以实现更好的性能,但是,如何有效地利用深度信息来增强对象的识别尚未达成共识。图像间对应关系通常表示为聚类过程、匹配过程、传播过程或学习过程,然而,这些方法具有噪声敏感、耗时的缺点,因此,如何准确地捕获图像间对应深度感知信息是一个迫切需要解决的问题,此外,设计高效实时的系统也是值得研究的,在今后的研究中,将充分利用深度显著性检测的深度学习优点。

## 参 考 文 献

[1] Wang L, Xue J R, Zheng N N, et al. Automatic salient object extraction with contextual cue [C] // 2011 International Conference on Computer Vision,

表2 原图与边缘检测图在多尺度上的均方误差

Table 2 Mean square error of multi-scale on original and edge detection maps

Multi-scale spatial	Canny	Sobel	Roberts	Log	Prewitt
One	0.0013	0.0031	0.0034	0.0023	0.0023
Two	0.0015	0.0080	0.0087	0.0049	0.0052
Three	0.0062	0.0189	0.0214	0.0073	0.0076

影像和大尺度影像的优点,既能检测出显著图深度边缘的整体轮廓,又能对图像细节进行评价。

同理,对不同方法的显著图与融合后的深度边缘检测图  $F$  也进行多尺度评价,计算其均方误差,结果如表3所示。

November 6-13, 2011. Barcelona, Spain. IEEE, 2011.

- [2] Jung C, Kim C. A unified spectral-domain approach for saliency detection and its application to automatic object segmentation[J]. IEEE Transactions on Image Processing, 2012, 21(3): 1272-1283.
- [3] Vu K, Hua K A, Tavanapong W. Image retrieval based on regions of interest[J]. IEEE Transactions on Knowledge and Data Engineering, 2003, 15(4): 1045-1049.
- [4] Wan S H, Jin P Q, Yue L H. An approach for image retrieval based on visual saliency [C] // 2009 International Conference on Image Analysis and Signal Processing, April 11-12, 2009. Linhai, China. IEEE, 2009.
- [5] Chen H R, Qian K L, Wang B. Temporal coherent video segmentation with support vector machine and graph cut[J]. Journal of Computer-Aided Design & Computer Graphics, 2017, 29(8): 1389-1395.  
陈华榕, 钱康来, 王斌. 结合支持向量机和图割的视频分割[J]. 计算机辅助设计与图形学学报, 2017, 29(8): 1389-1395.
- [6] Zhang Y B, Song W. Block weighted image retrieval method based on visual features [J]. Journal of Frontiers of Computer Science & Technology, 2017, 11(3): 468-477.  
张玉兵, 宋威. 视觉特征的分块加权图像检索方法 [J]. 计算机科学与探索, 2017, 11(3): 468-477.
- [7] Cheng X, Zhang Y F, Liu Y, et al. Object tracking

- algorithm based on deep feature [J]. Journal of Southeast University (Natural Science Edition), 2017, 47(1): 1-5.
- 程旭, 张毅锋, 刘袁, 等. 基于深度特征的目标跟踪算法[J]. 东南大学学报(自然科学版), 2017, 47(1): 1-5.
- [8] Huang D M, Wu G J, Wei L F, et al. A secure outsourcing search scheme for encrypted remote sensing images [J]. Laser & Optoelectronics Progress, 2019, 56(3): 031001.
- 黄冬梅, 吴国健, 魏立斐, 等. 一种加密遥感图像的安全外包搜索方案[J]. 激光与光电子学进展, 2019, 56(3): 031001.
- [9] Yarbus A L, Haigh B. Eye movements and vision [J]. Quarterly Review of Biology, 1968, 43(3): 360.
- [10] Chen N E, Chen Y. Image saliency detection using background and centroid prior[J]. Journal of Chinese Computer Systems, 2016, 37(10): 2371-2374.
- 陈南而, 陈莹. 采用背景和中心先验的图像显著性检测[J]. 小型微型计算机系统, 2016, 37(10): 2371-2374.
- [11] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10): 1915-1926.
- [12] Shen X H, Wu Y. A unified approach to salient object detection via low rank matrix recovery[C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012. Providence, RI. IEEE, 2012.
- [13] Jiang P, Ling H B, Yu J Y, et al. Salient region detection by UFO: uniqueness, focusness and objectness[C] // 2013 IEEE International Conference on Computer Vision, December 1-8, 2013. Sydney, Australia. IEEE, 2013: 1976-1983.
- [14] 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.
- [15] Li Y, Tong G F, Yang J C, et al. 3D point cloud scene data acquisition and its key technologies for scene understanding [J]. Laser & Optoelectronics Progress, 2019, 56(4): 040002.
- 李勇, 佟国峰, 杨景超, 等. 三维点云场景数据获取及其场景理解关键技术综述[J]. 激光与光电子学进展, 2019, 56(4): 040002.
- [16] Li Z D, Zhong Y, Chen M, et al. Fast face image retrieval based on depth feature [J]. Acta Optica Sinica, 2018, 38(10): 1010004.
- 李振东, 钟勇, 陈蔓, 等. 基于深度特征的快速人脸图像检索方法 [J]. 光学学报, 2018, 38(10): 1010004.
- [17] Gu K, Zhai G T, Yang X K, et al. Automatic contrast enhancement technology with saliency preservation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(9): 1480-1494.
- [18] Li F Y, Huo H T, Bai J, et al. Hyperspectral target detection based on sparse representation and adaptive model [J]. Acta Optica Sinica, 2018, 38(12): 1228004.
- 李非燕, 霍宏涛, 白杰, 等. 基于稀疏表示和自适应模型的高光谱目标检测 [J]. 光学学报, 2018, 38(12): 1228004.
- [19] Ouerhani N, Hugli H. Computing visual attention from scene depth[C] // Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, Spain. IEEE Comput. Soc, 2000: 375-378.
- [20] Li S, Wang D, Lu Y T. Method for improving imaging resolution of digital holographic adaptive optical system[J]. Chinese Journal of Lasers, 2019, 46(7): 0709001.
- 李顺, 王地, 陆彦婷. 一种提高数字全息自适应光学系统成像分辨率的方法 [J]. 中国激光, 2019, 46(7): 0709001.
- [21] Harel J, Koch C, Perona P. Graph-based visual saliency [C] // Advances in Neural Information Processing Systems, 2006, 19: 545-552.
- [22] Liu T, Zheng N N, Ding W, et al. Video attention: learning to detect a salient object sequence[C] // 2008 19th International Conference on Pattern Recognition, December 8-11, 2008. Tampa, FL. IEEE, 2008: 1-4.
- [23] Wang J L, DaSilva M P, LeCallet P, et al. Computational model of stereoscopic 3D visual saliency [J]. IEEE Transactions on Image Processing, 2013, 22(6): 2151-2165.
- [24] Fang Y M, Wang J L, Narwaria M, et al. Saliency detection for stereoscopic images [J]. IEEE Transactions on Image Processing, 2014, 23(6): 2625-2636.
- [25] Hou J W. Spectral residual approach implementation[C] // IEEE Conference on Computer Vision and Pattern Recognition, June 17-22, 2007.



- [26] Hou X, Zhang L. Saliency detection: a spectral residual approach [C] // IEEE Computer Society Conference on Computer Vision & Pattern Recognition. IEEE, 2007.
- [27] Imamoglu N, Fang Y M, Yu W W, et al. 2D mel-cepstrum based saliency detection [C] // 2013 IEEE International Conference on Image Processing, September 15-18, 2013. Melbourne, Australia. IEEE, 2013.
- [28] Imamoglu N, Lin W S, Fang Y M. A saliency detection model using low-level features based on wavelet transform [J]. IEEE Transactions on Multimedia, 2013, 15(1): 96-105.
- [29] Yang C, Zhang L H, Lu H C, et al. Saliency detection via graph-based manifold ranking [C] // 2013 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013. Portland, OR, USA. IEEE, 2013: 3166-3173.
- [30] Li N C, Yuan N Y, Cai N W, et al. Robust saliency detection via regularized random walks ranking [C] // IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2015.
- [31] Zhou L, Yang Z H, Zhou Z T, et al. Salient region detection using diffusion process on a two-layer sparse graph [J]. IEEE Transactions on Image Processing, 2017, 26(12): 5882-5894.
- [32] Ye T. Analysis of an intelligent monitoring technology based on severe area detection [J]. Technology and Economic Guide, 2017(28): 77.  
叶涛. 浅析一种基于显著区域检测的智能监控技术 [J]. 科技经济导刊, 2017(28): 77.
- [33] Zhou L, Yang Z H, Yuan Q, et al. Salient region detection via integrating diffusion-based compactness and local contrast [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3308-3320.
- [34] Fu H Z, Cao X C, Tu Z W. Cluster-based co-saliency detection [J]. IEEE Transactions on Image Processing, 2013, 22(10): 3766-3778.
- [35] Xing Y Y. Group image co-saliency analysis based on multi-instance learning [D]. Beijing: Beijing Jiaotong University, 2016.  
邢妍妍. 基于多示例学习的群组图像协同显著性分析 [D]. 北京: 北京交通大学, 2016.