

# Research for Scene Depth Structure of Monocular Image

Dan Sun, Kewei Wu, Yongxuan Sun, Zhao Xie

School of Computer and Information, Hefei University of Technology, Hefei Anhui  
Email: sundan\_1014@163.com, wu\_kewei1984@163.com

Received: Apr. 6<sup>th</sup>, 2018; accepted: Apr. 21<sup>st</sup>, 2018; published: Apr. 28<sup>th</sup>, 2018

---

## Abstract

Due to the inaccurate depth perception of the existing single depth cues, inspired by the architecture of collecting images and successfully building three-dimensional information created by Marr, an approach combined with multiple depth cues is presented for depth ordering of monocular images in outdoor scenes. First of all, the input monocular image is segmented to determine the ordering objects. Then those objects are labeled with regional attribute, which is labeled as one of ground, sky and vertical. The depth of the regions labeled as sky are fixed to the furthest, and the depth of the regions labeled as ground are fixed to the nearest, and other regions are ordered by local occlusion cues and global depth cues. Furthermore, based on above depth cues, we construct a graph model, and belief propagation is leveraged to enforce global depth reasoning. Finally, considering appearance characteristics of the regions, we group regions belonging to the same objects on the same depth layer to account for over-segmentation issues, and a hybrid evolution algorithm is utilized to minimize global energy. The properties of depth ordering of the proposed method are evaluated on the BSDS500 dataset, and the experimental results demonstrate that the depth ordering in this paper is outperform the depth ordering results of GM2014.

## Keywords

Occlusion Detection, Ground Contact Points, Belief Propagation, Depth Ordering

---

# 单目图像深度结构恢复研究

孙 丹, 吴克伟, 孙永宣, 谢 昭

合肥工业大学计算机与信息学院, 安徽 合肥  
Email: sundan\_1014@163.com, wu\_kewei1984@163.com

收稿日期: 2018年4月6日; 录用日期: 2018年4月21日; 发布日期: 2018年4月28日

## 摘要

针对现有的单个深度线索对深度感知不准确的问题,以Marr创建的由采集图像到成功构建三维信息的体系结构为基础,本文提出了一种结合多种深度线索的单目图像深度排序方法,实现室外场景中的单目深度排序任务。首先,对输入的单目图像进行超像素分割,确定要排序的目标。然后将分割的区域进行区域属性标记,分别标记为地面、天空、垂直物三个类别中的某一类。对于标记为天空的区域我们将其深度固定为最远,而标记为地面的区域将其深度值固定为最近,对于标定为垂直物的区域,结合局部区域之间的遮挡关系和消失点线索综合考虑。其次,构建图模型,结合区域与地面接触点位置关系和区域之间的遮挡关系构建图模型,然后使用置信度传播进行全局推理,获取一致的深度排序结果。最后利用区域外观特征,把具有外观像素的超像素区域进行合并,并使用混合进化算法进行全局能量优化,得到最终的深度排序结果。在BSDS500数据集上,验证了本文方法的深度排序性能,实验结果能够说明本文的深度排序表达优于GM2014获取的深度排序结果。

## 关键词

遮挡检测, 地面接触点, 置信度传播, 深度排序

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

场景深度结构恢复是估计出单幅单目二维图像中任一点与摄像机之间的距离关系进而获取真实的场景结构的空间布局表达。人类视觉系统作为基本的双目视觉能够直接感知场景的空间深度结构,也被很快应用到各项计算机视觉研究中。然而,现实生活中,我们获取的大部分图像都是有单个摄像机拍摄的二维图像或二维视频,这些图像或视频缺乏多视角信息,研究起来更具有挑战性。单目视觉场景的深度结构是场景物体间相对位置关系的集合,其基本任务就是判定场景中的物体对之间的相对位置远近关系,也就是单目图像深度排序,能够有效的指导人类从二维图像重构出具有几何信息和照片真实感的三维结构,获得图像的三维立体重构,进行场景的三维现实模拟,同时在视频监控、机器人导航等实际领域中具有广泛的应用。

单目场景深度排序就是判定目标距离摄像设备之间更远或更近的关系。针对这一课题,许多方法被提出,其中大部分工作研究关注在双目视觉系统下视差的计算,或通过学习参数的方法估计深度信息。当前用于获取单目图像深度的方法大致可以分为两类:基于深度线索的深度信息获取和基于传感器的深度信息采集。

单目图像深度估计主要依赖于特定场景中的单目线索,如亮度、阴影、遮挡、凸形、消失点、纹理梯度等低层线索。Palou 等人[1]根据 T 连接中接近平角的区域更可能是遮挡区域以及凸形更可能出现在前景目标区域中,获取目标之间的遮挡关系,并利用该遮挡关系构建深度排序概率图进行全局推理。凸形作为 T 连接进行遮挡判定的补充,可以得到更鲁棒的相对深度估计。然而,仅依靠单一的深度线索是不够的,特别是复杂的场景 Hoiem 等人[2]把图像划分为不同的平面区域,包括地面、天空、垂直物,并在此基础上为深度估计提供全局约束。Zeng 等人[3]除了依靠局部遮挡线索外,另外计算了地面接触点和

消失点之间的相对距离关系以此构建 MRF 模型, 获取全局的深度排序关系, 但该方法依赖于获取可靠的区域分割, 分割的效果直接影响最终的深度排序结果, 对于比较复杂的场景无法获取较好的结果。

由于深度传感器的推广使用, 大量深度数据被获得。郭连朋等人[4]直接使用 Kinect 相机直接获取目标到相机的绝对深度值。Wang 等人[5]使用一个可调整角度的一致性图片在光场图像构建一个遮挡预测框架进行深度估计的算法。Liu 等人[6]使用深度学习的方法来进行深度估计。利用卷积神经网络和连续条件随机场提出深度卷积神经场模型, 在深度卷积神经网络中联合学习条件随机场的一元和二元项。在 3D 设备基础上, 非参传递的数据驱动方法被应用。Karsch 等人[7]采用非参采样方式实现深度估计, 给定输入图像, 该算法利用最近邻方法从深度数据库中选取匹配的候选图像, 并通过时空信息光流和运动估计对扭曲处理后的深度图像进行最优化, 从而实现未知图像的深度估计。此外 Liu 等人[8]也将单目深度估计视为最优化问题, 通过非参方式在深度数据库中检索获取与输入图像相似的深度图, 并利用这些深度图构建数据项进行深度推理, 从而实现深度估计。

本文启发于 Marr 视觉理论框架, 符合 Gestalt 视觉感知过程, 构建了合理的、一致的场景深度布局结构, 体现了计算机视觉、生物视觉、物理学的学科交叉。从场景结构中的高层信息出发, 结合局部深度线索与全局属性的深度排序, 提出了一种新的深度传播机制, 以协调多种深度线索。本文的深度排序框架如图 1 所示。与现有的深度排序方法相比, 本文的贡献主要在于:

- 1) 针对单个深度线索无法获取全局一致的深度排序结果, 从场景结构中的高层信息出发, 我们结合局部深度线索与全局线索估计相对深度关系, 如图 1 中的全局深度线索和绝对深度线索部分。
- 2) 结合多个深度线索, 提出一种深度传播机制, 协调多种深度线索, 并结合区域外观线索, 确保全球一致性的深度排序。
- 3) 在 BSDS500 数据集上, 与 GM2014 [9]结果进行了对比, 实验表明本文方法的性能要优于该对比方法。

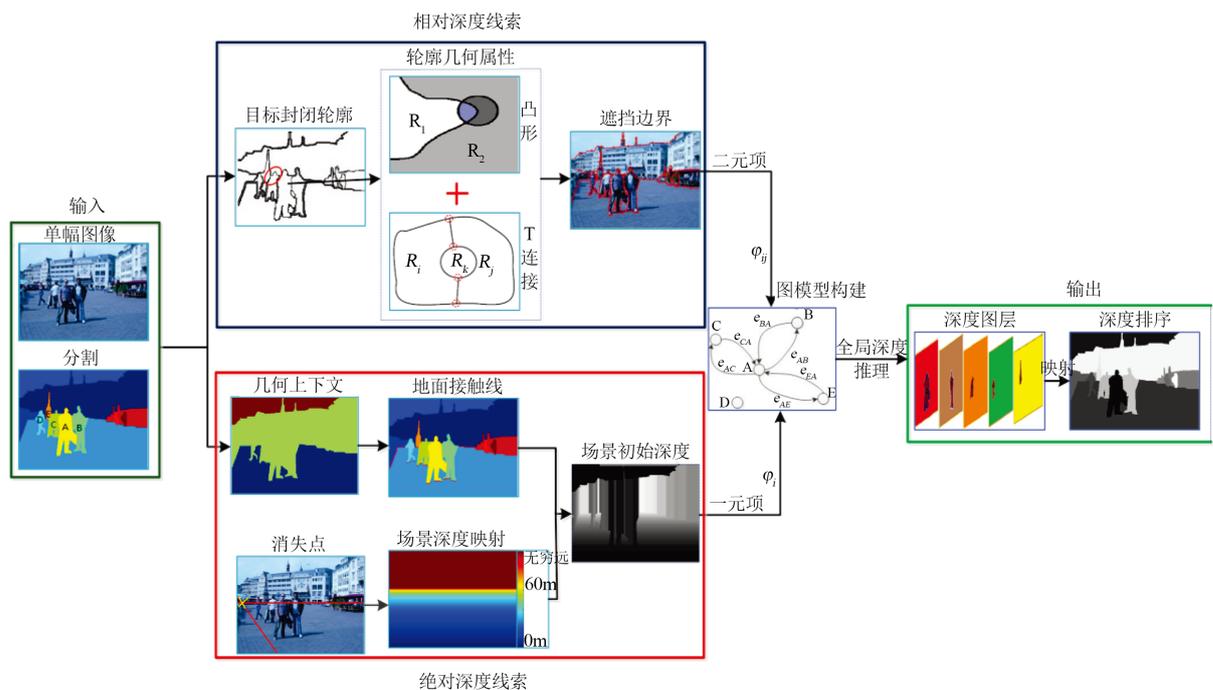


Figure 1. Framework of this paper for depth ordering  
图 1. 本文深度排序框架

## 2. 结合全局和局部深度线索的深度排序

当前主要使用深度线索获取单目图像的深度排序关系。遮挡作为确定局部区域之间深度关系的重要依据被广泛应用。现实生活中各个目标至今存在着各种各样的遮挡，通过有效的全局推理可以获得全局的局部排序关系。然而，遮挡作为深度线索进行全局深度排序存在以下问题：1) 目标之间的遮挡关系判定有误；2) 局部区域之间确定的深度关系无法确定全局目标之间关系；3) 区域之间的外观相似且遮挡信息缺失，无法确定遮挡关系。

另外，消失点作为全局深度线索也被用于深度排序。地面作为真实场景中目标之间相对位置之间关系重要的依据，被考虑到单目深度排序模型中。对于与地面有接触的区域，直接根据地面接触点的位置的前后关系进行遮挡关系判定，然而这一线索在解决深度排序时同样面临一些问题 1) 消失点的位置无法确定；2) 对于悬浮的目标不存在地面接触点；3) 有些场景中无法确定地面区域。对于这些场景来说，消失点这一深度线索有效判定场景目标之间深度关系估计。

本文考虑到上述遮挡和消失点线索对深度的优点和缺点，从场景结构中的高层信息出发，联合考虑遮挡局部线索和消失点全局线索构建图模型进行全局深度推理。本文的算法流程如算法 1 所示。用  $N$  表示分割区域数目， $m$  表示能量下降迭代次数， $n$  是解决方案的个数，因此本文的算法复杂度是  $O(mnN)$  与 GM2014，与 GM2014 [9] 的复杂度一样。

**Algorithm 1.** Algorithm of depth ordering on monocular image

**算法 1.** 单目图像深度排序算法

输入：原始图像  $I$ 、图像分割结果  $I_{seg}$

输出：深度排序标记  $\{l_i\}_{i=1}^N$

1. 对于原始输入图像，进行消失点  $v_j$  和消失线  $line$  的估计。
  2. 对中心视野消失线  $line_{cv}$  上的各像素，根据公式  $D_i = \theta/h_i$  获得场景中各水平线上的真实世界深度， $D_i$  为物体的真实距离， $\theta$  为摄像机参数， $h$  为标定物成像物体高度
  3. 对图像分割区域，依次提取图像区域的颜色、纹理、位置等  $M_{train}$  特征，训练 Adaboost 分类器，实现图像区域的多类标记  $label_{test} = M_{train}(feature_{test})$  预测。
  4. 以标签  $label_{test}$  为垂直物的分割区域为对象，获取区域的边界像素集合  $Boundary = r_{vertical} - erode(r_{vertical})$
  5. 分析边界像素两侧的标记，将边界下侧标记为地面的边界像素点保留，获取地面接触点像素集合  $Boundary_g = \{b_{i,j} | label(b_{i,j+1}) = ground, b_{i,j} \in Boundary\}$
  6. 获得像素对应的深度估计集合  $Dset = DepthMap(Boundary_g)$ ，计算其平均值，作为该垂直区域的期望深度估计  $Dexp = \frac{1}{|Dset_b|} \sum_b Dset_b$
  7. 利用区域轮廓几何属性，对区域之间遮挡关系进行判定  $\{e_{ij}\}_{i,j=1}^N$ ， $N$  为分割区域总数
  8. 构建图模型  $G = \langle V, E \rangle$ ， $V$  为分割区域作为图模型节点， $E$  为区域之间的关系
  9. 全局深度优化，通过置信度传播实现
- Endfor
10. 深度图层间信息传递，提取区域外观特征  $f_{color}$
  11. 区域合并，外观相似且深度关系相近的区域合并到同一深度图层  $l$  上
  12. 利用外观特征，重复步骤 10
  13. 所有深度图层合并完毕，得到深度排序标签集合  $\{l_i\}_{i=1}^N$

### 2.1. 构建图模型

给定一张图像和对应的分割图，本文的目标是为了推断这些目标之间的深度排序关系。本文的输入是图像的超像素分割区域，因为在复杂的场景中，不容易获取图像真实的分割，且图像的分割质量对本文方法并无直接的影响。用  $s_1, s_2, \dots, s_N$  表示图像的  $N$  个分割区域，使用  $1, 2, \dots, d_i, d_{i+1}, \dots, D$  表示  $D$  个深度标签 ( $L \leq N$ )，根据区域距离摄像设备之间的距离进行排序。本文考虑全局深度线索来进行缓解仅使用遮挡关系确定的局部深度关系，以排序目标为节点构建图模型，实现全局和局部深度线索的结合，最终根据置信度传播进行能量优化。

为了协调发挥图像局部深度线索和全局深度线索确定的全局深度排序的优势，构建能量函数：

$$E_{graph}(d) = E_c(d; c) + E_b(d, d; e) \tag{1}$$

$E_c(d; c)$  是根据全局深度线索确定的地面接触点位置构建的数据能量函数， $c$  为目标区域与地面区域接触的地面接触点的位置， $E_b(d, d; e)$  为根据局部遮挡线索构建的边界约束项，相邻区域之间的遮挡关系用  $e$  表示。

### 2.2. 全局深度线索

为了获取地面接触点的位置，单幅图像进行区域分割后，首先对分割后的图像中每个区域的属性进行标记，将其标记为地面、天空、垂直物中的某一类，以区域属性标记获得的垂直物为对象，使用尺度为  $3 \times 3$  的圆盘结构算子，进行形态学腐蚀，然后根据形态学腐蚀前后的图像中对应区域之间的差异，获取图像中区域的边界像素集合，根据边界像素的梯度方向，分析边界像素两侧的区域标记，将边界下侧区域属性为地面的边界像素低点保留，以此来获取地面接触点的像素集合。图 2(d)给出了图 2(a)的地面接触点的检测结果。

使用[10]提出的基于期望最大化的算法来估计消失点和消失线，找出属于场景成像区域内部的消失点位置，作为本文深度标定的消失点。根据消失点的位置，获取图像中与该垂直位置相同的水平消失线，来确定中心视野消失线的位置，进而来计算垂直目标的绝对深度。假设相机在同一地面上，地面上某一位置像素点  $pix_i$  的深度可以用  $d = \frac{fh_c}{pix_i - v_o}$  表示，其中  $f$  表示相机焦距， $h_c$  表示相机高度， $v_o$  是地面水平

线的位置。利用该深度估计模型，可以计算各种情况下垂直目标的深度关系映射表。图 2(e)给出了图 2(a)的分割区域深度映射图结果。然而，现实生活中并不是所有的垂直物都与地面有接触，对于与地面没有接触的垂直物，无法使用地面深度映射关系获取深度，因此遮挡线索同样被采用。

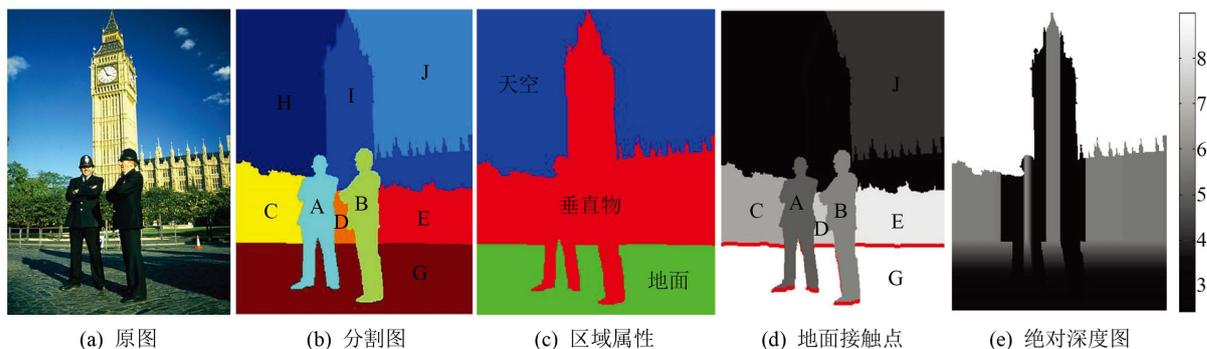


Figure 2. Global depth ordering cues  
图 2. 全局深度线索图

### 2.3. 局部深度线索

根据遮挡确定局部区域间的相对位置关系，一般情况下，遮挡区域相对被遮挡区域来说距离摄像设备更近。遮挡主要发生于 T 连接和凸形中，通常情况下，T 连接的 3 个区域中靠近观察者的区域与其他两个区域边界形成的角度近似为平角，而另外两个区域与其他区域边界形成的角度对应任意角度，同样，具有凸形的自然目标被认为是前景目标，而凹的自然目标更趋向于是背景目标。为了恢复出区域之间的遮挡边界，本文采用 Hoiem 等人[2]的方法，计算区域的边界概率，并使用边界概率的平均值表示边界强度，另外考虑局部区域之间的颜色、位置等特征，相邻的颜色相似的超像素区域更可能对应同一个目标，位置相近的超像素区域也可能属于同一目标，把不大可能属于遮挡边缘的边界进行删除，恢复出区域之间的遮挡关系。

### 2.4. 深度置信度传播

本节综合考虑局部遮挡线索和全局消失点线索，综合两种线索的优势，构建图模型将两种线索结合，并通过置信度传播进行全局推理，得到初始的深度排序结果。本文构建的图模型如图 3 所示。

为了发挥全局消失点线索在单目深度估计中的优势，本文采用垂直物底部的位置估计每个区域被分配在不同的深度图层上的概率似然，以垂直物区域的深度的均值和方差，构建高斯模型，估计分割区域  $s_i$  在为某一深度标签的概率，公式表示如下：

$$E_c(d; c) = \sum_{i=1}^N \varphi_i(d_i, c_i) \tag{2}$$

$$\varphi_i(d_i, c_i) = \varphi_i(d_i; \mu_i, \sigma_i) = -\log \left( \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(d_i - \mu_i)^2}{2\sigma_i^2}} \right) \tag{3}$$

其中  $\mu_i$  为分割区域  $s_i$  的深度均值， $\sigma_i^2$  是分割区域  $s_i$  的深度方差， $d_i$  为区域  $s_i$  的深度标签。本文发现使用高斯分布进行拟合比简单的将所有的可能性分配给每个垂直物区域更加鲁棒。另外，需要注意的是每个图层都被分配了一个非零值，这样可以避免在后期的推理过程中不收敛的情况。

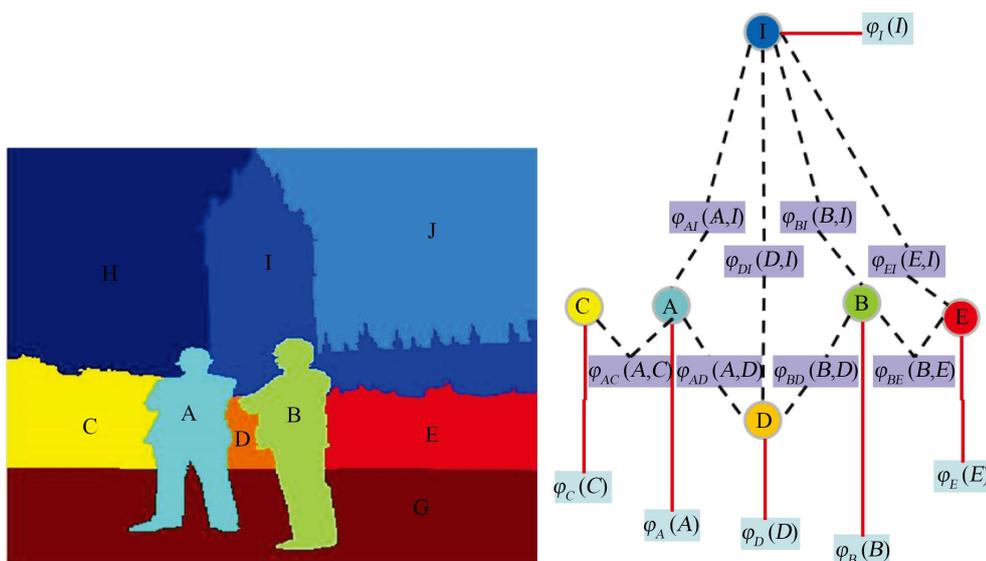


Figure 3. Model of graph  
图 3. 图模型

区域之间边界约束项表征了成对区域之间的遮挡关系。当两个分割区域之间存在遮挡关系，但却给它们分配在同一深度图层上，此时会设置惩罚项来约束此类错误。另外当获取的区域遮挡关系与分配的深度图层相矛盾时，同样会设置惩罚项。边界约束项的设置为：

$$E_b(d, d; e) = \sum_{(i,j)} \varphi_{ij}(d_i, d_j; e_{ij}) \tag{4}$$

$$\varphi_{ij}(d_i, d_j; e_{ij}) = 1 - e^{-b_{ij}^2}, (d_i - d_j) \cdot e_{ij} < 0 \tag{5}$$

这里  $e$  定义了区域之间的边界遮挡关系， $e_{ij}$  定义了区域  $s_i$  遮挡区域  $s_j$ ，相反  $e_{ji}$  表示区域  $s_j$  遮挡区域  $s_i$ ，在本文中，我们采用两种不同类型的遮挡线索：T 连接和凸形。

为了获取合理的单目深度排序结果，本文使用置信度传播进行有效的全局深度推理。置信度传播主要思想是对于图模型中的每个节点，通过并行的消息传播把该节点的状态信息传递给临近的节点，从而影响临近节点的状态信息，经过一定的迭代次数，所有节点的状态信息不再发生变化，每个节点的状态信息趋于稳定，此时每个节点的标签就是我们所求的最优标签。但是对于存在环状结构的图，无法获取最终的收敛结果，为此我们采用[11]中的方法来获取近似收敛结果。通过设定一定的阈值，多次迭代后保证所有节点的置信度总和小于该阈值，我们就认为达到了收敛结果。

### 2.5. 深度图层

给定一张由 Arbeláez 等人[12]提出的算法得到的过分割图像，上面所述的深度线索仅仅考虑了过分割区域之间的局部关系，当两个超像素区域有一样的深度，需要把它们合并到同一深度层。使用  $1, 2, \dots, l_i, l_{i+1}, \dots, L$  表示  $L$  个深度图层 ( $L \leq N$ )，分割区域  $s_i$  在深度图层  $k$  层上表示对应的区域目标被  $1, \dots, k-1$  层上的区域遮挡，并遮挡深度在  $k+1, \dots, L$  上的区域，多个不同的超像素区域可能被分配在同一个图层上，能量函数为：

$$E_{\text{our}}(l) = E_D(l) + E_{SL}(l) \tag{6}$$

$E_D$  表示根据局部遮挡关系和全局消失点信息确定的区域之间的深度排序关系， $E_{SL}$  表示外观相似的超像素区域之间更趋向于在同一深度层。其中：

$$E_D(l) = \sum_{i=1}^{nd} E_d(\lambda_d, D_i^d l_i) \tag{7}$$

其中  $\lambda_d > 0$  表示区域的前景/背景特征的权重系数。 $D = [d_1, d_2, \dots, d_{nd}]$ ，表示超像素区域的深度标签，其中  $nd$  表示超像素区域的深度标签的个数。

$$E_{SL}(l) = \sum_{i=1}^{nh} E_h(\lambda_h, H_i^d l_i) + \sum_{i=1}^{nw} E_w(\lambda_w, W_i^d l_i) \tag{8}$$

其中  $\lambda_h$ ， $\lambda_w$  是颜色特征和纹理特征的权重系数。 $H = [h_1, h_2, \dots, h_{nh}]$ ，表示超像素区域的颜色特征，其中  $nh$  根据颜色特征个数。 $W = [w_1, w_2, \dots, w_{nw}]$  表示纹理特征， $nw$  表示根据纹理特征个数。当判定的所有超像素区域之间的深度关系是正确的， $H_i^d l_i$  和  $W_i^d l_i$  的值都是零，对应的  $E_{SL}(l)$  取最小值，此时是我们寻找的最优结果。我们定义参数化的指数函数来表示能量函数：

$$E(\lambda, x) = e^{-\lambda x} \tag{9}$$

为了获取全局能量优化结果，获得全局一致的深度排序，本文采用混合进化算法[9]来最小化能量函数。

### 3. 实验结果及分析

本文的验证实验室基于 BSDS500 数据集的。该数据集是加利福尼亚大学伯克利分校[13]提出的图像分割数据集，主要是基于室外场景的，有 500 张图像，其中 200 张图像用于训练，200 张图像用于测试，其余的 100 张用于验证试验。该数据集给出了图像分割和边缘检测的真实结果，因此当前许多遮挡检测和单目深度估计方法采用该数据集作为定性评价的数据集。为了与 GM2014 进行比较，本文同样选择 BSDS500 中的 100 张图像进行实验，并给出了与 GM2014 比较的定量和定性结果。测试过程中同样以像素点之间深度关系与真实标记是否一致作为最终的评价标准来验证本文方法的性能。该评价标准沿用了 GM2014 提出的深度排序估计方法，即当成对像素之间的深度关系与真实标记一致时，设置为 1，不一致时，设置为 0，统计 1 的个数占总的像素关系的比例就是该图像的正确率。

本文结合了多个局部特征获取深度排序结果，通常，具有相似外观特征的超像素区域更可能属于同一深度图层。本文把置信度传播获取的深度排序关系作为一种前景/背景约束，另外计算每个超像素区域的颜色和纹理特征，把具有外观相似同时满足深度值相近的超像素区域放在同一深度图层上。对于不同的图像类型，每个特征的重要性都不相同，每个特征权重参数的最优值是由给定的数据集决定的。为了获取 BSDS500 数据集上，每个特征的权重参数，本文采用[9]中所述的 BFGS 方法获取最优的结果，对应的最优的参数为  $\lambda_d = 0.1$ ， $\lambda_h = 4$ ， $\lambda_w = 1$ 。

本文使用[9]中混合进化优化方法来获取最终的深度排序结果。图 4 和表 1 给出了不同方法使用不同的深度线索在 BSDS 数据集上的深度排序结果。针对该定量结果，可以得出：1) 在总体能量层面：本文方法在 BSDS500 的 100 张图像上取得了更好的性能，平均正确率达到 85.1%，同样优于 GM2014 的平均正确率 83.4%；2) 在单个线索层面：本文方法验证了不同的特征线索对深度排序结果的影响，在加入消失点线索后获得了更好的结果，验证了消失点作为深度排序线索的有效性。

针对上述获得的定量结果进行分析：1) 相较于 GM2014 [9]仅采用低层的图像特征进行区域深度排序，本文方法加入高层语义信息进行区域属性标记，并且加入区域的地面接触点与消失点之间的位置关系作为全局限制条件，能够有效减小由于区域之间遮挡关系判定有误导致深度排序错误的问题，因此，本文的深度排序性能优于 GM2014 方法；2) 由于 GM2014 中缺乏高层语义信息，可能会把属于同一目标的超像素区域放在不同的深度图层。由于没有高层语义信息，仅仅依靠区域之间的同层关系不能把这些

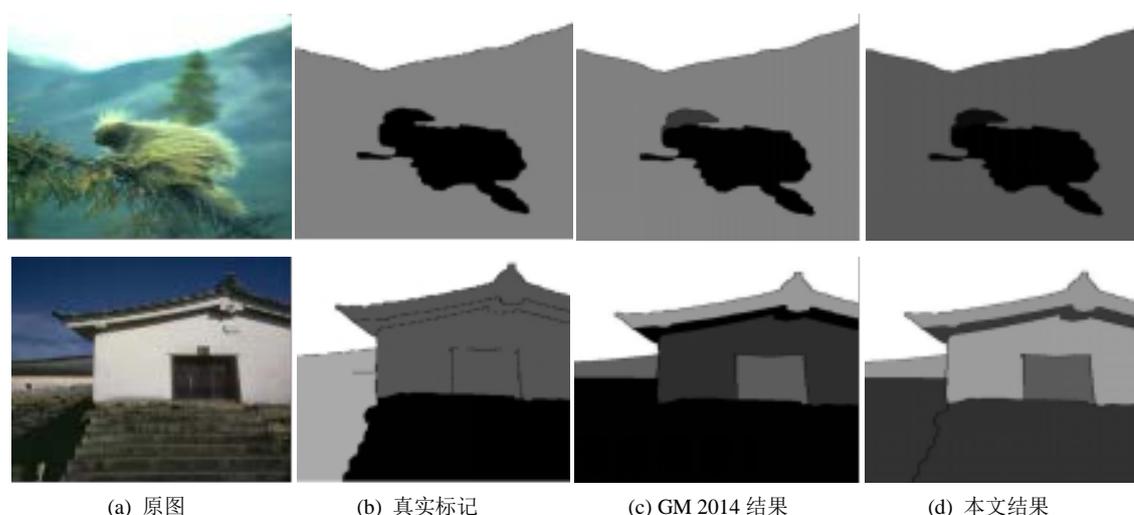


Figure 4. Examples for depth ordering on BSDS500 data set  
图 4. BSDS500 数据集上深度排序实例

**Table 1.** Accuracy of depth ordering with different depth cues  
**表 1.** 不同深度线索对深度排序的影响

检测方法	线索							
	消失点	区域面积	显著性	凸形	T 连接	纹理	颜色	所有线索
GM2014	—	—	61.29	—	48.66	44.73	51.21	83.4
本文方法	61.87	—	—	75.82	41.95	44.22	53.02	85.1

区域进行合并，而本文方法进行了区域属性标记，可以直接把天空和地面区域直接标记出来，同时使用了深度传播机制能够获取更加一致的全局深度排序，综合来看，本文的深度排序优势更明显。

#### 4. 结束语

现有的仅使用单个深度线索进行深度排序存在不足，针对该问题，本文联合考虑局部遮挡线索和全局布局约束，通过构建图模型，研究多个深度线索对深度排序的影响，最后在数据集 BSDS500 上证明了本文方法的有效性。

#### 致 谢

感谢国家重点研发计划(NO. 2017YFB1002203)和国家自然科学基金(NO.61503111; NO.61501467)的支持。

#### 参考文献

- [1] Palou, G. and Salembier, P. (2013) Monocular Depth Ordering Using T-Junctions and Convexity Occlusion Cues. *IEEE Transactions on Image Processing*, **22**, 1926-1939. <https://doi.org/10.1109/TIP.2013.2240002>
- [2] Hoiem, D., Efros, A.A. and Hebert, M. (2011) Recovering Occlusion Boundaries from an Image. *International Journal of Computer Vision*, **91**, 328-346. <https://doi.org/10.1007/s11263-010-0400-4>
- [3] Zeng, Q., Chen, W., Wang, H., et al. (2015) Hallucinating Stereoscopia from a Single Image. *Computer Graphics Forum*, **34**, 1-12. <https://doi.org/10.1111/cgf.12536>
- [4] 郭连朋, 陈向宁, 刘彬. Kinect 传感器的彩色和深度相机标定. *中国图象图形学报*, 2014, 19(11): 1584-1590.
- [5] Wang, T.C., Efros, A.A. and Ramamoorthi, R. (2016) Depth Estimation with Occlusion Modeling Using Light-Field Cameras. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **38**, 1-1. <https://doi.org/10.1109/TPAMI.2016.2515615>
- [6] Liu, F., Shen, C., Lin, G., et al. (2016) Learning Depth from Single Monocular Images Using Deep Convolutional Neural Fields. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **38**, 1-1. <https://doi.org/10.1109/TPAMI.2015.2505283>
- [7] Karsch, K., Liu, C. and Kang, S.B. (2012) Depth Extraction from video Using Non-Parametric Sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **36**, 775-788. [https://doi.org/10.1007/978-3-642-33715-4\\_56](https://doi.org/10.1007/978-3-642-33715-4_56)
- [8] Liu, M., Salzmann, M. and He, X. (2014) Discrete-Continuous Depth Estimation from a Single Image. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, IEEE, 716-723. <https://doi.org/10.1109/CVPR.2014.97>
- [9] Yu, C.C., Liu, Y.J., Wu, M.T., et al. (2014) A Global Energy Optimization Framework for 2.1D Sketch Extraction from Monocular Images. *Graphical Models*, **76**, 507-521. <https://doi.org/10.1016/j.gmod.2014.03.015>
- [10] Saxena, A., Sun, M. and Ng, A.Y. (2009) Make3D: Learning 3D Scene Structure from a Single Still Image. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **31**, 824-840. <https://doi.org/10.1109/TPAMI.2008.132>
- [11] Heskes, T., Albers, K. and Kappen, B. (2003) Approximate Inference and Constrained Optimization. *Proceedings of the 19th Conference in Uncertainty in Artificial Intelligence*, Acapulco, 313-320.
- [12] Arbeláez, P., Maire, M., Fowlkes, C., et al. (2011) Contour Detection and Hierarchical Image Segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **33**, 898. <https://doi.org/10.1109/TPAMI.2010.161>

- 
- [13] Martin, D.R., Fowlkes, C. and Malik, J. (2001) Learning to Detect Natural Image Boundaries Using Brightness and Texture. *Proceedings of the 15rd Annual Conference on Neural Information Processing Systems*, British Columbia, **3**, 1255-1262.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2161-8801, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [csa@hanspub.org](mailto:csa@hanspub.org)