

## 基于视差和条纹调制度分析的多物体场景分割

李小晗, 陈璐, 周翔

(西安交通大学机械工程学院, 陕西西安 710049)

**摘要:** 噪声是影响图像分割的重要因素, 文中提出了一种能够在含噪声的真实场景中准确提取出多个物体区域的分割方案。利用基于正弦条纹投影的双目结构光系统, 得到包含目标物体的相位图和视差图。将视差图映射到 U-视差图中, 利用物体和噪声区域在该视差空间的不同形态特征, 采用闭合区域检测算法初步得到各个物体的分割区域, 并结合条纹调制度阈值分析法进一步去除阴影区域的噪声, 最终得到精确的分割结果。客观评价的数据分析表明, 文中提出的分割算法, 不仅对噪声的鲁棒性较好, 还可以有效地将物体与水平支撑面分割开, 在不同场景下具有计算复杂度低, 抗干扰能力强的优势, 分割准确率均在 90% 以上, 最高可达到 99.2%, 平均运行时间为 27 ms。

**关键词:** 正弦条纹投影轮廓术; U-视差图分析; 条纹调制度分析; 物体分割

**中图分类号:** TP931 **文献标志码:** A **DOI:** 10.3788/IRLA20200085

## Segmentation in multi-objects scene based on disparity and fringe modulation analysis

Li Xiaohan, Chen Lu, Zhou Xiang

(School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

**Abstract:** Noise is an important problem that affects image segmentation. A novel scheme was proposed that could accurately extract multiple objects in a noisy real-world scene. The phase map and disparity map were obtained by using the binocular structured light system based on sinusoidal fringe projection. Firstly, a disparity map was transformed into the corresponding U-disparity map. Then, according to the different projection characteristics of object and noise regions in the disparity map, the preliminary segmentation regions were obtained by using the closed region detection algorithm. In addition, the fringe modulation analysis method was used to remove the noise in the shadow region, and the accurate segmentation results were finally obtained. Experimental results and objective evaluation data indicate that the proposed segmentation algorithm in this paper is not only robust to noise but also can effectively separate the object from the horizontal support surface. It has the advantages of low computational complexity and strong anti-interference ability in different scenarios. The average segmentation accuracy is above 90%, and the best accuracy can achieve 99.2%. The average running time is almost 27 ms.

**Key words:** sinusoidal fringe projection profilometry; U-disparity analysis; fringe modulation analysis; object segmentation

收稿日期: 2020-03-01; 修订日期: 2020-04-29

基金项目: 国家科技重大专项 (2017ZX04002001)

作者简介: 李小晗 (1993-), 女, 博士生, 主要从事深度场景理解和机器人三维视觉控制等方面的研究。Email: lixiaohan1993@stu.xjtu.edu.cn

导师简介: 周翔 (1978-), 男, 副教授, 博士生导师, 博士, 主要从事光学三维成像技术, 机器人三维视觉导航等方面的研究。

Email: zhouxiang@xjtu.edu.cn(通讯联系人)

## 0 引言

基于条纹投影轮廓术的三维形状测量和三维成像技术的快速发展<sup>[1]</sup>,进一步地推动了三维信息(例如视差,深度和点云等)在物体识别,三维场景理解和机器人避障等研究领域的应用和发展。特别是三维场景分割技术,其作为计算机视觉领域的一个主要研究方向,得到了广泛关注。

图像分割是指将图像分成若干具有相似性质的区域的过程,是图像语义理解的重要步骤。传统的三维场景分割方法借鉴 RGB 图像分割方法,以相邻像素之间在某些方面的一致性关系,如颜色、纹理等特征作为分割依据,提出了基于边缘信息的方法<sup>[2]</sup>、基于区域信息的方法<sup>[3,4]</sup>和基于阈值的方法<sup>[5]</sup>。这类方法在噪声较少的场景中具有良好的分割效果,然而当目标物体周围的噪声区域较多或者具有支撑面时,则会干扰算法对分割界限的判断和特征提取,出现过分割或欠分割的问题。基于深度学习的三维场景分割方法<sup>[6,7]</sup>将分割问题转化为像素级的分类问题。此类方法可以准确地对存在遮挡,结构复杂的场景进行有效分割。但是,具有计算量大和泛化能力欠缺的问题。

为了更好地解决分割问题,Derek Hoiem 等人<sup>[8]</sup>认为单个物体的一致性特征应该由物理上的连通性来定义,在三维场景中表现为某个区域的视差或者深度连续性。其中,U-V 视差图的提出,使研究者们开始关注视差连续性在物体检测中的重要性。具体来说,通过统计视差图中每一行或每一列的各个视差值所包含的像素点个数,可以得到 U-视差图和 V-视差图。在 V-视差图中,三维空间中的水平面可被投影成斜率不为 0 的倾斜直线。因此在自动驾驶或机器人导航领域,常通过拟合这条直线间接得到场景中的可通行区域<sup>[9]</sup>,如路面等。在 U-视差图中,非 0 区域包含了垂直于道路的障碍物区域,例如灌木丛,花坛,行人和车辆等物体<sup>[10,11]</sup>,经常用于辅助道路检测。值得注意的是,水平面在该图像空间的投影像素值比较小,可通过图像处理的方法进行去除。不难看出,利用 U-视差图的投影特性,即便物体与水平面具有支撑关系,也可以将其正确分离出来。

文中的研究对象是基于正弦条纹投影的结构光设备所得到的相位图和视差图。目前,针对具有相位

信息的三维场景分割的研究还较少。Ji Deng 等人<sup>[2]</sup>首次提出了一种利用包裹相位的分割算法,其认为物体边缘的深度点具有相移不变性,通过检测不同频率的相位图的跳变点 (abrupt-change points) 得到物体之间的边界线。这种方法可以有效分割不同深度的物体,然而噪声区域会干扰算法对边界点的提取。再者,视差图或者深度图是由一系列计算得到的,由于误差的传递,反映在深度图上的噪声区域的构成是复杂的。因此除了依赖相位信息,有必要直接对结果图的噪声形态进行分析和研究。

针对现有三维场景分割算法的局限性,文中提出了一种适用于含噪声的多物体场景的分割算法。其贡献点主要分为两点:充分利用物体的视差连续的特点,对其在 U-视差图的投影特性进行深入研究和分析,初步完成了对目标物体的位置定位和区域分割;还引入了相位信息,通过计算分割区域对应的相位对比度,利用噪声和物体的不同值域,进一步对分割结果进行噪声去除。

## 1 基本原理

### 1.1 正弦条纹投影三维测量原理

文中采用的是基于相移轮廓术 (Phase Shifting Profilometry, PSP) 的结构光三维测量系统,具有非接触,成本低和精度高的优点。投影仪依次投射不同频率的正弦条纹编码图案到被测物体表面,结合相移法获取相位分布。采用四步等步长相移,步长为 $\pi/2$ ,同一频率下变形光栅条纹图像 $I_1 \sim I_4$ 的灰度表达式为

$$I_m(x, y) = A(x, y) + B(x, y) \cos[\phi(x, y) + 2\pi m/M] \quad (1)$$

式中: $m=1,2,3,4$ ;  $M$  是相移步数,文中采用四步相移,相移间隔为 $\pi/2$ 。 $A(x, y)$  为背景光强; $B(x, y)$  为光栅振幅; $\phi(x, y)$  为与物体表面高度相对应的变形条纹相位。由公式 (1) 可得:

$$\phi(x, y) = 2\pi f x + \varphi(x, y) = \frac{\sum_{m=1}^M I_n \sin(2\pi m/M)}{\sum_{m=1}^M I_n \cos(2\pi m/M)} \quad (2)$$

公式 (2) 中的 $\varphi(x, y)$  是包裹相位 (或相位主值),取值范围为 $[-\pi, \pi]$ ,  $f$  为光栅条纹的空间频率。为了获得真实相位值,使用得到的不同频率的包裹相位图,结合相位解包裹技术获得具有连续相位值的相位图。

对于双目结构光来说,能同时获得极线校正后行对准的左相位图 $\phi_l(x_l, y_l)$ 和右相位图 $\phi_r(x_r, y_r)$ 。利用视差原理可以得到视差图 $d(x, y)$ ,用公式表示为:

$$d = x_l - x_r \quad (3)$$

### 1.2 U-视差图的计算和投影特性

U-视差图的概念由 Zheng 等人<sup>[10]</sup>提出,用于 3D 路面驾驶环境分析。如图 1 所示,通过统计视差图中每一列视差值的相同的像素点个数,构建出 U-视差图 $U_d$ 。像素点 $U_d(u, \tilde{d})$ 表示视差图 $d(u, v)$ 中第  $u$  行中视差值为 $\tilde{d}$ 的像素点个数。

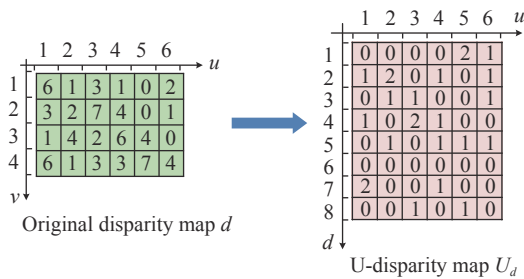


图 1 视差图到 U-视差图的转换示意图

Fig.1 Transformations from the disparity map to U-disparity map

如图 2 所示,假设一个双目视觉系统的左右相机的参数完全相同,位置共面且与地面平行,相对于地面俯仰角为 $\theta$ ,即左右相机光轴 $Z_l$ (或 $Z_r$ )与地面存在夹角。水平面作为三维场景中的一个重要平面,其对分割结果的影响不可忽略。双目视觉系统中理想的水平面,每一行( $v$ 轴)的视差值相等,每一列( $u$ 轴)的视差值随距离相机的远近逐渐变化,因此按列统计的相同视差值的数目不多,对应的像素值比较小。

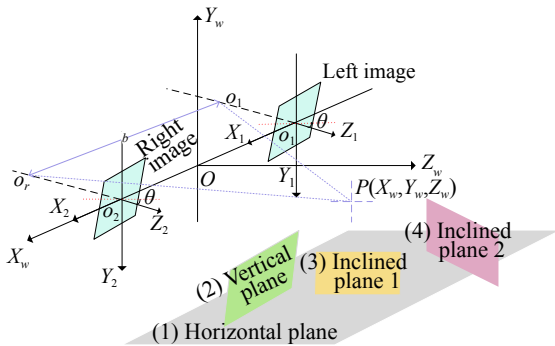


图 2 不同平面与世界坐标系的示意图

Fig.2 Several planes in the world coordinate system

除了水平面,双目系统的世界坐标系中某一非水

平的平面上的三维点 $P(X_w, Y_w, Z_w)$ 与 U-视差图 $u_d(u, d)$ 具有一定的投影关系<sup>[10-11]</sup>,例如以下几种平面:

(1) 垂直面。该平面与 $O-X_w Y_w$ 平行,任一点 $P$ 的 $Z_w$ 值均相等,即 $Z_w = p$ 。其在 U-视差图的投影形态可表示为:

$$d \approx f \frac{b}{p} \cos \theta \quad (4)$$

(2) 倾斜面 1。该平面与 $O-Y_w Z_w$ 平行,任一点的 $X_w$ 值相等,即 $X_w = r$ 。其在 U-视差图的投影形态可表示为:

$$d = (u - u_0) \frac{2b}{2r + b} \quad (5)$$

(3) 倾斜面 2。该平面在三维场景中是最常见的,投影到 $O-X_w Z_w$ 面是一条斜率不为 0 的倾斜直线,则该面上的任一点的 $X_w$ 和 $Z_w$ 的关系可以简化为 $Z_w = \rho X_w + \tau$ 。其在 U-视差图的投影形态可表示为:

$$d \approx \frac{2b}{\rho b - 2\tau} [\rho(u - u_0) - f] \quad (6)$$

公式中 $u_0$ 是图像主点坐标 $(u_0, v_0)$ 的横坐标, $p, r, \rho$ 和 $\tau$ 为常数,用于表述上述平面的特点。公式(4)~(6)可以看作是关于 $u$ 和 $d$ 的二元一次直线方程。换句话说,上述这些非水平的平面在 U-视差图上会被映射为一条直线段,长度由该平面的视差范围决定。

## 2 算法框架

文中所提出的分割算法主要分为两部分:利用 U-视差图分析法,初步获取各个物体所在区域;结合质量导向引导技术<sup>[12]</sup>,利用包裹相位图的调制度(fringe modulation),去除分割结果中的质量不好的区域。算法框架如图 3 所示。

### 2.1 基于 U-视差图分析的初步分割

基于对不同平面在 U-视差图的投影形态的分析做出如下推论:如果将一个物体表面看作由许多不同的小平面组成的集合,结合视差连续性的特点,则该物体会被投影成一个由斜率不同的线段构成的闭合区域。

为验证该推论,图 4 中,对一个石膏模型进行视差计算,并构建出相应的 U-视差图(图 4(c)和(d))。文中的视差图的灰度级均被标准化到 0~255。通过对该物体区域的灰度级进行直方图统计,像素值不为

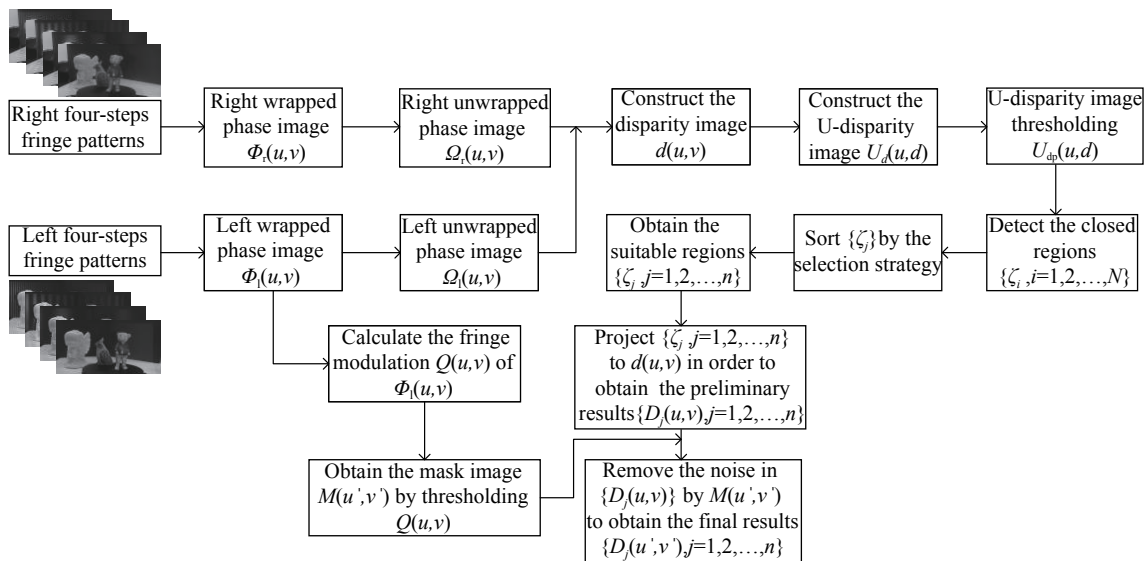


图 3 算法流程示意图

Fig.3 Flowchart of the proposed algorithm

0 的灰度级范围为 [83, 96], 对应 U-视差图的纵轴  $d$  的范围, 横轴  $u$  的范围对应其在视差图上横坐标的范围。不难看出, 相对于图 4(a)(或图 4(b)) 中石膏模型较为复杂的边缘结构, 其映射在 U-视差图的形态接近一个简单的闭合区域, 大大降低了检测的难度。

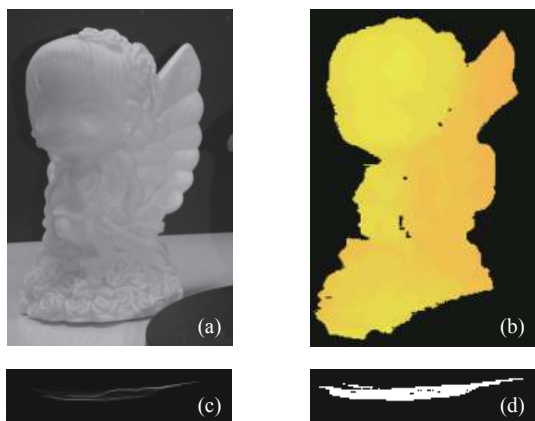


图 4 (a) 石膏模型的灰度图像; (b) 伪彩处理后的视差图; (c) 物体在 U-视差图上的投影形态; (d) 图 (c) 的阈值化图像

Fig.4 (a) Gray image of the plaster model; (b) pseudo color disparity map; (c) projection result on the U-disparity map; (d) binary image of (c)

结合上述分析, 将单个物体的 U-视差图投影特性应用到多物体场景中, 图 3 已给出主要的算法步骤。对 U-视差图  $U_d(u, v)$  的阈值化处理分为两步, 首先进行视差值截取, 去除一部分远距离的孤立点或块; 再

进行二值化处理, 简化区域轮廓提取和面积计算。用公式表示为:

$$U_{dp}(u, d) = \begin{cases} 255, & \text{if } U_d(u, v) \geq \mu \\ 0, & \text{else} \end{cases} \quad (7)$$

式中:  $\mu$  为阈值常数, 用于去除 U-视差图中较小的像素值, 文中设置为 3。得到  $U_{dp}(u, d)$  后, 就可以检测图中的多个闭合区域。然而, 文中使用的是基于光栅投影的三维重建系统, 由于投影光能量的衰减以及环境随机噪声等原因, 视差图中的部分噪声区域的视差值与目标物体相近, 仅靠公式 (7) 并不能将其去除。

针对上述问题, 首先统计各个区域  $\{\zeta_i, i = 1, 2, \dots, N\}$  的轮廓点数目进行排序和筛选; 然后计算对应应在视差图的区域像素点数目  $\{R(\xi_j), j = 1, 2, \dots, n\}$ , 对区域进行再一次筛选。具体策略方法如下:

$$size(\zeta_i) > \alpha \quad (8)$$

$$R(\xi_j) = \frac{S_{nozero}}{S_{total}} \times 100\% > \beta \quad (9)$$

式中: 常数  $\alpha$  和  $\beta$  为阈值, 文中给出的建议值分别是 15 和 0.08;  $size(\cdot)$  表示区域  $\zeta_i$  边缘的离散点数;  $R(\cdot)$  表示  $\zeta_i$  反向投影到视差图后的区域  $\xi_j$  的非 0 像素点百分比。公式 (8) 利用噪声在 U-视差图的投影轮廓面积较小的特点, 去除相关区域; 针对视差范围跨度较大但是投影较小的噪声区域, 如图 5 所示, 使用公式 (9) 在视差图的反向投影过程中予以去除。



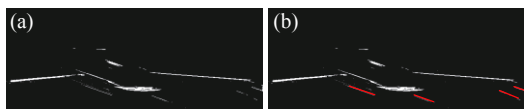


图 5 (a) 二值化后的 U-视差图; (b) 红色斜线段标注的视差范围跨度较大的噪声

Fig.5 (a)Binary U-disparity map; (b)noise with large disparity range labeled with red oblique lines

### 2.2 基于条纹调制度计算的噪声去除

利用物体在 U-视差图的投影特性可以初步完成场景分割。然而,分割结果中仍存在小部分噪声分布在目标物体周围,视差范围与物体相近,并且其 U-视差图中的投影区域往往被包含于物体的闭合区域中,依靠公式 (7)~(9) 无法将其去除干净。物体的阴影区域常会导致此类噪声,其对应在对比度的相位图,因此算法还引入了光栅条纹的调制度信息。条纹调制度与光栅的虚部和实部有关,可定义为:

$$Q(x, y) = \left\{ \left\{ \sum_{m=1}^M I_n \sin(2\pi m/M) \right\}^2 + \left\{ \sum_{m=1}^M I_n \cos(2\pi m/M) \right\}^2 \right\}^{1/2} \quad (10)$$

将公式 (1) 代入公式 (10), 单位统一起见, 用像素坐标表示为:

$$Q(u, v) = \frac{M}{2} B(u, v) \quad (11)$$

由公式 (11) 可知, 调制度与条纹振幅相关。找到上一章节中的分割结果  $\{\xi_j(u, v)\}$  对应条纹图的位置, 并利用公式 (10) 计算调制度, 去除不符合条件的像素值, 并二值化为掩码图像  $M(u', v')$ , 公式表示如下:

$$\begin{cases} M(u', v') = 1, & \text{if } Q^2(u, v) \geq \lambda \\ M(u', v') = 0, & \text{if } Q^2(u, v) < \lambda \end{cases} \quad (12)$$

文中  $\lambda$  取 512。利用公式 (12), 去除不符合条件的像素点, 得到最终的分割结果  $\{\xi_j(u', v')\}$ 。

需要注意的是, 在计算条纹调制度时需要将图像进行畸变矫正和极线校正, 才能满足相位图和视差图的像素一一对应的关系。另外, 该算法的研究对象是全场三维场景。如果直接用公式 (12) 处理视差图再进行分割, 虽然可以去除大部分噪声区域, 但也会导致场景中部分目标物体的缺失。

## 3 实验结果与分析

实验中, 三维测量系统所用的投影仪的像素分辨

率为 1 920 pixel×720 pixel, 投影仪最大投影速度为 60 Hz, 左右相机的像素分辨率均为 640 pixel×360 pixel, 极线距离为 130 mm。为充分利用视场, 目标物体被放置在系统前 0.2~1 m 处。采用四频四步 (频率分别为 1 Hz, 4 Hz, 16 Hz 和 64 Hz, 步长间隔为  $\pi/2$ ) 的方式向场景投影正弦条纹图案, 利用相移轮廓术获取相位图和视差图。文中通过采集实际环境的三维场景对提出的分割算法的可行性和运行速度进行评估。实验的主要工作环境是 Inter(R) Core(TM) i7-8550U CPU 和 8 GB 内存的计算机。算法编程采用 C++ 语言, 并在 Visual Studio 2015 软件上运行。

为了更客观地分析算法的准确性和鲁棒性, 在人眼主观感知的基础上, 使用五种评价指标对分割图与参考图 (Ground Truth) 进行 pixel-to-pixel 比较, 公式表示为:

$$Precision = \frac{S_{TP}}{S_{TP} + S_{FP}} \times 100\% \quad (13)$$

$$Recall = \frac{S_{TP}}{S_{TP} + S_{FN}} \times 100\% \quad (14)$$

$$Fscore = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \quad (15)$$

$$Jaccard = \frac{S_{TP}}{S_{TP} + S_{FN} + S_{FP}} \times 100\% \quad (16)$$

$$Conformity = \left( 3 - \frac{2}{Fscore} \right) \times 100\% \quad (17)$$

式中:  $S_{TP}$  表示分割结果中的正确像素点数目;  $S_{FP}$  表示分割结果中的错误像素点数目;  $S_{FN}$  表示非分割结果中的错误像素点数目。不难看出,  $S_{TP} + S_{FP}$  即分割结果的像素点数目,  $S_{TP} + S_{FN}$  即参考结果的像素点数目,  $S_{TP} + S_{FN} + S_{FP}$  即分割结果和参考结果的并集。公式 (13)~(17) 评估的是分割结果和 Ground Truth 的一致性。数值越大, 说明结果越接近真实情况, 准确性越高。

图 6~8 分别展示了三种典型的三维场景, 具有不同的物体数目, 遮挡程度和噪声占比。其中, 图 6 的物体个数最少, 但是物体周围的噪声干扰较多; 图 7 的两个模型存在遮挡关系, 同样受阴影噪声的干扰; 图 8 的物体个数最多, 被放置在两个不同的支撑面, 存在支撑面 (黑色转台) 上阴影噪声的干扰。

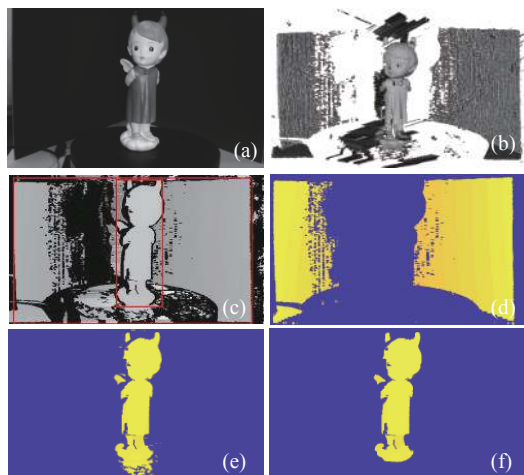


图 6 第一组三维场景分割结果。(a) 灰度图; (b) 三维重建结果; (c) 分割结果在视差图上的标注框; (d),(e) 初步分割结果; (f) 最终分割结果

Fig.6 Scene segmentation results of group 1. (a) Gray image; (b) 3D reconstruction result; (c) labeled regions of segmentation results in the disparity map; (d),(e) preliminary segmentation result; (f) final segmentation result

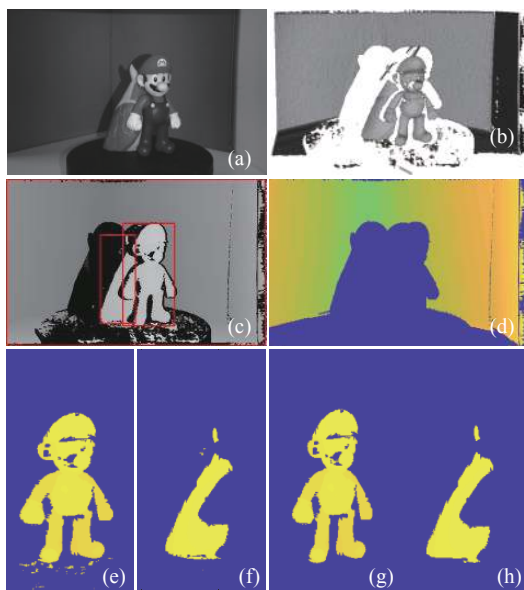


图 7 第二组三维场景分割结果。(a) 灰度图; (b) 三维重建结果; (c) 分割结果在视差图上的标注框; (d)-(f) 背景板和两个模型的初步分割结果; (g),(h) 两个模型的最终分割结果

Fig.7 Scene segmentation results of group 2. (a) Gray image; (b) 3D reconstruction result; (c) labeled regions of segmentation results in the disparity map; (d)-(f) preliminary segmentation results of the background board and the two models; (g),(h) final segmentation results of the two models

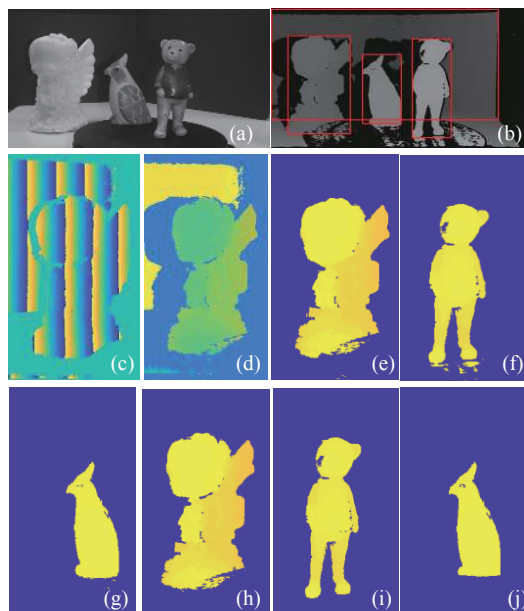


图 8 第三组三维场景分割结果。(a) 灰度图; (b) 分割结果在视差图上的标注框; (c) 分割结果 (e) 在对应的相位图的结果; (d) 图 (c) 对应在原视差图的结果; (e)-(g) 初步分割结果; (h)-(j) 最终分割结果

Fig.8 Scene segmentation results of group 3. (a) Gray image; (b) labeled regions of segmentation results in the disparity map; (c) segmentation region of (e) in the corresponding phase image; (d) segmentation region in the corresponding disparity image of (c); (e)-(g) preliminary segmentation result; (h)-(j) final segmentation result

第一组实验在噪声复杂的场景中进行。图 6(b) 中, 匹配误差以及环境光等因素导致了三维重建结果中物体周围的噪声和无效区域较多。U-视差图分析法可以将目标物体提取出来, 但仍有与其视差范围相近的噪声残留, 如图 6(e) 所示。对照图 6(a) 场景的灰度图发现, 这部分噪声主要是由转台上物体的倒影纹理导致, 结合公式 (12), 利用条纹调制度可进一步予以去除。图 6(e) 的噪声解决思想同样适用于图 7(e) 和图 8(g)。从图 6 的分割结果中可以看出, 文中提出的分割算法对信噪比较低的三维场景具有较好的鲁棒性, 分割准确率不仅能够保留和分割场景中包含背景板和模型在内的有效区域, 还有效解决了三维重建中最常见的阴影噪声问题。

第二组实验在存在遮挡的场景中进行。如图 7(a) 所示, 距离相机较远的模型的部分区域被遮挡, 并且两个模型之间的距离很近。仅仅依靠纹理信息, 很难

将其分离开。然而,结合“同一物体的视差是连续的,不同物体的视差范围也不同”的特点,即文中提出的 U-视差图分析法,可以准确地分割开这两个物体,如图 7(e) 和 (f) 所示。这也说明了视差分析在三维场景分割中的重要性。

第三组的实验场景中不同物体的支撑面也不相同,如图 8(a) 所示。在视差图中,水平支撑物和被支撑物的接触区域往往视差值相近且空间物理距离重合,这些问题都增加了分割的难度。结合 1.2 章节的分析,水平支撑面在 U-视差图的投影区域具有像素值小的特点,因此使用公式 (7) 的阈值化处理,可以减小其对分割结果的影响,最终结果如图 8(h)~(j) 所示。另外,图 8(c) 对应分割结果图 8(e) 对应包裹相位的位置,此时已经过条纹调制度的处理,图 8(d) 是该区域在视差图中的区域,通过计算图 8(d) 和图 8(e) 的交集即可得到最终的分割结果(如图 8(f) 所示)。图 8 表明文中算法在具有多个支撑面的三维场景中,仍能保持较高的分割准确率,以及较好的分割结果。

表 1 分别对图 6~8 的结果进行了客观评价,其中 ①表示只利用 U-视差图分析的分割结果,②表示文中算法的最终分割结果。结合对比三组实验在①和②的结果,文中提出的算法,通过结合相位调制度,可以提高分割结果的准确率,尤其是在噪声较多的场景下。

表 1 图 6~8 的评价指标数据对比

Tab.1 Comparison of evaluation data of Fig.6-8

Evaluation index	Group 1		Group 2		Group 3	
	①	②	①	②	①	②
Precision	91.46%	99.92%	98.37%	99.24%	99.12%	99.17%
Recall	98.65%	98.28%	93.66%	94.55%	98.18%	98.88%
F-score	94.92%	99.09%	96.37%	96.42%	97.95%	98.16%
Jaccard	90.33%	98.20%	92.99%	93.09%	95.99%	95.11%
Conformity	89.29%	98.16%	92.46%	92.58%	95.82%	95.32%

为了验证算法的可行性,还统计了算法的运行速度,三组实验数据的运行时间分别是 33 ms, 21 ms 和 27 ms, 达到了实时分割的速度。由此可见,算法的运行速度与噪声区域面积和物体数目有关,其中由于噪声导致的无效区域增多,将较大程度影响算法的处理

速度。为了不失一般性,使用 Realsense D415 深度相机获取视差图序列,利用算法中 U-视差图分析的部分对场景结构更为复杂的连续帧图像进行分割,如图 9 所示。结果显示,平均耗时为 60.2 ms, 处理速度约为 17 fps, 达到了快速分割的速度。



图 9 提出的算法对 Realsense D415 深度相机获取的视差图序列(第 2 帧)的处理结果。(a) RGB 图像;(b) 视差图;(c) 分割结果在伪彩视差图上的标注框

Fig.9 Disparity sequence obtained by Realsense D415 depth camera by the proposed method. (a) RGB image; (b) disparity map; (c) labeled regions of segmentation results in the rendered disparity map

## 4 结 论

近年来,面向运动物体的实时三维成像技术的研究和应用成果显著。值得注意的是,在结构光三维重建中,相移轮廓术也在向高精度实时重建的方向发展。这类方法同时结合了时域信息和频域信息,将会开启目标检测和跟踪,三维场景理解,人机交互等领域的新的方向。图像分割作为计算机视觉任务的重要步骤,也被广泛关注。在这种趋势下,文中基于物体具有视差连续性的特性,提出了一种针对多物体场景的分割算法。通过构建 U-视差图,降低了对不同物体的提取难度,简化了计算复杂度;并利用条纹调制度对阴影的辨别能力,进一步去除噪声区域,提高了分割精度。不同环境下的实验结果充分验证了提出算法的可行性和有效性,可以快速高效分离出场景中的有效区域,并对噪声、遮挡等问题具有一定的鲁棒性。

文中首次提出了基于 U-视差图的图像分割方法,在几个方面有待进一步的研究和改进:(1) 利用视差信息,分割出的物体存在部分边缘缺失的问题。需要进一步研究边缘等高频区域在 U-视差图的投影特点。(2) 条纹调制度可以解决的噪声问题具有局限性,阈值处理方法单一,其对噪声去除的范围是不可控的。因此如何更充分地利用相位信息提高分割精

度,将是未来需要解决的问题。

#### 参考文献:

- [1] Yin W, Feng S, Tao T, et al. High-speed 3D shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system [J]. *Optics Express*, 2019, 27(3): 2411–2431.
- [2] Deng R J, Li J, Feng H, et al. Flexible depth segmentation method using phase-shifted wrapped phase sequences [J]. *Optics and Lasers in Engineering*, 2019, 122: 284–293.
- [3] Silberman N, Fergus R. Indoor scene segmentation using a structured light sensor[C]//2011 IEEE international conference on computer vision workshops (ICCV workshops). IEEE, 2011: 601-608.
- [4] Liu Ji, Zeng Wenliang, Liang Xiaosheng. Depth assisted active contour segmentation algorithm [J]. *Application Research of Computers*, 2020, 37(1): 287–290.
- [5] Lee D S. Effective Gaussian mixture learning for video background subtraction [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(5): 827–832.
- [6] Zhang Y, Qiu Z, Yao T, et al. Fully convolutional adaptation networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 6810-6818.
- [7] Danielczuk M, Matl M, Gupta S, et al. Segmenting unknown 3d objects from real depth images using mask r-cnn trained on synthetic data[C]//2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019: 7283-7290.
- [8] Hoiem D, Efros A A, Hebert M. Recovering occlusion boundaries from an image [J]. *International Journal of Computer Vision*, 2011, 91(3): 328–346.
- [9] Leng J, Liu Y, Du D, et al. Robust obstacle detection and recognition for driver assistance systems [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 21(4): 1560–1571.
- [10] Hu Z, Uchimura K. UV-disparity: an efficient algorithm for stereovision based scene analysis[C]//IEEE Proceedings. Intelligent Vehicles Symposium. IEEE, 2005: 48-54.
- [11] Li Xiaohan. Research on real-time detection technology of passable area based on binocular stereovision[D]. Beijing: University of Chinese Academy of Sciences, 2018.
- [12] Zhao M, Huang L, Zhang Q, et al. Quality-guided phase unwrapping technique: comparison of quality maps and guiding strategies [J]. *Applied Optics*, 2011, 50(33): 6214–6224.