# 基于深度图像利用遮挡信息确定下一最佳观测方位

张世辉",2) 刘建新"孔令富",2)

<sup>1)</sup>(燕山大学信息科学与工程学院 河北 秦皇岛 066004)
 <sup>2)</sup>(河北省计算机虚拟技术与系统集成重点实验室 河北 秦皇岛 066004)

**摘 要** 文中提出一种新颖的利用深度图像中的遮挡信息确定下一最佳观测方位的方法.该方法首先从某一观测 方位获取视觉目标的一幅深度图像,然后根据已获得的深度图像中的遮挡信息确定出下一最佳观测方位.主要贡 献在于:(1)提出深度图像中最大深度差相邻点的概念,利用其与深度图像中的遮挡边界点可获取遮挡区域外接表 面信息;(2)一种基于投影降维思想的遮挡区域外接表面最佳小平面集合的确定方法,用于确定下一最佳观测方 位;(3)一种基于最佳小平面集合的下一最佳观测方位确定算法.所提方法无需预先获取视觉目标的先验知识及将 摄像机的观测位置限定在固定表面上,适用于具有不同型面的视觉目标.实验结果验证了所提方法的可行性和有 效性.

关键词 深度图像;遮挡信息;下一最佳观测方位;最大深度差相邻点;最佳小平面集合 中图法分类号 TP391 **DOI 号** 10.11897/SP. J. 1016. 2015. 02450

# Determining Next Best View Based on Occlusion Information in Depth Image

ZHANG Shi-Hui<sup>11,2)</sup> LIU Jian-Xin<sup>1)</sup> KONG Ling-Fu<sup>11,2)</sup>

<sup>1)</sup> (School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004) <sup>2)</sup> (The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Qinhuangdao, Hebei 066004)

**Abstract** This paper proposes a novel next best view approach based on occlusion information in a depth image. At first, a depth image of visual object is obtained from one view, and then the next best view is determined by the occlusion information in depth image. This work is distinguished by three contributions. The first contribution is the concept of maximal depth difference adjacent point, which is combined with the occlusion boundary point to obtain external surface information of the occluded region. The second contribution is a determining method based on projection dimension reduction for the optimal small plane set of the occluded region's external surface, which will be used in the process of determining the next best view. The third contribution is a specific next best view algorithm based on the optimal small plane set. The proposed approach does not need the priori knowledge of visual object or limit the observation position of the camera on a fixed surface. In addition, the approach is suitable for the visual object with different surface. Experimental results demonstrate its feasibility and effectiveness.

**Keywords** depth image; occlusion information; next best view; maximal depth difference adjacent point; optimal small plane set

收稿日期:2013-08-26;最终修改稿收到日期:2015-06-05.本课题得到国家自然科学基金(61379065)、河北省自然科学基金(F2014203119)和机器人技术与系统国家重点实验室开放研究基金(SKLRS-2010-ZD-08)资助.张世辉,男,1973年生,博士,教授,博士生导师,中国计算机学会(CCF)会员,主要研究领域为视觉信息处理、模式识别.E-mail:sshbzz@ysu.edu.cn.刘建新,男,1988年生,硕士研究生,主要研究方向为视觉信息处理.孔令富,男,1957年生,博士,教授,博士生导师,中国计算机学会(CCF)高级会员,主要研究领域为机器视觉、智能信息处理.

# 1 引 言

下一最佳观测方位(Next Best View,NBV)的 确定一直是三维重建、机器人导航、自动装配、目标 识别等领域重要且困难的研究课题之一,它根据当 前观测到的信息,确定出摄像机(传感器)的下一观 测方向和位置,使得从该方向和位置能够最大量地 获取所希望观测到的场景或视觉目标的未知信息.

到目前为止,NBV 确定算法所基于的图像信息 类型主要有两种:图像亮度信息和图像深度信息.基 于图像亮度信息的 NBV 确定算法相对较少,比较 有代表性的如:Bottino 等人<sup>[1]</sup>提出一种由轮廓得到 形状的重建算法,该方法基于一个最优重建条件来 判断是否需要进一步观测,若需要,则根据特定的规 则计算出下一最佳观测方位;Trummer 等人<sup>[2]</sup>提出 一种综合在线方法用于规划下一最佳观测方位,从 而达到优化任意目标三维重建结果精度的目标. Haner 和 Heyden<sup>[3]</sup>提出一种利用协方差传播理论 确定下一最佳观测方位的方法.

同 2D 的亮度图像相比, 2.5D 的深度图像更利 于获得场景的三维信息,因此,现有的 NBV 确定算 法一般都基于深度图像实现. Connolly<sup>[4]</sup> 是较早对 NBV 问题进行研究的学者之一,他采用八叉树模型 来描述视觉目标,并根据每个节点实际代表的空间 是否为空将其进行标记,进而确定出下一最佳观测 方位. Maver 和 Bajcsy<sup>[5]</sup> 通过对视觉目标的遮挡区 域进行多边形建模,提出一种利用遮挡信息确定下 一最佳观测方位的方法. Whaite 和 Ferrie<sup>[6]</sup>利用深 度数据建立物体的参数相似模型,将深度数据与当 前模型拟合最差的视角方向作为下一最佳视角方 向. Pito<sup>[7]</sup>通过应用 PS(Positional Space)算法确定 当前观测方位下观测不到的未知区域,进而从大量 潜在的观测方位中确定出下一最佳观测方位. Banta 等人<sup>[8]</sup>提出一种组合方法来确定下一最佳观测方 位. Chen 和 Li<sup>[9]</sup>提出一种基于模型的视点规划方 法,该方法对视点设置了一系列约束条件,通过 HG (Hierarchical Genetic)算法和 Christofides 算法来 解决视点规划问题. Li 和 Liu<sup>[10]</sup>提出一种基于信息 熵的视点规划方法,下一最佳观测方位为视觉传感 器获得模型最大信息熵的位置. Scott<sup>[11]</sup>提出一种基 于模型的 3M(Modified Measurability Matrix)算法 来解决视点规划问题.

然而,上述方法在适用范围、遮挡处理或计算效

率等方面存在一些不足. 文献[1]所提方法不适合具 有复杂表面的视觉目标,因此方法的适用范围有限. 文献[2,4,6]所提方法未考虑遮挡因素,因此当视觉 目标出现遮挡现象时,这些方法将无法保证所得结 果的正确性. 文献[5]所提方法针对于特定的设备. 文献[7-8,10]所提方法确定出的观测位置被限定在 一个固定的表面(如球体表面、圆柱体表面等),因此 方法的通用性受到了限制. 文献[3,7,9]所提方法计 算耗时较大,实时性较低. 文献[8-9,11]所提方法需 要预先获取场景的先验知识(如场景中视觉目标的 尺寸等),因此不适用于未知场景.

针对已有 NBV 方法的不足,本文提出一种基 于深度图像利用遮挡信息确定下一最佳观测方位的 方法:该方法首先从某一观测方位获取视觉目标的 一幅深度图像,然后根据已获得的深度图像中的遮 挡信息确定出下一最佳观测方位,从而使摄像机在 所求观测方位下能够最大量地观测到当前观测方位 下视觉目标上的遮挡区域,在方法的设计过程中,我 们主要解决了两个问题:(1)如何建立视觉目标遮 挡区域的模型;(2)如何利用获得的遮挡区域信息 确定下一最佳观测方位.针对第1个问题,我们提出 了最大深度差相邻点的概念,通过利用深度图像中 遮挡边界点及其最大深度差相邻点的三维信息实现 了对遮挡区域的建模,从而获得了遮挡区域外接表 面信息.针对第2个问题,基于已获得的遮挡区域外 接表面信息,我们首先提出了一种基于投影降维思 想确定遮挡区域外接表面最佳小平面集合的方法, 然后提出了一种利用遮挡区域外接表面最佳小平面 集合确定下一最佳观测方位的方法.实验结果验证 了所提方法的可行性和有效性.

本文第2节概述所提方法的基本思想;第3节 详细论述如何依据深度图像中的遮挡信息确定出下 一最佳观测方位;第4节给出实验结果及分析;第5 节总结全文.

# 2 方法概述

#### 2.1 下一最佳观测方位描述

观测方位即摄像机(观测者)观测视觉目标时的 方向和位置.图1所示为一个理想的视觉目标,由于 在当前观测方位下存在遮挡现象,为了获取视觉目 标上遮挡区域(如图1中所示的灰色区域)内的未知 表面信息,需要调整摄像机对其进行下一步观测.其 中,遮挡区域内的未知表面包括两个部分:一是视觉

信息的观测方位.

目标的一部分背面;二是视觉目标上的被遮挡表面. 在所有可能的下一观测方位中,下一最佳观测方位



图 1 下一最佳观测方位描述

为了求解下一最佳观测方位,本文首先提出了 最大深度差相邻点的概念,在此基础上,最终给出了 基于深度图像利用遮挡信息确定下一最佳观测方位 的方法.

## 2.2 最大深度差相邻点、遮挡区域外接表面及其

小平面

对于当前观测方位下存在遮挡现象的视觉目标,由于遮挡区域的信息是未知的,因此需要利用遮挡区域外接表面近似地代表遮挡区域内的未知表面<sup>[5]</sup>.为了结合深度图像中的遮挡信息给出遮挡区域外接表面的定义,本文提出了最大深度差相邻点的概念,其定义为深度图像中任一像素点8个邻域像素点中与其深度差最大的像素点.图2(a)所示为

一个理想视觉目标,其由△ABC (记为面 U)以及矩 形 BCDE(记为面 L)组成.在摄像机当前的观测方 位下,面 U 遮挡了面 L,其对应的深度图像如图 2 (b)所示.在图 2(b)中,由遮挡边界点构成的线段称 为遮挡边界,其为视觉目标深度图像中发生遮挡现 象的多个部分之间的分界线,对应图 2(a)中的线段 AB 和线段 AC;与遮挡边界对应的线段为由各遮挡 边界点的最大深度差相邻点构成的线段,称为邻接 边界,对应图 2(a)中的线段 FB 和线段 FC,在图 2 (a)中,由线段 AB、线段 AC、线段 FB 和线段 FC 所 确定的四边形 ABFC 称为遮挡区域外接表面,其可定 义为由深度图像中的遮挡边界和邻接边界对应的视觉 目标上的直线段(曲线段)所确定的多边形(曲面).

即能够使摄像机获取到最大量遮挡区域内未知表面



图 2 视觉目标遮挡区域外接表面及视觉目标深度图像

将图 2(a)中所示的遮挡区域外接表面单独提 取出来并将其划分为若干足够窄的小平面(如图 3 所示),则每个小平面(图 3 中各虚线框区域)就可以 用一条线段近似地表示,该线段的两个端点分别对 应一个遮挡边界点及其最大深度差相邻点,我们称 这样的线段为遮挡线段.任取遮挡区域外接表面上 的一个小平面 Patch<sub>i</sub>,根据文献[12],Patch<sub>i</sub>的最佳 观测方向应为垂直指向 Patch<sub>i</sub>的方向,最佳观测中 心点(摄像机主光轴正对的场景点)应为遮挡线段的 中点,即在该观测方位下摄像机能够观测到最大面 积的 Patch<sub>i</sub>.



## 2.3 下一最佳观测方位问题分析

基于下一最佳观测方位问题及遮挡区域外接表 面的描述可知,摄像机对遮挡区域内未知表面的观 测可近似等价为其对遮挡区域外接表面的观测,故 能够最大量地观测到遮挡区域外接表面信息的观测

2453

方位即可视为下一最佳观测方位,其中,信息量的大 小通过表面积来衡量.据此可给出下一最佳观测方 位问题的描述如下:记任意下一观测方位为 $x_{nv}$ ,记 摄像机在该观测方位下所能观测到的遮挡区域外接 表面的面积 $S_{nv} = f_s(x_{nv})$ ,遮挡区域外接表面的总 面积为 $S_t$ ,所有可能的下一观测方位组成的集合为  $\{X\}_{nv},则下一最佳观测方位<math>x_{nv}$ 可定义为

 $\boldsymbol{x}_{nbv} = \underset{\boldsymbol{x}_{nv} \in \{X\}_{nv}}{\arg\max} f_s(\boldsymbol{x}_{nv}), \quad \text{s. t. } 0 \leq f_s(\boldsymbol{x}_{nv}) \leq S_t (1)$ 

分析可知,在理想情况下{X}<sub>nv</sub>可视为无穷集, 因而直接求解下一最佳观测方位非常困难.虽然能 够通过设定约束条件限制摄像机的观测方位以获取 有穷的{X}<sub>nv</sub>,从而基于搜索的方式从{X}<sub>nv</sub>中确 定下一最佳观测方位,但摄像机的灵活性会受到限 制.同时,在{X}<sub>nv</sub>较大的情况下,搜索过程耗时也 较大.

针对上述情况,本文提出了下一最佳观测方位 问题的对偶问题,具体描述如下:记 $r_i$ 为当前观测方 位下整个遮挡区域外接表面, $x_v$ 为 $r_i$ 的任一观测方 位,摄像机在 $x_v$ 下所观测到的 $r_i$ 的面积 $S_v = g_s(x_v)$ , 则 $r_i$ 的最佳观测方位 $x_{bv}$ 应满足

 $x_{bv} = \underset{x_v \in \{X\}_v}{\operatorname{arg max}} g_s(x_v)$ , s.t.  $0 \leq g_s(x_v) \leq S_t(2)$ 其中, $\{X\}_v 为 r_t$ 所有可能的观测方位组成的集合. 由于摄像机在  $x_{nbv}$ 及  $x_{bv}$ 下均能够观测到最大面积 的遮挡区域外接表面并且 $\{X\}_v \subset \{X\}_{nv}$ ,因此可将 下一最佳观测方位的求解转化为遮挡区域外接表面 最佳观测方位的求解, $r_t$ 的最佳观测方位  $x_{bv}$ 即为下 一最佳观测方位  $x_{nbv}$ .

由于遮挡区域外接表面 r<sub>i</sub>可能比较复杂,直接 求其最佳观测方位比较困难.所以,基于分治思想, 可利用前述遮挡线段将 r<sub>i</sub>划分为若干足够窄的小平 面 Patch<sub>i</sub>并分别计算各小平面的最佳观测方位信 息,在此基础上,确定一个最佳小平面集合{P}<sub>nbv</sub>, 使 $\{P\}_{nbv}$ 中的各小平面所组成的表面为摄像机在 $r_i$ 的最佳观测方位下观测到的遮挡区域外接表面, 通过对 $\{P\}_{nbv}$ 中各小平面的最佳观测方位信息加权 求和可求出 $r_i$ 的最佳观测方位,即下一最佳观测方 位.记 $\{P\}_i$ 为划分 $r_i$ 所得全部小平面组成的集合,  $\{\Phi\}_i$ 为 $\{P\}_i$ 的所有子集组成的集合, $r_{nv}$ 为摄像机在 下一观测方位 $\mathbf{x}_{nv} \in \{X\}_v$ )所能观测到的遮挡 区域外接表面,对 $r_{nv}$ 进行划分可得小平面集合  $\{P\}_{nv}(\{P\}_{nv} \in \{\Phi\}_i), \end{subscript{abs}}, \end{subscript{bbs}}, \end{s$ 

$$\{P\}_{nbv} = \operatorname*{arg\,max}_{\{P\}_{nv} \in \{\Phi\}_{t}} f_{\Sigma}(\{P\}_{nv}),$$
  
s.t.  $0 \leq f_{\Sigma}(\{P\}_{nv}) \leq S_{t}, \{P\}_{nv} = f_{seg}(x_{nv}),$   
 $\mathbf{x}_{rv} \in \{X\}_{v}$  (3)

由于 $\{P\}_i$ 是有限集(其中的元素个数与深度图 像中遮挡边界点的个数相等),因此 $\{\Phi\}_i$ 也为有限 集.又由于最佳小平面集合 $\{P\}_{nbv} \in \{\Phi\}_i$ ,因此通过 设计适当的搜索策略即可从 $\{\Phi\}_i$ 中确定出 $\{P\}_{nbv}$ . 在此基础上,分别计算 $\{P\}_{nbv}$ 中各小平面的最佳观 测方位,通过对其加权求和即可计算出下一最佳观 测方位.

### 2.4 方法总体思想

基于上述分析,本文提出一种基于深度图像利 用遮挡信息确定下一最佳观测方位的方法.方法总 体思想如下.首先,获取视觉目标的一幅深度图像, 同时记录摄像机当前的内外参数.其次,计算深度 图像中的遮挡信息并提取最大深度差相邻点.然 后,利用上述得到的信息计算遮挡区域外接表面 上各小平面的最佳观测方向及观测中心点.最后, 计算出最佳小平面集合,基于得到的最佳小平面集 合计算出下一最佳观测方位.所提方法流程如图 4 所示.



图 4 下一最佳观测方位确定流程

由于目前已很容易获得视觉目标的深度图像 (如用 Kinect),且可以利用已有方法检测到深度图 像中的遮挡信息<sup>[13-16]</sup>,故本文不再对深度图像获取 及遮挡检测方法进行探讨,而是重点论述如何利用 遮挡信息确定下一最佳观测方位.

# 3 利用遮挡信息确定下一最佳观测方位

## 3.1 提取最大深度差相邻点

获得深度图像及其遮挡信息后,下一步工作是

提取最大深度差相邻点.依据前述所提最大深度差 相邻点的概念,记深度图像 I 中任一像素点 p(i,j)的深度值为 d(i,j),则该像素点的最大深度差值  $f_{MDD}(i,j)$ 的计算公式为

$$f_{MDD}(i,j) = \max_{x,y}(d(x,y) - d(i,j)),$$

s.t.  $i - 1 \le x \le i + 1, j - 1 \le y \le j + 1$  (4)

提取一幅深度图像的最大深度差相邻点的具体 方法为:遍历整幅深度图像中的各点,按式(4)计算 各点的最大深度差值,记录下各最大深度差值对应 的邻域点的像素坐标.遍历完整幅图像后,即可得到 每个像素点的最大深度差相邻点的像素坐标.

## 3.2 计算小平面的最佳观测方向及观测中心点

得到最大深度差相邻点的像素坐标后,还需要 获取深度图像中每个像素点的三维坐标,对于已获 取的单幅深度图像,可以利用拍摄该幅深度图像时 摄像机的内外参数信息,依据投影变换原理,对该幅 深度图像进行反投影变换,从而重建出深度图像中 每个像素点在世界坐标系中的三维坐标.得到三维 坐标信息之后,即可利用深度图像中的遮挡边界点 及其对应的最大深度差相邻点的三维坐标以及摄像 机的当前观测方位计算遮挡区域外接表面上各小平 面对应的最佳观测方向及观测中心点,设摄像机当 前的观测方向为V<sub>camera</sub>,在当前观测方位下,任取一 个遮挡边界点 A,坐标为( $x_A$ , $y_A$ , $z_A$ ),其对应的最 大深度差相邻点 F 的坐标为( $x_F, y_F, z_F$ ),记由点 A和点 F 确定的遮挡线段为 l<sub>AF</sub>,由 l<sub>AF</sub>表示的小平面 为 Patch AF, 根据 2.2 节的描述, 小平面 Patch AF 的 最佳观测方向 $V_{thv}$ 应垂直于 $l_{AF}$ 且指向其观测中心 点 P, P 应为  $l_{AF}$  的中点,具体计算方法如下.

(1) 计算最佳观测中心点 *P*. 设其坐标为(*x<sub>P</sub>*, *y<sub>P</sub>*,*z<sub>P</sub>*),计算公式为

$$(x_P, y_P, z_P) = \frac{1}{2}(x_A + x_F, y_A + y_F, z_A + z_F)$$
 (5)

(2) 计算最佳观测方向  $V_{pbv}$ .由于  $V_{pbv}$ 垂直于代表 Patch<sub>AF</sub> 的遮挡线段  $l_{AF}$ ,因此可首先求得 Patch<sub>AF</sub>上的两个方向向量,然后通过计算所得向量 的叉积得到  $V_{pbv}$ .目前 Patch<sub>AF</sub>上可利用的方向向量 仅有  $V_{AF}$ (点 A 和点 F 确定的向量),Patch<sub>AF</sub>上另 一个方向向量  $V_x$ 可通过计算  $V_{camera}$ 与  $V_{AF}$ 的叉积得到.在此基础上,即可利用  $V_x$ 与  $V_{AF}$ 进一步求得  $V_{pbv}$ .分析发现,摄像机与遮挡区域空间位置关系有 两种情形,各情形对应的  $V_{pbv}$ 求解方法有所不同.其中,情形 1 为摄像机主光轴穿过遮挡区域,与遮挡区域,与遮挡区域;情形 2 为摄像机主光轴未穿过遮挡区域,与

遮挡区域不相交.为了区别这两种情形,图 5(a)和 图 5(b)分别给出了 *Patch*<sub>AF</sub>的最佳观测方向及观测 中心点的示意图.



(b)摄像机主光轴与遮挡区域不相交

图 5 不同情形下小平面最佳观测方向和观测中心点示意

由图 5(a)分析可知,当摄像机主光轴与遮挡区 域相交时, $V_{AF}$ 和 $V_x$ 的叉积 $V_{temp}$ 垂直于线段  $l_{AF}$ 且指 向遮挡区域,故将其直接作为  $Patch_{AF}$ 的最佳观测 方向 $V_{pbv}$ .而由图 5(b)分析可知,当摄像机主光轴 与遮挡区域不相交时,由于三角形 ABC 的遮挡,摄 像机在观测方位  $V_{temp}$ 下将无法观测到遮挡区域,此 时,合理的 $V_{pbv}$ 应该为 $-V_{temp}$ .综合考虑上述两种情 形, $V_{pbv}$ 的确定方法如下.

首先,计算 $V_x$ 和 $V_{temp}$ ,并将 $V_{temp}$ 作为临时观测 方向. $V_x$ 和 $V_{temp}$ 的计算公式分别为

$$\boldsymbol{V}_{x} = \boldsymbol{V}_{\text{camera}} \times \boldsymbol{V}_{AF} \tag{6}$$

和

$$\boldsymbol{V}_{\text{temp}} = \boldsymbol{V}_{AF} \times \boldsymbol{V}_{x} \tag{7}$$

其次,将 $V_{temp}$ 单位化,得到其对应的单位向量  $v_{temp}$ ,将式(5)中计算出的观测中心点P作为 $v_{temp}$ 的 终点,计算 $v_{temp}$ 的起点Q的坐标( $x_Q, y_Q, z_Q$ ).此处 使用 $v_{temp}$ 而不直接使用 $V_{temp}$ 来确定Q点的坐标是 为了保证后续对Q点进行投影变换时,其所对应的 投影点位于深度图像的平面内.

再次,利用已得到的摄像机内外参数,对Q点

进行投影变换,可以获得 Q 点在深度图像平面内的 投影点  $Q_{\text{project}}$ .图 6(a)和图 6(b)分别给出了图 5(a)和图 5(b)所示两种情形下  $Q_{\text{project}}$ 相对于遮挡边界点 A 及其最大深度差相邻点 F 的投影点  $A_{\text{project}}$ 和  $F_{\text{project}}$ 的位置关系.其中,图 6(a)所示为摄像机主光 轴与遮挡区域相交的情形,而图 6(b)所示为摄像机 主光轴与遮挡区域不相交的情形.



(a)摄像机主光轴与遮挡区域相交(b)摄像机主光轴与遮挡区域不相交图 6 不同情形下Q点在深度图像平面的投影点位置示意

记  $A_{\text{project}}$  和  $Q_{\text{project}}$  所形成的向量与  $A_{\text{project}}$  和  $F_{\text{project}}$ 所形成的向量的夹角为 $\theta$ . 由图 6(a)分析可 知,当摄像机主光轴与遮挡区域相交时, $\theta$ 角小于 90°,而由图 6(b)分析可知,当摄像机主光轴与遮挡 区域不相交时, $\theta$ 角大于 90°. 因此,可以根据 $\theta$ 角的 大小来确定  $V_{\text{pbv}}$ 的方向. 记  $A_{\text{project}}$ 和  $Q_{\text{project}}$ 形成的向 量为  $V_{A^*Q^*}$ ,记  $A_{\text{project}}$ 和  $F_{\text{project}}$ 形成的向量为 $V_{A^*F^*}$ , 则 $\theta$ 角的计算公式为

$$\theta = \arccos\left(\frac{\boldsymbol{V}_{A^{*}Q^{*}} \cdot \boldsymbol{V}_{A^{*}F^{*}}}{\|\boldsymbol{V}_{A^{*}Q^{*}}\| \times \|\boldsymbol{V}_{A^{*}F^{*}}\|}\right)$$
(8)

最后,根据计算出的 $\theta$ 角的值确定出小平面 Patch<sub>AF</sub>的最佳观测方向 $V_{phv}$ ,方法为

$$\mathbf{V}_{pbv} = \begin{cases} \mathbf{V}_{temp}, & \theta \leq 90^{\circ} \\ -\mathbf{V}_{temp}, & \theta > 90^{\circ} \end{cases}$$
(9)

3.3 确定下一最佳观测方位

3.3.1 遮挡区域外接表面最佳小平面集合的确定 思路

计算出遮挡区域外接表面中各小平面的最佳观

测方向及观测中心点后,下一步工作是利用这些信 息确定出一个最佳小平面集合,该小平面集合所对 应的观测方位即为所求的下一最佳观测方位,依据 最佳小平面集合的定义(详见 2.3 节),记r,为当前 观测方位下整个遮挡区域外接表面, $\{P\}_t$ 为划分 $r_t$ 所得全部小平面组成的集合, $\{\phi\}$ ,为 $\{P\}$ ,的所有子 集组成的集合,最佳小平面集合为 $\{P\}_{nbw}$ ,由于  $\{P\}_{nbv} \in \{\Phi\}, \mathbb{L}\{\Phi\},$ 为有限集,因此可通过遍历  $\{\Phi\}_{t}$ 确定出最佳小平面集合.事实上,我们并不需要 对整个 $\{\phi\}$ ,进行处理,而只需要关注其中一部分有 效的小平面集合,其定义如下:记任意下一观测方位 为 x<sub>nn</sub>,摄像机在 x<sub>nn</sub>下所观测到的遮挡区域外接表 面为r<sub>nn</sub>,则划分 r<sub>nn</sub>所得小平面集合称为有效的小 平面集合.虽然通过仅遍历全部有效的小平面集合 可缩小查找范围,从而减少了一部分计算量,但若是 在三维空间中确定最佳小平面集合还需要对大量三 维信息进行处理,计算过程仍比较复杂,故在此基础 上,本文提出一种基于投影降维思想的搜索策略来 确定最佳小平面集合,具体描述如下.

首先将 $\{P\}_i$ 中各小平面的最佳观测方向投影到 一个指定坐标系的 xoy 平面,要求该指定坐标系的 z 轴正方向与摄像机当前的观测方向相同.这种投 影方式具有两个优点:(1)可以使得所有投影向量 不重叠,也就是说,每个小平面的最佳观测方向与其 在该平面内的投影向量是一一对应的;(2)记 $\{P\}_{valid}$ 为任一有效的小平面集合, $\{P\}_{valid}$ 中各小平面最佳 观测方向在 xoy 平面的投影向量组成集合 $\{V\}_{pro}$ , 则 $\{V\}_{pro}$ 中任意两个投影向量的夹角应小于等于一 个固定角度  $\gamma$ ,称 $\{V\}_{pro}$ 为一个有效的投影向量集 合.如图 7 所示,当摄像机在某个下一观测方位对遮 挡区域外接表面进行观测时,可将其可视空间(视 野)视为投影平面内的一个梯形区域,若该梯形区域 两个腰的夹角为 $\beta$ ,则 $\gamma$ 应为 $\pi - \beta$ . 一般情况下, $\beta$ 取 为摄像机的水平视场角.



图 7 摄像机观测遮挡区域外接表面时的可视空间及小平面最佳观测方向投影示意

→ 投影向量 投影向量平移

基于这种投影方式可得全部投影向量的分布情况如图 8(a)所示,其中的梯形框区域为摄像机观测 遮挡区域外接表面时可视空间对应的投影区域.若 将所有投影向量的起点都移动到指定坐标系 *xoy* 平面的原点(如图 8(b)所示),记各投影向量与 *x* 轴 正方向的夹角为 $\alpha, \alpha \in [0^\circ, 360^\circ), 任取 \phi \in [0^\circ, 360^\circ),$ 则夹角  $\alpha$  的值在  $[\phi, \phi + \gamma]$ 范围(如图 8(b)所示扇 形区域)内的所有投影向量所组成的集合即为一 个有效的投影向量集合.因此,每给定一个角度  $\phi$ 即可计算出一个有效的投影向量集合,通过该集合



即可确定一个有效的小平面集合.在此基础上,遍历 角度 \$ 的所有可能取值,每获得一个有效的投影向 量集合,计算其对应小平面集合中各小平面面积之 和,遍历完成后,面积值最大的小平面集合即为 最佳小平面集合.此处需要说明的是,由于本文基 于积分思想将遮挡区域外接表面划分为若干足够 窄小平面,每个小平面均可用其对应的遮挡线段 表示.因此,各小平面的面积值即为其对应遮挡线 段的长度值.下面给出最佳小平面集合的具体求解 过程.



(b) 平移后投影向量的分布情况及遍历方式

图 8 投影向量的分布情况及遍历示意

3.3.2 确定最佳小平面集合

3.3.2.1 计算投影向量及其夹角 α

为了确定最佳小平面集合,基于投影降维思想, 首先将计算出的各小平面的最佳观测方向投影到一 个指定坐标系的 xoy 平面(要求该坐标系的 z 轴正 方向为摄像机当前的观测方向),然后,计算所有小 平面各自的最佳观测方向在该坐标系 xoy 平面内 的投影与x 轴正方向的夹角 $\alpha$ ,从而得到投影向量 的分布情况.任取一个小平面的最佳观测方向 $V_{pbv}$ , 其在指定坐标系 xoy 平面的投影  $V_{project}$  的位置如 图 9 所示,夹角 $\alpha$ 为x 轴正方向沿箭头所示方向与  $V_{project}$ 的夹角,其中 $\alpha \in [0^\circ, 360^\circ).$ 



夹角 α 的具体计算方法如下.

首先,确定指定坐标系的 x 轴、y 轴和 z 轴正方向 的单位向量.任取一个遮挡边界点 A 及其最大深度差 相邻点 F 的三维坐标,二者可以确定一个向量  $V_{AF}$ ,记 摄像机当前观测方向为  $V_{camera}$ ,令  $V_x = V_{camera} \times V_{AF}$ , 将 $V_x$ 单位化可得到x轴正方向的单位向量 $e_x$ ,将  $V_{camera}$ 单位化可得到z轴正方向的单位向量 $e_z$ ,最后,y轴正方向单位向量则为 $e_y = e_z \times e_x$ .

然后,计算  $V_{pbv}$  在指定坐标系 xoy 平面的投影 向量. 在世界坐标系下,记  $e_x$ ,  $e_y$ 和  $e_z$ 的坐标分别为  $(x_{e_x}, y_{e_x}, z_{e_x})$ 、 $(x_{e_y}, y_{e_y}, z_{e_y})$ 和 $(x_{e_z}, y_{e_z}, z_{e_z})$ , $V_{pbv}$ 坐 标为 $(x_c, y_c, z_c)$ ,则  $V_{pbv}$ 在指定坐标系的坐标  $(x'_c, y'_c, z'_c)$ 的计算公式为

$$\begin{bmatrix} x_c' \\ y_c' \\ z_c' \end{bmatrix} = \begin{bmatrix} x_{e_x} & x_{e_y} & x_{e_z} \\ y_{e_x} & y_{e_y} & y_{e_z} \\ z_{e_x} & z_{e_y} & z_{e_z} \end{bmatrix}^{-1} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$
(10)

记 V<sub>pbv</sub> 在指定坐标系 xoy 平面的投影向量为 V<sub>project</sub>,坐标为(x<sub>project</sub>,y<sub>project</sub>),则

$$(x_{\text{project}}, y_{\text{project}}) = (x'_{c}, y'_{c})$$
(11)

最后,计算 $V_{\text{project}}$ 与x轴正方向的夹角 $\alpha$ .其计算公式为

$$\alpha = \begin{cases} \arccos\left(\frac{x_{\text{project}}}{\sqrt{x_{\text{project}}^{2} + y_{\text{project}}^{2}}}\right), & y_{\text{project}} \ge 0\\ 360^{\circ} - \arccos\left(\frac{x_{\text{project}}}{\sqrt{x_{\text{project}}^{2} + y_{\text{project}}^{2}}}\right), & y_{\text{project}} < 0 \end{cases}$$
(12)

## 3.3.2.2 计算最佳小平面集合

得到所有小平面最佳观测方向投影向量的夹角后,即可计算最佳小平面集合. 令 $\phi \in [0^\circ, 360^\circ)$ ,记  $\phi_i \end{pmatrix} \phi$ 的任一取值,固定角度 $\gamma$ ,夹角 $\alpha$  在 $[\phi_i, \phi_i + \gamma]$ 范围内的投影向量组成的集合为 $\{V\}_{\text{pro}}, \{V\}_{\text{pro}}$ 对应 的小平面集合为 $\{P\}_{valid}$ ,代表 $\{P\}_{valid}$ 中各小平面的 遮挡线段组成的集合为 $\{L\}_{\phi_i}$ ,则 $\{P\}_{valid}$ 对应的遮挡 区域外接表面面积  $S_{\phi_i}$ (即 $\{P\}_{valid}$ 中各小平面的面积 之和)为

$$S_{\phi_{i}} = \sum_{l_{i} \in \langle L \rangle_{\phi_{i}}} area(l_{i})$$
(13)

其中,  $area(l_i)$ 为集合 { $L_i$ }, 中第 i 条遮挡线段  $l_i$ 所 代表的小平面的面积, 其值为  $l_i$ 的长度值. 在此基础 上, 记最佳小平面集合 {P}<sub>nbv</sub>对应的遮挡区域外接 表面面积为  $S_{max}$ , 固定迭代步长为 step, 其  $S_{max}$ 计算 公式为

 $S_{\max} = \underset{\phi_i \in [0^{\circ}, 360^{\circ})}{\operatorname{maximize}} S_{\phi_i}, \quad \text{s. t. } \phi_{i+1} = \phi_i + step (14)$ 

求得  $S_{max}$ 后,记录下  $S_{max}$ 对应的  $\phi_i$ 值,记为  $\phi_{begin}, \Leftrightarrow \phi_{end} = \phi_{begin} + \gamma,则对应夹角 \alpha 在[\phi_{begin}, \phi_{end}]$ 范围内的投影向量对应的小平面组成的集合即为最 佳小平面集合.

3.3.3 确定下一最佳观测方位

确定出最佳小平面集合后,即可在此基础上进 一步计算出下一最佳观测方位.记下一最佳观测方 位 $V_{nbv}$ 及观测中心点 $P_{view}$ 的坐标分别为 $(x_{nbv}, y_{nbv},$  $z_{nbv}$ )和 $(x_{view}, y_{view}, z_{view})$ ,摄像机当前的观测方向  $V_{camera}$ 的坐标为 $(x_c, y_c, z_c)$ ,求得的最佳小平面集合 为 $\{P\}_{nbv}$ ,其对应遮挡区域外表面面积为 $S_{max}$ .为了 兼顾遮挡现象不显著或不存在遮挡现象的情况(即  $S_{max}$ 比较小的情况),我们设定了一个遮挡区域外接 表面面积阈值 $S_{thres}$ (可根据具体应用的实际情况进 行调整),根据 $S_{max}$ 与 $S_{thres}$ 的大小关系,分两种情况 进行讨论.

(1)当 $0 \le S_{max} < S_{thres}$ 时,即当前观测方位下遮 挡现象不显著或不存在遮挡现象时,我们认为在当 前观测方位下的场景已经得到了充分的观测,如果 将当前观测方位获取到的场景表面视为视觉目标的 正面,要想尽可能多地获取场景的未知信息,显然需 要从视觉目标的背面进行进一步观测,即将当前观 测方向的反方向作为下一最佳观测方向.同时,为了 尽可能多地观测到视觉目标的背面,将当前观测方 位下视觉目标的中心点作为观测点,从而确定出该 情况下的下一最佳观测方位.记摄像机在当前观测 方位下获得的视觉目标点总数为  $N_{model}$ ,观测到的 任意视觉目标点  $M_i$ 的坐标为 $(x_{M_i}, y_{M_i}, z_{M_i}),则 <math>V_{nbv}$ 的计算公式为

 $(x_{nbv}, y_{nbv}, z_{nbv}) = (-x_c, -y_c, -z_c)$  (15)  $P_{view}$ 的计算公式为

$$(x_{\text{view}}, y_{\text{view}}, z_{\text{view}}) = \frac{1}{N_{\text{model}}} \left( \sum x_{M_i}, \sum y_{M_i}, \sum z_{M_i} \right)$$
(16)

(2) 当  $S_{\text{max}} \ge S_{\text{thres}}$ 时,即当前观测方位下遮挡现 象较显著时,基于求得的最佳小平面集合确定最佳 观测方向.记{P}<sub>nbv</sub>中任意小平面的最佳观测方向  $V_i$ 和观测中心点 $P_i$ 的坐标分别为( $x_i, y_i, z_i$ )和 ( $x_{P_i}, y_{P_i}, z_{P_i}$ ), $V_i$ 的投影向量夹角 $\alpha \in [\phi_{\text{begin}}, \phi_{\text{end}}]$ , 则 $V_{\text{nbv}}$ 的计算公式为

$$(x_{nbv}, y_{nbv}, z_{nbv}) = \left(\sum \omega_{i} x_{i}, \sum \omega_{i} y_{i}, \sum \omega_{i} z_{i}\right) (17)$$

$$P_{view} 的计算公式为$$

$$(x_{view}, y_{view}, z_{view}) = \left(\sum \omega_{i} x_{P_{i}}, \sum \omega_{i} y_{P_{i}}, \sum \omega_{i} z_{P_{i}}\right) (18)$$

式(17)和(18)中, $\omega_i$ 为 $V_i$ 、 $P_i$ 对应的权值,记 $V_i$ 对应的遮挡线段为 $l'_i$ ,则 $\omega_i$ 的计算公式为

$$\omega_i = \frac{area\left(l'_i\right)}{S_{\max}} \tag{19}$$

最后,根据计算出的下一最佳观测方向和观测 中心点计算摄像机的观测位置 P<sub>camera</sub>.图 10 给出了 下一最佳观测方向和观测中心点与摄像机观测位置 之间的关系.



图 10 下一最佳观测方向和观测中心点与 摄像机观测位置之间的关系

记  $P_{\text{camera}}$ 的坐标为 $(x_{P_{\text{camera}}}, y_{P_{\text{camera}}}, z_{P_{\text{camera}}})$ ,摄像 机相对于观测中心点的观测距离为  $d_{\text{camera}}$ .设由点  $P_{\text{camera}}$ 和点  $P_{\text{view}}$ 确定的向量为  $\mathbf{V}'_{\text{nbv}}$ ,坐标为 $(x_{\text{view}} - x_{P_{\text{camera}}}, y_{\text{view}} - y_{P_{\text{camera}}}, z_{\text{view}} - z_{P_{\text{camera}}})$ .由图 10 可知,  $\|\mathbf{V}'_{\text{nbv}}\| = d_{\text{camera}}, \mathbf{V}'_{\text{nbv}} \in \mathbf{V}_{\text{nbv}}$ 的平行向量,且方向相同, 记  $\mathbf{V}'_{\text{nbv}}$ 的坐标为 $(x'_{\text{nbv}}, y'_{\text{nbv}}, z'_{\text{nbv}})$ ,由平行向量的性质 可得

$$(x'_{\text{nbv}}, y'_{\text{nbv}}, z'_{\text{nbv}}) = \frac{\|V'_{\text{nbv}}\|}{\|V_{\text{nbv}}\|} (x_{\text{nbv}}, y_{\text{nbv}}, z_{\text{nbv}})$$
$$= \frac{d_{\text{camera}}}{\|V_{\text{nbv}}\|} (x_{\text{nbv}}, y_{\text{nbv}}, z_{\text{nbv}}) (20)$$

则 P<sub>camera</sub>的计算公式为

$$(x_{P_{camera}}, y_{P_{camera}}, z_{P_{camera}}) = (x_{view}, y_{view}, z_{view}) - \frac{d_{camera}}{\|V_{nbv}\|} (x_{nbv}, y_{nbv}, z_{nbv}) (21)$$
至此,已确定出下一最佳观测方位  $x_{nbv} = (V_{nbv}, P_{camera}).$ 

#### 3.4 下一最佳观测方位确定算法

算法1. 下一最佳观测方位确定算法.
输入:深度图像、遮挡边界和摄像机内外参数
输出:下一最佳观测方向和摄像机位置的三维坐标
1. /\*提取最大深度差相邻点\*/
For 深度图像中每个像素点
根据式(4)计算最大深度差并记录最大深度差
相邻点的坐标

End For

- 计算深度图像中每个像素点的三维坐标并记录三 维坐标的总数.
- 根据式(5)~(9)计算各小平面的最佳观测方向及 观测中心点.
- 根据式(10)~(12)计算各小平面最佳观测方向的投影向量及其夹角 α.
- 5. / \* 确定最佳小平面集合 \* /
  - For 每个角度值 *φ<sub>i</sub>*,0°≤*φ<sub>i</sub>*<360°,*φ<sub>i+1</sub>*=*φ<sub>i</sub>*+step (step 为固定步长)

获取夹角  $\alpha$  在[ $\phi_i$ ,  $\phi_i$  +  $\gamma$ ]范围内的投影向量对应 的小平面集合 {*P*}<sub>valid</sub>;

根据式(13)计算 $\{P\}_{valid}$ 对应遮挡区域外接表面的面积  $S_{\phi_i}$ ;

End For

- 根据式(14)计算最大面积  $S_{max}$ 并记录其对应的最佳 小平面集合 $\{P\}_{nbv}$ .
- 6. / \* 确定下一最佳观测方位 \* /
  - If S<sub>max</sub>小于面积阈值 S<sub>thres</sub>

根据式(15)、(16)计算下一最佳观测方向及观测 中心点

Else If Smax大于或等于面积阈值 Sthres

获取{P}<sub>nbv</sub>中各小平面的最佳观测方向及观测中 心点;

根据式(17)~(19)计算下一最佳观测方向及观测 中心点;

End If

依据式(20)、(21)计算摄像机的位置.

 输出计算出的下一最佳观测方向和摄像机位置的 三维坐标.

# 4 实验及分析

#### 4.1 实验环境

为了验证本文所提方法的效果,分别进行了仿 真实验和真实实验,实验硬件环境为 CPU Pentium (R) Dual-Core 2.93 GHz、内存 2.0 GB.确定下一最 佳观测方位的程序采用C++编程实现.在仿真实验 阶段,我们用 http://range.informatik.uni-stuttgart. de/htdocs/html 上 Stuttgart Range Image Database 中提供的三维实物模型进行了实验.实验过程中,通 过 OpenGL 模拟摄像机观测三维实物模型并获取 其深度图像,设置投影矩阵的参数为(60,1,200, 600),视窗大小设定为400×400.在真实实验阶段, 我们应用 Kinect 作为深度图像获取设备,选取了现 实生活中若干实物作为视觉目标进行了实验,深度 图像的分辨率为640×480.

### 4.2 实验结果及分析

4.2.1 所提方法的实验结果及分析

我们应用所提方法对不同的场景进行了实验. 仿真实验过程中设定摄像机相对于观测中心点的观 测距离为 250 mm,固定角度 γ 为 120°(摄像机的水 平视场角为 60°), ∮ 的遍历步长为 1°,面积阈值为 10 mm<sup>2</sup>,计算遮挡边界的方法采用文献[16]中所提 的遮挡检测方法.真实实验过程中设定摄像机相对 于观测中心点的观测距离为 1200 mm,固定角度 γ 为 123°(Kinect 的水平视场角为 57°),其余设置与 仿真实验一致.

图 11 给出了部分实验结果,从上到下 9 种不同 复杂程度的视觉目标名称分别为 Bunny、Duck、 Mole、Rocker、Dragon、Banana、Printer、Kettle 和 Panda.其中,前 6 个为仿真实验对应的视觉目标, 后 3 个为真实实验对应的视觉目标.图 11 所示第 1 列为视觉目标名称,第 2 列为摄像机在当前观测方 位下拍摄到的视觉目标的深度图像,第 3 列为视觉 目标深度图像中的遮挡边界点及其最大深度差相邻 点,第 4 列为计算出的各小平面的最佳观测方向(数 量多且长度均等的向量)及最终的下一最佳观测方 向(唯一且长度最长的向量),第 5 列为所求下一最 佳观测方位对应的遮挡线段(片状分布的线段)及观 测中心点(遮挡线段上成带状分布的点),第 6 列为摄 像机移动到下一最佳观测方位时获取到的深度图像.

由图 11 可以看出,对于当前观测方位下存在遮 挡区域且遮挡区域外接表面面积大于给定面积阈值 的视觉目标(如 Bunny、Duck、Mole、Rocker、Dragon、 Printer、Kettle 和 Panda),所提方法在获得视觉目 标深度图像中的遮挡边界点及其最大深度差相邻点 的基础上,能够根据上述信息计算出合理的小平面 最佳观测方向及观测中心点,并最终确定出下一最 佳观测方位,使得在所求观测方位下能够尽可能 多地观测到当前观测方位下的遮挡区域.对于当前

Bunny					
Duck	3				
Mole					
Rocker	$\bigcirc$			Î	
Dragon	ч?	N.			No.
Banana					
Printer					
Kettle					
Panda	**	**			X
(a) 视觉目标名称	(b)当前观测方位下 视觉目标的深度图像	(c)遮挡边界点及其 最大深度差相邻点	(d) 各小平面的最佳观测 方向和下一最佳观测方向	(e)下一最佳观测方位 对应的遮挡线段及 观测中心点	(f)下一最佳观测方位 下的视觉目标深度图像

图 11 下一最佳观测方位实验结果

观测方位下不存在遮挡区域或遮挡区域外接表面面 积小于给定面积阈值的视觉目标(如 Banana),将当 前观测方向的反方向作为下一最佳观测方向,将当 前观测方位下视觉目标的中心点作为观测点,从而 得到了下一最佳观测方位.分成上述两种情况计算, 可以保证观测的充分性,从而为后续的三维重建等 领域的实际应用提供高精度的视觉目标点集.同时, 由图 11 还可以看出,无论是仿真实验还是真实实 验,所提方法计算出的下一最佳观测方位均与人类 的视觉习惯相符,说明了所提方法的可行性和有效 性,此外,需指出的是,在图 11 的真实实验中,由于 Kinect 获取的深度图像带有噪声,使得视觉目标 Printer、Kettle 和 Panda 的深度图像中的部分深度 值产生偏差,进而导致深度图像中的各视觉目标检 测出的遮挡边界产生偏差,从而在一定程度上会影 响到下一最佳观测方位确定的结果. 但对本文方法 而言,遮挡检测过程为预处理过程,遮挡检测结果只 是本文方法的输入数据,所以,只要在图像采集及遮 挡检测过程中对噪声加以有效处理,便可进一步改 善本文方法在噪声干扰下的效果.

考虑到目前不存在 NBV 方面的标准对照图 (Ground Truth),因此很难通过与标准结果比较的 方式来评价方法的优劣.但为了进一步评估所提方 法的合理性,我们对上述视觉目标对应的实验结果 进行了量化分析,分析结果如表 1 所示.其中, $N_{cv}$ 为 当前观测方位下获取到的视觉目标表面点的三维坐 标集合  $Set_{cv}$ 中元素的个数, $N_{nbv}$ 为下一最佳观测方 位下获取到的视觉目标表面点的三维坐标集合  $Set_{nbv}$ 中元素的个数, $N_{overlap}$ 为  $Set_{cv} \cap Set_{nbv}$ 中元素的 个数,即为上述两个集合中重合点的个数, $N_{new} =$  $N_{nbv} - N_{overlap}$ ,即所求下一最佳观测方位下新增点的 个数, $R_{overlap}$ ,即所求下一最佳观测方位下新增点的 个数, $R_{overlap}$ 为重合率,计算公式为  $N_{overlap}/N_{nbv}$ , $R_{new}$ 为新增率,计算公式为( $N_{nbv} - N_{overlap}$ )/ $N_{nbv}$ ,即 1- $R_{overlap}$ .评估所求下一最佳观测方位效果时,首先关 注的是新增点数  $N_{new}$ ,其次才关注新增率  $R_{new}$ .

表 1 下一最佳观测方位实验结果量化评估

视觉目 标名称	$N_{cv}$	${N}_{ m nbv}$	$N_{ m overlap}$	$N_{ m new}$	$R_{ m overlap}/\%$	$R_{ m new}/\%$
Bunny	13813	23485	2229	21256	9.49	90.51
Duck	20594	18812	4593	14219	24.42	75.58
Mole	14432	12268	1396	10872	11.38	88.62
Rocker	3728	9639	360	9279	3.73	96.27
Dragon	4737	11666	405	11261	3.47	96.53
Banana	7996	10733	3	10730	0.03	99.97
Printer	12460	34133	5850	28283	17.13	82.87
Kettle	12014	17685	8154	9531	46.11	53.89
Panda	3748	6196	1777	4419	28.68	71.32

由表1可以看出,在下一最佳观测方位下,获取 到的场景点具有较多的新增点数和较高的新增率, 说明在所求观测方位下,摄像机的确能够最大量地 观测到当前观测方位下观测不到的区域.对于当前 观测方位下存在遮挡区域且遮挡区域外接表面面积 大于给定面积阈值的视觉目标(如 Bunny、Duck、 Mole、Rocker、Dragon、Printer、Kettle 和 Panda),其 对应的重合率略高,也就是说,在确定出的下一最佳 观测方位下观测这些视觉目标时,获取到的视觉目 标点中包含了少部分已经观测过的点.这是由于当 前观测方位下的遮挡区域与非遮挡区域相接,因此, 在下一最佳观测方位下获取到的视觉目标点集中会 包含一部分已经观测过的视觉目标点.同时,由于每 个场景在当前观测方位下所对应的遮挡区域的面积 和形状存在差异,因此导致每个场景对应的重合率 也会有所不同.其中,Kettle的重合率相对较高(即 新增率相对较低),其主要是由于 Kettle 发生遮挡 的区域在壶柄部位,遮挡区域较小.同时,Kettle对 应的下一最佳观测方位相对于当前的观测方位的变 化也较小(与人类观测壶柄处遮挡的行为相符),从 而导致摄像机在下一最佳观测方位观测到的新增视 觉目标表面点相对较少.对于当前观测方位下不存 在遮挡区域或遮挡区域外接表面面积小于给定面积 阈值的视觉目标(如 Banana),由于直接将当前观测 方向的反方向作为下一最佳观测方向,即从视觉目 标的背面进行观测,故其对应的新增点数较多,新增 率较高,重合率较低.

此外,表2中给出了在我们的实验环境下所提 方法针对不同视觉目标确定下一最佳观测方位的时 间消耗.每个视觉目标对应的时间消耗为10次实验 结果的平均值.由表2可以算出,所提方法的平均时 间消耗为21.57ms,其中,仿真实验的平均时间消耗 为21.31ms,真实实验的平均时间消耗为22.08ms.

表 2 本文方法确定下一最佳观测方位的时间消耗

视觉目标名称	视觉目标模型	时间消耗/ms
Bunny	10	25.49
Duck	2	21.72
Mole		21.91
Rocker	-	21.46
Dragon		21.27
Banana	~	16.02
Printer		21.09
Kettle	1	22.83
Panda	Å.	22.31

4.2.2 与已有方法的实验比较及分析

为了更好地评估本文所提方法的效果,我们将

其与基于深度图像、对设备无依赖性且考虑了遮挡 信息的文献[8,10]中的方法进行了实验比较.其中, 各方法在相同的观测方位下计算下一最佳观测方 位.图12给出了不同方法在所求下一最佳观测方位 下获得的视觉目标深度图像.其中,第1行为当前观 测方位下视觉目标的深度图像,第2行为文献[8]方 法计算出的下一最佳观测方位下视觉目标的深度 图像,第3行为文献[10]方法计算出的下一最佳观测方位下视觉目标的深度图像,第4行为本文方法 计算出的下一最佳观测方位下视觉目标的深度 图像.同时,表3给出了不同方法在所求下一最佳观 测方位下对应的视觉目标表面点个数 N<sub>nbv</sub>、重合点 个数 N<sub>overlap</sub>、新增点个数 N<sub>new</sub>、重合率 R<sub>overlap</sub>和新增 率 R<sub>new</sub>.



图 12 不同方法在所求下一最佳观测方位下获得的视觉目标深度图像

表 3 不同方法的下一最佳观测方位实验结果量化评估

视觉目	文献[8]方法				文献[10]方法				本文方法						
标名称	$N_{ m nbv}$	$N_{\rm overlap}$	$N_{\rm new}$	$R_{ m overlap}/\%$	$R_{ m new}/\%$	$N_{ m nbv}$	$N_{\mathrm{overlap}}$	$N_{\rm new}$	$R_{ m overlap}/\%$	$R_{ m new}/\%$	$N_{ m nbv}$	$N_{ m overlap}$	$N_{\rm new}$	$R_{ m overlap}/\%$	$R_{ m new}/\%$
Bunny	18839	1	18838	0.01	99.99	17601	5053	12548	28.71	71.29	23485	2229	21256	9.49	90.51
Duck	21580	25	21555	0.12	99.88	21838	9769	12069	44.73	55.27	18812	4593	14219	24.42	75.58
Mole	17612	0	17612	0	100.00	14767	6391	8376	43.28	56.72	12268	1396	10872	11.38	88.62
Rocker	4523	0	4523	0	100.00	4383	1157	3226	26.40	73.60	9639	360	9279	3.73	96.27
Dragon	8017	12	8005	0.15	99.85	9557	585	8972	6.12	93.88	11666	405	11261	3.47	96.53
Banana	8273	110	8163	1.33	98.67	8368	6475	1893	77.38	22.62	10733	3	10730	0.03	99.97

根据 2.1 节的描述可知,对于当前观测方位下 的视觉目标,其表面可分为两部分:可见表面和未知 表面.其中,未知表面又包括背面和被遮挡表面.而 视觉目标遮挡区域内的未知表面则包括一部分背面 和被遮挡表面.同时,由方法原理可知,文献[8]中方 法计算出的下一最佳观测方位主要观测当前观测方 位下视觉目标遮挡区域外的背面,文献[10]中方法 计算出的下一最佳观测方位主要观测当前观测方位 下信息增益较大点的相邻未知表面.而本文方法则 按遮挡现象的显著性分为两种情况计算,两种情况 下计算出的下一最佳观测方位分别主要观测遮挡区 域内的未知表面和视觉目标的背面.在此基础上,结 合图 12 和表 3 分析可知,与文献[8]方法相比,对于

视觉目标 Duck 和 Mole 而言,由于其遮挡区域内的 未知表面较小,遮挡区域外的背面较大,故本文方法 的新增点数相对较少,新增率相对较低;对于视觉目 标 Bunny、Rocker 和 Dragon 而言,由于其遮挡区域 内的未知表面较大,遮挡区域外的背面较小,故本文 方法的新增点数明显较多(尽管新增率略低);对于 视觉目标 Banana 而言,由于其在当前观测方位下不 存在遮挡区域,因此本文方法计算出的下一最佳观 测方位主要观测视觉目标的背面,故本文方法的新 增点数和新增率与文献[8]中方法相当且略好.文献 [10]方法通过相关信息增益的计算来综合考虑背面 和被遮挡表面,根据已观测点中信息增益较大点及 其法向量计算下一最佳观测方位,将计算出的最佳 观测这些已观测点中信息增益较大点的观测方位作 为下一最佳观测方位.结合图 12 和表 3 可以看出, 对于视觉目标 Duck、Mole、Banana 而言,由于其表 面平滑或无明显遮挡,致使计算出的下一最佳观测 方位与当前观测方位相差相对较小,所以其新增点 数相对较少,新增率相对较低;对于视觉目标 Bunny、 Rocker 和 Dragon 而言,由于其表面复杂或有明显遮 挡,致使计算出的下一最佳观测方位与当前观测方位 相差相对较大,所以其新增点数相对较多,新增率相 对较高.与本文方法相比,由于文献[10]方法注重对 视觉目标信息增益较大点的相邻未知表面的观测,而 本文方法注重对视觉目标遮挡部分的观测,根据实验 结果来看,在新增点以及新增率上,本文方法都比较 好.综合来看,本文方法与已有方法侧重不同,对于存 在显著遮挡现象的视觉目标,本文方法适用性较好.

此外,我们也将本文方法与文献[8,10]中的方 法进行了实时性方面的对照比较,各方法确定下一 最佳观测方位所需要的时间信息如表4所示.

表 4 不同方法确定下一最佳	≌观测方位的时间消耗
----------------	------------

方法	时间消耗/ms								
名称	10	2		-	- A	~	消耗/ms		
文献[8]	19671.65	19740.58	19613.92	19549.33	19561.14	19661.27	19632.98		
文献[10]	9334.64	30072.07	28098.03	4056.27	4847.11	5158.38	13594.42		
本文	25.49	21.72	21.91	21.46	21.27	16.02	21.31		

通过表 4 给出的不同方法的时间消耗信息可以 看出,文献[8,10]和本文方法的平均时间消耗分别为 19632.98ms(约19.63s)、13594.42ms(约13.59s) 和 21.31 ms. 与文献 [8] 和文献 [10] 中所提方法相 比,本文所提方法具有较好的实时性.需说明的是, 实验中,在利用本文所提方法确定下一最佳观测方 位之前,我们使用了文献「16]中所提的方法来计算 深度图像中的遮挡边界,实验中该方法计算遮挡边 界的平均时间消耗为 193.77 ms. 因此,实验总的平 均时间消耗为215.08ms,其中,计算遮挡边界的时 间消耗占总时间消耗的 90.09%,而确定下一最佳 观测方位的时间消耗仅占 9.91%. 由此可见,即使 将计算遮挡边界的时间消耗也加以考虑,利用所提 方法确定下一最佳观测方位时的实时性仍然较好. 在实际应用中,也可以通过采用其他的遮挡检测方 法来进一步减少总的时间消耗,从而进一步提高具 体应用的实时性.

# 5 结 论

本文提出一种利用遮挡信息确定下一最佳观测 方位的方法.所提方法的贡献主要体现在3个方面: (1)提出深度图像中最大深度差相邻点的概念.应 用遮挡边界点及其最大深度差相邻点信息即可对视 觉目标遮挡区域的外接表面进行建模并可将遮挡 区域的外接表面划分为若干小平面,进而可计算出 各小平面的最佳观测方向及观测中心点,为确定最 佳小平面集合及下一最佳观测方位奠定了基础; (2)提出一种遮挡区域外接表面最佳小平面集合的 确定方法.该方法基于各小平面的最佳观测方向,结 合积分原理,采用投影降维思想,将三维空间内的问 题转化到二维平面内进行求解,计算出最佳小平面 集合.该集合中的小平面组合在一起即为摄像机在 下一最佳观测方位所能观测到的遮挡区域外接表 面:(3)提出一种下一最佳观测方位确定算法.该算 法基于计算出的最佳小平面集合,通过综合考虑集 合中各小平面的影响,以小平面的面积作为其对最 终计算结果影响力的衡量标准,对集合中各小平面 的最佳观测方向及观测中心点加权求和求出下一最 佳观测方位.与已有方法相比,所提方法的优点是: (1)所需输入数据为深度图像、该图像中的遮挡边 界以及获取该幅深度图像时摄像机的内外参数,无 需预先获取视觉目标的先验知识;(2)基于分治思 想,依据深度图像中的遮挡边界点及其最大深度差 相邻点的信息确定下一最佳观测方位,对视觉目标 表面形状无依赖关系,适用于具有不同型面的视觉 目标;(3)以遮挡区域为研究对象,利用遮挡区域外 接表面最佳小平面集合实现了对下一最佳观测方位 简便、高效地求解,整个过程未对摄像机位置做特殊 限制.最后,需指出的是,所提方法在噪声处理和遮 挡区域外接表面面积阈值自适应选取方面存在一定 的不足,未来我们将重点针对方法的抗噪性和自适 应性展开研究.

#### 参考文献

[1] Bottino A, Laurentini A. What's NEXT? An interactive next best view approach. Pattern Recognition, 2006, 39(1): 126-132

- [2] Trummer M, Munkelt C, Denzler J. Online next-best-view planning for accuracy optimization using an extended E-criterion //Proceedings of the 20th International Conference on Pattern Recognition. Istanbul, Turkey, 2010: 1642-1645
- [3] Haner S. Heyden A. Covariance propagation and next best view planning for 3D reconstruction//Proceedings of the 12th European Conference on Computer Vision. Florence, Italy, 2012: 545-556
- [4] Connolly C. The determination of next best views//Proceedings of the IEEE International Conference on Robotics and Automation. Missouri, USA, 1985: 432-435
- [5] Maver J, Bajcsy R. Occlusion as a guide for planning the next view. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(5): 417-432
- [6] Whate P, Ferrie F P. Autonomous exploration: Driven by uncertainty. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(3): 193-205
- [7] Pito R. Solution to the next best view problem for automated surface acquisition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(10): 1016-1030
- [8] Banta J E, Wong L M, Dumont C, Abidi M A. Next-bestview system for autonomous 3-D object reconstruction. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 2000, 30(5): 589-598
- [9] Chen S Y, Li Y F. Automatic sensor placement for modelbased robot vision. IEEE Transactions on Systems, Man,



ZHANG Shi-Hui, born in 1973, Ph. D., professor, Ph. D. supervisor. His current research interests include visual information processing and pattern recognition. and Cybernetics, Part B: Cybernetics, 2004, 34(1): 393-408

- [10] Li Y F, Liu Z G. Information entropy-based viewpoint planning for 3-D object reconstruction. IEEE Transactions on Robotics, 2005, 21(3): 324-337
- [11] Scott W R. Model-based view planning. Machine Vision and Applications, 2009, 20(1): 47-69
- [12] Banta J E, Zhien Y, Wang X Z, et al. Best-next-view algorithm for three-dimensional scene reconstruction using range images//Proceedings of the International Society for Optical Engineering. Philadelphia, USA, 1995; 418-429
- [13] Zhang Shi-Hui, Zhang Yu-Jie, Kong Ling-Fu. Self-occlusion detection algorithm combining depth image and optimal segmentation threshold iteration. Chinese High Technology Letters, 2010, 20(7): 754-757(in Chinese)
  (张世辉,张煜婕,孔令富. 结合深度图像和最佳分割阈值迭 代的自遮挡检测算法. 高技术通讯, 2010, 20(7): 754-757)
- [14] Zhang S, Gao F, Kong L. A self-occlusion detection approach based on range image of vision object. ICIC Express Letters, 2011, 5(6): 2041-2046
- Zhang S, Yan S. Depth estimation and occlusion boundary recovery from a single outdoor image. Optical Engineering, 2012, 51(8): 087003; 1-11
- [16] Zhang S, Liu J. A self-occlusion detection approach based on depth image using SVM. International Journal of Advanced Robotic Systems, 2012, 9(12): 230: 1-8

LIU Jian-Xin, born in 1988, M.S. candidate. His current research interest is visual information processing.

**KONG Ling-Fu**, born in 1957, Ph. D., professor, Ph. D. supervisor. His current research interests include machine vision and intelligent information processing.

#### Background

Next best view (NBV) planning is always an important issue in the fields of three-dimensional reconstruction, robot navigation, automatic assembly, object recognition and others. It means to determine the next sensor direction and position which will extract the greatest amount of unknown scene information according to the current viewing information, thus the complete information about the visual object surface can be obtained at the fewest viewpoints. The existing NBV methods are inadequate in aspects such as dealing with occlusion, not suitable for visual object with complex surface and without generality etc. Therefore, how to design a more general approach which can quickly find the next meaningful view remains to be a challenge. This paper proposes a novel NBV approach based on the depth image of a visual object and its occlusion information, which has a good performance and generality.

This work was supported by the National Natural Science Foundation of China with the titles "Research on Occlusion Detection and Avoidance Approach Based on the Visual Object Depth Information for Active Vision Tasks" (No. 61379065) and the Natural Science Foundation of Hebei Province with the title "Research on Occlusion Key Technologies for Visual Object Based on Depth Image" (No. F2014203119) etc. These foundations focus on occlusion detecting and avoiding. Our group has been working on this field for many years and some achievements have been published or accepted by journals at home and abroad.