

基于单幅深度图像遮挡信息的 下一最佳观测方位确定方法

张世辉^{1,2}, 张钰程¹

(1. 燕山大学信息科学与工程学院, 河北秦皇岛 066004; 2. 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: 如何根据当前观测到的信息确定摄像机的下一最佳观测方位是视觉领域一个具有挑战性的问题. 本文提出一种基于单幅深度图像利用遮挡信息求解下一最佳观测方位的方法. 该方法首先利用当前观测方位下获得的深度图像中的遮挡信息对遮挡区域外接表面进行四边形剖分, 从而建立遮挡区域外接表面模型; 然后通过综合考虑下一观测过程中的可见四边形信息以及观测损失信息构造下一最佳观测方位模型; 最后采用梯度下降法求解所建模型得到下一最佳观测方位. 与已有方法相比, 所提方法无需将摄像机位置固定于某一表面, 也无需获取视觉目标的先验知识. 实验结果验证了所提方法的可行性和有效性.

关键词: 下一最佳观测方位; 深度图像; 遮挡信息; 遮挡区域外接表面; 梯度下降法

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2016)02-0445-08

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2016.02.028

Determining Next Best View Based on Occlusion Information of a Single Depth Image

ZHANG Shi-hui^{1,2}, ZHANG Yu-cheng¹

(1. School of Information Science and Technology, Yanshan University, Qinhuangdao, Hebei 066004, China;

2. The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Qinhuangdao, Hebei 066004, China)

Abstract: How to determine camera's next best view based on current information is a challenging problem in visual field. A next best view approach was proposed based on the occlusion information of a single depth image. Firstly, to establish the model for occlusion region external surface, the quadrilateral meshes for occlusion region external surface were obtained according to the occlusion information of a depth image in initial view. Secondly, the model for next best view was constructed by considering both the visible quadrangle and the loss information in next view. Finally, the next best view was achieved by solving the model with gradient descent method. Compared with the existing methods, the proposed approach does not limit the camera position on a fixed surface or need the priori knowledge of visual object. Experimental results demonstrate its feasibility and effectiveness.

Key words: next best view; depth image; occlusion information; external surface of occlusion region; gradient descent method

1 引言

下一最佳观测方位(Next Best View, NBV)的确定一直是场景探测、三维重建、目标识别等领域重要研究问题之一,它是基于当前观测方位下摄像机获取的信息,确定出摄像机的下一观测方位,使摄像机能够在下一次观测中最大量地获取到当前观测方位下未知的信息.

目前,下一最佳观测方位问题的研究主要基于两种图像:彩色图像和深度图像. 基于彩色图像的下一最

佳观测方位确定算法相对较少,并且存在一些局限性,算法效率有待提高. 代表性的方法有: Bottino^[1]提出一种交互式重建方法确定下一最佳观测方位,其基于最优重建条件判断是否需要继续观测,但该方法不适用于复杂表面物体. Haner^[2]基于协方差传播理论估计不确定性从而确定下一最佳观测方位,但计算耗时较大.

与彩色图像不同,深度图像中的像素值表示场景中某一点到摄像机成像平面的距离,这意味着基于深度图像更易获取场景的三维信息,故现有的下一最佳观测方位确定

算法大多基于深度图像实现. 代表性的成果有: Connolly^[3] 采用八叉树模型描述视觉目标, 对每个节点进行标记从而确定出下一最佳观测方位, 但该方法未考虑遮挡因素, 所以遮挡现象越严重, 越无法保证方法的正确性. Maver^[4] 提出一种利用遮挡信息确定下一最佳观测方位的方法, 用多边形近似遮挡区域, 但该方法基于特定设备, 不具有良好的普适性. Pito^[5] 基于 PS (Positional Space) 算法从大量潜在观测方位中确定出下一最佳观测方位, 但摄像机位置被限制在固定表面, 且 PS 算法在观测光线的投影过程中会花费大量的时间. Banta^[6] 提出一种基于体素块的不同标记以及整体观测策略确定下一最佳观测位置的方法, 该方法中摄像机的位置同样被限制在固定表面, 并且需要预先获取场景的先验知识. Chen^[7] 提出一种基于模型的利用 HG (Hierarchical Genetic) 算法和 Christofides 算法解决视点规划问题的方法, 该方法不仅需要获取场景的先验知识, 而且计算耗时也较大. Li^[8] 提出一种基于 B 样条计算信息熵的视点规划方法, 把摄像机能够获得模型最大信息熵的方位作为下一最佳观测方位, 但该方法计算出的摄像机位置同文献 [5, 6] 一样也被限定在一个固定表面, 且方法耗时受已观测点云数量的影响较大. Scott^[9] 提出一种基于模型的 3M (Modified Measurability Matrix) 算法解决视点规划问题, 该方法同样需要获取场景的先验知识. Vasquez-Gomez^[10] 通过效用函数 (Utility Function) 对生成的候选观测方位进行排序求得下一最佳观测方位, 但其视点评价过程需要花费大量的时间.

综上所述可知, 现有下一最佳观测方位确定方法普遍存在摄像机位置受限、需要先验知识或耗时较大等方面的不足. 鉴于此, 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. 该方法首先从当前观测方位下获取一幅深度图像, 然后基于遮挡信息构造下一最佳观测方位模型, 最后利用梯度下降法求解所建模型得到下一最佳观测方位. 实验结果验证了所提方法的可行性和有效性.

2 方法概述

2.1 下一最佳观测方位问题分析

观测方位是指摄像机观测视觉目标时的方向和位置. 在观测过程中, 遮挡现象几乎无处不在. 图 1 为摄像机观测一个由矩形 $ABCD$ 和矩形 $GHIJ$ 构成的理想视觉模型的示意图. 在当前观测方位下产生了遮挡现象, AB 、 AD 和 BC 为遮挡边界, 五面体 $ABCFED$ 内的空间区域为遮挡区域, 面 ADE 、面 BCF 和面 $ABFE$ 构成遮挡区域的外接表面. 为了获取视觉目标遮挡区域内的信息则需要调整摄像机的观测方位进行进一步观测. 在所有可能的下一观测方位中, 能够最大量观测遮挡区域内未知信息的观测方位即为下一最佳观测方位.

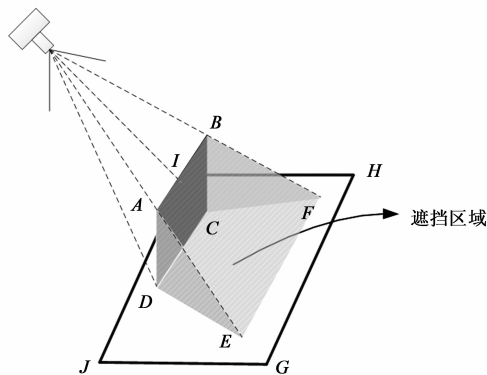


图1 摄像机观测理想视觉模型示意图

分析可知, 摄像机和遮挡区域是下一最佳观测方位问题所涉及的两个主要研究对象. 如果以摄像机为研究对象求解下一最佳观测方位问题, 则需要预先给出摄像机所有可能的观测方位, 因此, 为了获取有穷的观测方位集合就需要对摄像机的观测位置加以限制. 同时, 由于摄像机可能的观测方位一般都比较, 因此在大量可能的观测方位中确定下一最佳观测方位的过程通常耗时较大. 但是, 如果以遮挡区域为研究对象求解下一最佳观测方位问题, 则所需信息主要为当前观测方位下视觉目标的遮挡信息, 因此在确定下一最佳观测方位的过程中可以避免上述问题.

由于遮挡区域内的信息在当前观测方位下是未知的, 因此需要利用遮挡区域的外接表面信息近似地代表遮挡区域内的信息^[4], 根据下一最佳观测方位问题的描述, 则求解观测遮挡区域内未知信息的最佳观测方位问题可以近似地转化为求解观测遮挡区域外表面的最佳观测方位问题.

2.2 方法总体思想

基于上述对下一最佳观测方位问题的分析, 本文提出一种基于单幅深度图像利用遮挡信息求解下一最佳观测方位的方法. 方法总体思想如下: 首先, 在当前观测方位下获取视觉目标的深度图像. 其次, 检测和计算深度图像中遮挡信息, 并对遮挡区域外接表面进行四边形剖分, 同时计算得到各四边形的中心点、法向量和面积. 然后, 基于四边形信息构造下一最佳观测方位模型. 最后, 求解所建模型得到下一最佳观测方位. 所提方法总体流程如图 2 所示.

3 基于遮挡信息确定下一最佳观测方位

3.1 遮挡区域建模

3.1.1 获取与遮挡边界对应的下邻接边界

本文利用深度图像中的遮挡信息确定下一最佳观测方位, 由于获取深度图像中遮挡信息的方法已有文献可借鉴^[11-13], 故本文不再对获取遮挡信息的遮挡检测过程

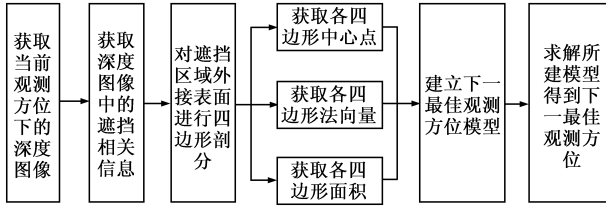


图2 求解下一最佳观测方位总体流程

进行论述,而是重点论述基于遮挡信息的下一最佳观测方位确定过程.获取深度图像中的遮挡边界后,为了确定下一最佳观测方位,下一步需要计算与遮挡边界对应的下邻接边界.以图1中理想视觉模型的遮挡边界 AB 为例进行分析, AB 对应的下邻接边界应为 EF ,但在真实的三维空间中 EF 在当前观测方位下并不可见,所以本文选取与 EF 相邻的可见边界近似作为 AB 对应的下邻接边界.图3为图1中摄像机获取的深度图像,网状线条为遮挡边界,点状线条为遮挡边界对应的下邻接边界,其中,点 A 与点 E 重合,点 B 与点 F 重合.

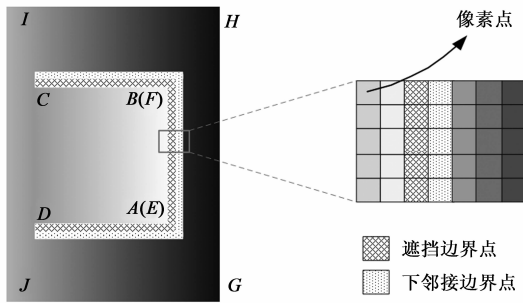


图3 深度图像及局部像素点示意图

计算深度图像中遮挡边界对应的下邻接边界的具体方法如下.

首先为每一个遮挡边界像素点寻找到一个下邻接边界像素点与之对应.记遮挡边界像素点集合为 $\{P\}_o$,记下邻接边界像素点集合为 $\{P\}_d$.在 $\{P\}_o$ 中任取一点记为 P_o ,其深度值记为 d_o ,在 P_o 的八邻域内寻找与 P_o 对应的点 P_d ,其深度值记为 d_d .为了保证集合 $\{P\}_o$ 与集合 $\{P\}_d$ 中的元素一一对应,需要满足 P_d 既不是遮

挡边界像素点,也不与其他遮挡边界像素点对应.同时,考虑到为了确定出最优的下邻接边界,本文定义 P_o 的下邻接边界像素点 P_d 的深度值 d_d 满足 $\arg \max_{d_d} (d_d - d_o)$.通过上述方法即可求得每一个遮挡边界像素点对应的下邻接边界像素点.所有遮挡边界像素点对应的下邻接边界像素点形成的边界即为遮挡边界对应的下邻接边界.在三维空间中,遮挡边界与其对应的下邻接边界沿摄像机观测方向形成的曲面则为遮挡区域外接表面.

3.1.2 基于边界信息建立遮挡区域外接表面模型

为了建立遮挡区域外接表面模型,基于遮挡边界和下邻接边界信息,依次从遮挡边界中取出第 i 对相邻的两个遮挡边界点 P_{OL_i} 、 P_{OR_i} ,同时取出两个分别与之对应的下邻接边界点 P_{DL_i} 、 P_{DR_i} ,上述四个点所围成的空间四边形 $Patch_i$ 即为剖分遮挡区域外接表面得到的四边形之一.基于此思想,即可实现对遮挡区域外接表面的四边形剖分.图4给出了对遮挡区域外接表面进行分析的示意图.其中,图4(a)为对遮挡区域外接表面进行四边形剖分后的结果,图4(b)为四边形对应的三维空间结构,图4(c)为四边形对应的深度图像中的结构.

记四边形 $Patch_i$ 的四个顶点 P_{OL_i} 、 P_{OR_i} 、 P_{DL_i} 、 P_{DR_i} 的三维坐标分别为 $(x_{OL_i}, y_{OL_i}, z_{OL_i})$ 、 $(x_{OR_i}, y_{OR_i}, z_{OR_i})$ 、 $(x_{DL_i}, y_{DL_i}, z_{DL_i})$ 、 $(x_{DR_i}, y_{DR_i}, z_{DR_i})$,则四边形 $Patch_i$ 的中心点坐标 $M_i(x_{M_i}, y_{M_i}, z_{M_i})$ 的计算公式定义为

$$\begin{bmatrix} x_{M_i} \\ y_{M_i} \\ z_{M_i} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} x_{OL_i} & x_{OR_i} & x_{DL_i} & x_{DR_i} \\ y_{OL_i} & y_{OR_i} & y_{DL_i} & y_{DR_i} \\ z_{OL_i} & z_{OR_i} & z_{DL_i} & z_{DR_i} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad (1)$$

由于空间四边形的四个顶点不一定共面,并且本文前述将深度图像中与遮挡边界相邻的可见边界近似作为与其对应的下邻接边界,故为了减小遮挡区域外接表面建模的误差,本文选取 P_{DL_i} 与 P_{DR_i} 的中点与两个遮挡边界点 P_{OL_i} 和 P_{OR_i} 组成的三角形的法向量作为四边形 $Patch_i$ 的法向量(见图4(b)),则四边形法向量 v_i 的计算公式定义为

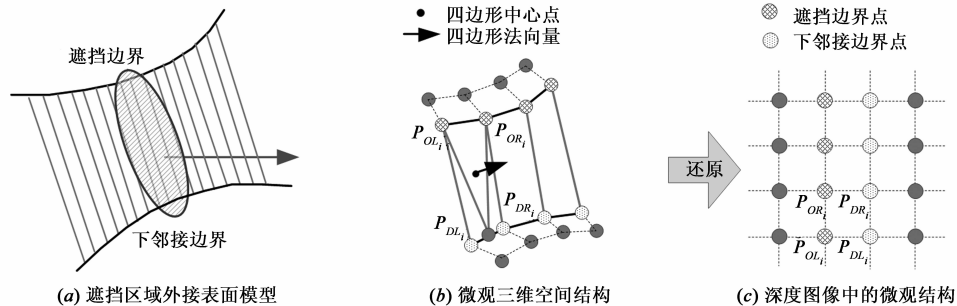


图4 遮挡区域外接表面模型微观分析

$$\mathbf{v}_i = \begin{vmatrix} i & j & k \\ \frac{x_{DL_i} + x_{DR_i}}{2} - x_{OL_i} & \frac{y_{DL_i} + y_{DR_i}}{2} - y_{OL_i} & \frac{z_{DL_i} + z_{DR_i}}{2} - z_{OL_i} \\ x_{OR_i} - x_{OL_i} & y_{OR_i} - y_{OL_i} & z_{OR_i} - z_{OL_i} \end{vmatrix} \quad (2)$$

基于四边形法向量 \mathbf{v}_i 的计算公式(2),则四边形面积 S_i 的计算公式可定义为

$$S_i = 1/2 \|\mathbf{v}_i\|_2 \times 2 = \|\mathbf{v}_i\|_2 \quad (3)$$

则在当前观测方位下遮挡区域外接表面总面积 S_Σ 的计算公式可定义为

$$S_\Sigma = \sum_{i=1}^n S_i \quad (4)$$

式中, n 为剖分遮挡区域外接表面所得四边形的个数.

3.2 基于遮挡区域外接表面信息构造下一最佳观测方位模型

对遮挡区域外接表面建模后,为了构造下一最佳观测方位模型,需要判断在下一最佳观测方位下剖分所得 n 个四边形中哪些四边形是可见的,且期望可以最大量地观测到四边形的信息.

记 $X_{NV}(x_{NV}, y_{NV}, z_{NV})$ 为下一观测方位下摄像机的位置,向量 $\mathbf{u}_{NV} = (p_{NV}, q_{NV}, t_{NV})$ 为下一观测方位下摄像机的观测方向,则第 i 个四边形的中点 $M_i(x_{M_i}, y_{M_i}, z_{M_i})$ 指向下一观测位置的向量为 $\mathbf{r}_i = (x_{NV} - x_{M_i}, y_{NV} - y_{M_i}, z_{NV} - z_{M_i})$. 若向量 \mathbf{r}_i 与第 i 个四边形的法向量 \mathbf{v}_i 的夹角 $\theta_i \geq \frac{\pi}{2}$, 则表明第 i 个四边形在下一观测方位下是不可见的,故本文定义当向量 \mathbf{r}_i 与第 i 个四边形的法向量 \mathbf{v}_i 的夹角 $\theta_i < \frac{\pi}{2}$ 时,该四边形在下一观测方位下是可见的.

在观测过程中,若可见的第 i 个四边形的中点到下一观测方位下摄像机平面的距离远大于或远小于其到当前观测方位下摄像机平面的距离,则意味着在当前观测方位下与下一观测方位下观测到的像素点所代表的三维空间中的真实信息相差较大,如果不加考虑则会导致对两次观测所得像素点进行三维匹配时产生较大的误差. 因此,本文首先通过将可见的第 i 个四边形的中点到下一观测方位下摄像机平面的距离 d_2 的平方与该点到当前观测方位下摄像机平面的距离 d_1 的平方的比值减去单位 1 再求平方的方式衡量两次观测距离的差异. 然后利用指数函数的性质对两次观测距离的差异进行归一化,归一化计算公式定义为

$$\delta_i(X_{NV}, \mathbf{u}_{NV}) = \begin{cases} e^{-(\frac{d_2}{d_1} - 1)^2}, & \theta_i < \frac{\pi}{2} \\ 0, & \text{其他} \end{cases} \quad (5)$$

式中, $d_2 = \frac{|\mathbf{u}_{NV} \cdot \mathbf{r}_i|}{\|\mathbf{u}_{NV}\|_2} = \frac{|\mathbf{u}_{NV} \cdot (X_{NV} - M_i)|}{\|\mathbf{u}_{NV}\|_2}$, $d_1 =$

$\frac{|\mathbf{u}_{CV} \cdot (X_{CV} - M_i)|}{\|\mathbf{u}_{CV}\|_2}$, θ_i 为向量 \mathbf{r}_i 与第 i 个四边形的法向量 \mathbf{v}_i 的夹角. d_1 中, X_{CV} 为当前观测方位下摄像机的位置, \mathbf{u}_{CV} 为当前观测方位下摄像机的观测方向.

由于在相同观测距离下,只有当摄像机观测方向与四边形法向量平行时,摄像机获取的四边形面积最大. 若记第 i 个四边形 $Patch_i$ 的面积为 S_i ,则摄像机在下一观测方位下观测到 $Patch_i$ 的面积 $S'_i(X_{NV})$ 可定义为

$$S'_i(X_{NV}) = \begin{cases} S_i \frac{\mathbf{v}_i \cdot \mathbf{r}_i}{\|\mathbf{v}_i\|_2 \|\mathbf{r}_i\|_2}, & \theta_i < \frac{\pi}{2} \\ 0, & \text{其他} \end{cases} \quad (6)$$

式中, \mathbf{v}_i 为第 i 个四边形的法向量, $\mathbf{r}_i = X_{NV} - M_i = (x_{NV} - x_{M_i}, y_{NV} - y_{M_i}, z_{NV} - z_{M_i})$ 为下一观测方位下第 i 个四边形中点 $M_i(x_{M_i}, y_{M_i}, z_{M_i})$ 指向下一观测位置 $X_{NV}(x_{NV}, y_{NV}, z_{NV})$ 的向量, θ_i 为向量 \mathbf{r}_i 与第 i 个四边形的法向量 \mathbf{v}_i 的夹角.

结合上述分析,若同时考虑摄像机在下一观测方位下能够最大量地观测到遮挡区域外接表面剖分所得四边形且能够最大量地降低观测误差两方面因素,则可以把公式(5)中归一化的结果 $\delta_i(X_{NV}, \mathbf{u}_{NV})$ 作为公式(6)中面积 $S'_i(X_{NV})$ 的权值,也就是说,将摄像机观测方位信息作为其获取到的第 i 个四边形面积的权值. 基于此思想,本文最终构造出的下一最佳观测方位模型为

$$S_{NBV}(X, \mathbf{u}) = \arg \max_{(X, \mathbf{u})} \sum_{i=1}^n \delta_i(X, \mathbf{u}) S'_i(X) \quad (7)$$

式中, $X(x_{NBV}, y_{NBV}, z_{NBV})$ 为下一最佳观测方位下摄像机的位置,向量 $\mathbf{u} = (p_{NBV}, q_{NBV}, t_{NBV})$ 为下一最佳观测方位下摄像机的观测方向.

3.3 下一最佳观测方位模型求解

分析式(7)模型可知其并非一个凸模型,故很难直接利用凸优化方法得到其全局最优解. 可行的解法有随机法、填充函数法、积分水平集法、区域分割法、遗传算法、梯度下降法等. 这些方法中,有的理论上就不能得到真正的全局最优解,如随机法、填充函数法、遗传算法和梯度下降法;有的虽然在理论上可得到全局最优解,但需要非常大的计算量,如积分水平集法和区域分割法. 就已有方法而言,除了某些特殊的非凸问题能够求得全局最优解外,对于一般的非凸问题,尚无求解全局最优的有效手段. 考虑到该模型较易计算梯度,且梯度下降法在迭代过程中需要的存储单元相对较少,在很大程度上可减少算法的空间复杂度,故综合考虑方法的可行性和时间消耗,本文采用梯度下降法求解式(7)的下一最佳观测方位模型.

为了采用梯度下降法求解式(7)模型,首先需要选取较好的初始迭代方位. 在不考虑摄像机视角场和物

体面面的理想情况下,遮挡区域外接表面剖分所得所有四边形法向量之和的反方向应为最优的初始迭代方向 $\mathbf{u}_0 = (p_{NBV_0}, q_{NBV_0}, t_{NBV_0})$. 相应地,初始迭代位置也应在此方向上选取. 为了确定初始迭代位置,以当前观测方位下所有已观测点的重心为起点,在初始迭代方向 \mathbf{u}_0 上求得一点 $X_0(x_{NBV_0}, y_{NBV_0}, z_{NBV_0})$,使点 X_0 到已观测点重心的距离与当前观测位置到已观测点重心的距离相等. 基于上述思路即可得到初始迭代方位 (X_0, \mathbf{u}_0) .

初始迭代方位 (X_0, \mathbf{u}_0) 确定后,即可基于式(7)模型对其六个变量分别求梯度. 若记迭代次数为 k ,迭代步长为 λ ,则其六个变量沿梯度方向的更新迭代公式为

$$\begin{aligned} x_{NBV_{k+1}} &= x_{NBV_k} + \lambda \frac{\partial S_{NBV_k}(X_k, \mathbf{u}_k)}{\partial x_{NBV_k}} \\ y_{NBV_{k+1}} &= y_{NBV_k} + \lambda \frac{\partial S_{NBV_k}(X_k, \mathbf{u}_k)}{\partial y_{NBV_k}} \\ z_{NBV_{k+1}} &= z_{NBV_k} + \lambda \frac{\partial S_{NBV_k}(X_k, \mathbf{u}_k)}{\partial z_{NBV_k}} \\ p_{NBV_{k+1}} &= p_{NBV_k} + \lambda \frac{\partial S_{NBV_k}(X_k, \mathbf{u}_k)}{\partial p_{NBV_k}} \\ q_{NBV_{k+1}} &= q_{NBV_k} + \lambda \frac{\partial S_{NBV_k}(X_k, \mathbf{u}_k)}{\partial q_{NBV_k}} \\ t_{NBV_{k+1}} &= t_{NBV_k} + \lambda \frac{\partial S_{NBV_k}(X_k, \mathbf{u}_k)}{\partial t_{NBV_k}} \end{aligned} \quad (8)$$

迭代终止时得到的六个变量 $X_{k+1}(x_{NBV_{k+1}}, y_{NBV_{k+1}}, z_{NBV_{k+1}})$ 、 $\mathbf{u}_{k+1} = (p_{NBV_{k+1}}, q_{NBV_{k+1}}, t_{NBV_{k+1}})$ 即为最终求得的下一最佳观测方位 (X, \mathbf{u}) .

3.4 下一最佳观测方位确定算法

Input: 摄像机当前观测方位下获取的深度图像、遮挡边界及摄像机内外参数

Output: 下一最佳观测方位

Step 1: /* 计算遮挡边界像素点对应的下邻接边界像素点 */

For 每个遮挡边界像素点

对其八邻域内非遮挡边界像素点的深度值从大到小排序
遍历排序结果寻找到第一个没有与其他遮挡边界像素点对应的非遮挡像素点
把找到的非遮挡像素点作为当前遮挡像素点对应的下邻接边界像素点

End For

Step 2: /* 计算四边形的中心、法向量和面积 */

For 每一对相邻的遮挡边界像素点

取出它们对应的下邻接边界像素点构成四边形
依据公式(1)计算四边形的中心
依据公式(2)计算四边形的法向量
依据公式(3)计算四边形的面积

End For

Step 3: /* 确定下一最佳观测方位 */

依据计算所得四边形信息构造下一最佳观测方位模型式(7)

计算当前观测方位下所有已观测点的重心

确定较好的初始迭代观测方位 (X_0, \mathbf{u}_0) 并代入式(7)模型

依据式(8)迭代求解观测方位中的六个变量

Step 4: 输出迭代终止时的六个变量,即为求得的下一最佳观测方位

4 实验及分析

4.1 实验环境

为了验证本文方法的效果,我们基于 <http://range.informatik.uni-stuttgart.de/htdocs/html> 上 Stuttgart Range Image Database (简称 SRID) 中提供的三维物体模型进行了实验. 实验硬件环境为 CPU Intel (R) Core (TM) i7-3770 3.40GHz, 内存 8.00GB. 采用 C++ 实现下一最佳观测方位确定程序,并结合 OpenGL 实现摄像机观测物体模型获取深度图像的过程. 获取深度图像过程中,设置 OpenGL 投影矩阵的参数为(60,1,200,600),窗口大小为 400×400.

4.2 实验及分析

基于本文所提方法,我们对 SRID 中不同的三维物体模型进行了实验. 实验过程中,遮挡边界检测采用文献[13]中所提方法实现,设定摄像机相对于观测中心点的距离为 250mm,步长 λ 采用固定步长 0.1,迭代过程中当 $S_{NBV_{k+1}}(X_{k+1}, \mathbf{u}_{k+1})$ 与 $S_{NBV_k}(X_k, \mathbf{u}_k)$ 的差值小于 0.01 或迭代次数大于 2000 则终止(步长 0.1,差值 0.01 和迭代次数 2000 均为经验值). 图 5 给出了部分实验结果. 其中,图 5(a)为视觉目标名称,图 5(b)为当前观测方位下视觉目标的深度图像,图 5(c)为遮挡边界(白色线条)及下邻接边界(黑色线条),图 5(d)为遮挡区域外接表面各四边形对应的法向量(白色线条),图 5(e)为下一最佳观测方位下可见的四边形集合(白色区域),图 5(f)为所求下一最佳观测方位下视觉目标的深度图像.

分析图 5 的实验结果可知,对视觉目标 Duck 而言,由于其遮挡现象并不明显,故在所求下一最佳观测方位下观测到的可见四边形区域并不是很大,即在所求下一最佳观测方位下会观测到较多的已观测点. 但对视觉目标 Rocker、Mole、Bunny 和 Dragon 而言,由于其遮挡现象比较明显且越来越严重,故在所求下一最佳观测方位下观测到的可见四边形区域越来越大,也就是说,遮挡现象越严重,本文方法效果越好. 同时,对照图 5(b)和图 5(f)中各视觉目标的深度图像可知,本文方法求得的下一最佳观测方位符合人类视觉的观测习惯.

由于目前并不存在下一最佳观测方位评测方面的标准对照图(Ground Truth),又由于不同方法各有特点,所以很难设定统一的标准对比分析不同下一最佳观测方位算法的优劣. 为了进一步验证本文方法的可行性和有效性,我们仅基于本文方法特点对图 5 的实验结果进行了量化分析. 分析结果如表 1 所示.

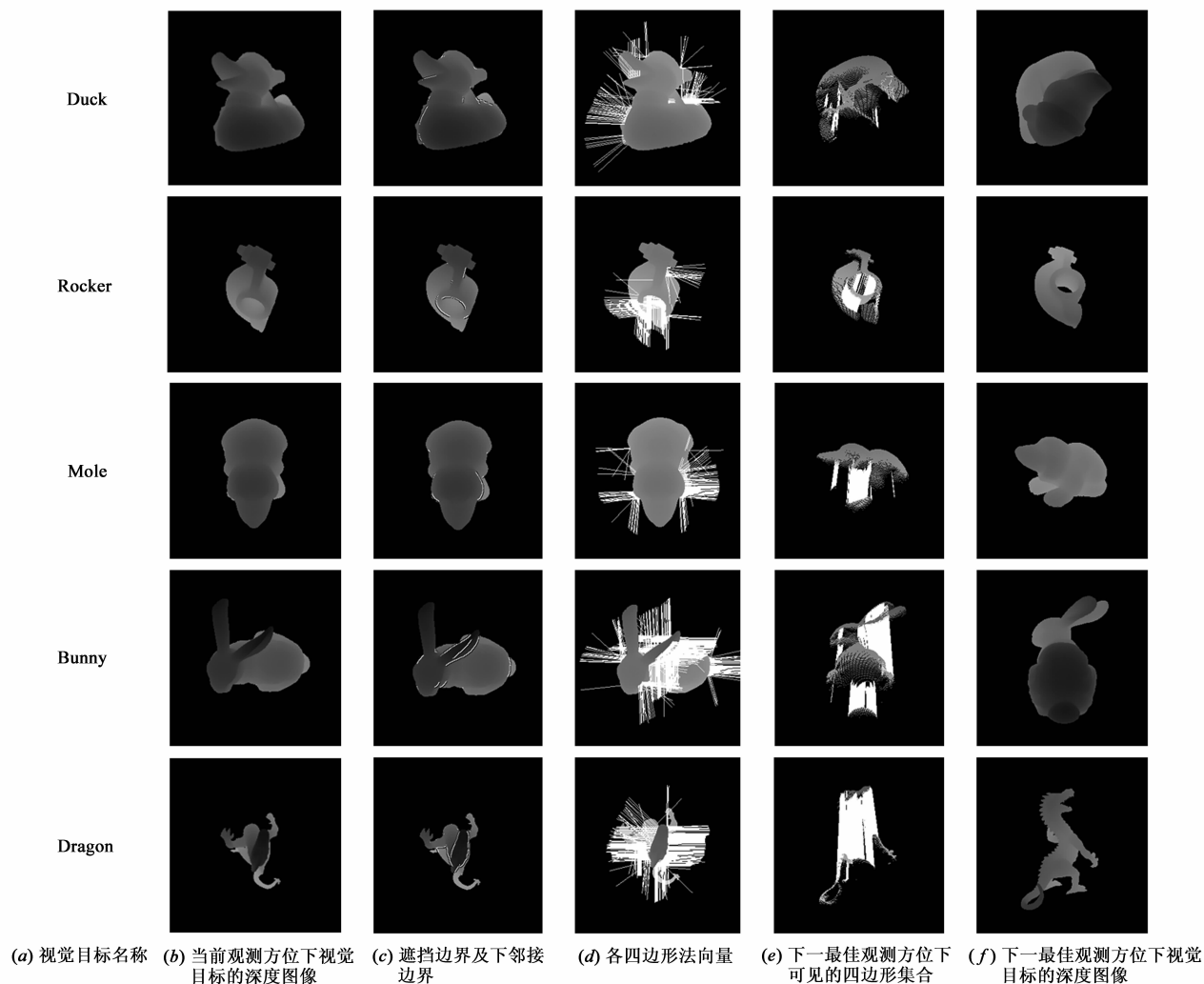


图5 下一最佳观测方位实验结果

表1中, N_{CV} 为当前观测方位下获取到的视觉目标表面点的个数, N_{NBV} 为所求下一最佳观测方位下获取到的视觉目标表面点个数, $N_{OVERLAP}$ 为两次观测中重复观测到的视觉目标表面点个数, R_{NEW} 为新增率, 计算公式为 $1 - N_{OVERLAP}/N_{NBV}$, S_{Σ} 为公式(4)计算结果, 表示当前观

测方位下视觉目标遮挡区域外接表面的总面积, S_{NBV} 为基于梯度下降法求解公式(7)所得结果, 表示在所求下一最佳观测方位下可见的四边形面积加权求和的结果, T 为求解下一最佳观测方位所需的时间消耗. 每个视觉目标对应的时间消耗为10次实验结果的平均值.

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_{NBV}	$N_{OVERLAP}$	R_{NEW}	S_{Σ}/mm^2	S_{NBV}/mm^2	T/ms
Duck	20594	17622	5178	70.62%	1911.14	940.47	40.03
Rocker	3728	8865	1802	79.67%	2810.05	1770.09	42.87
Mole	14432	12354	1930	84.38%	3025.11	1505.51	56.04
Bunny	13813	16071	1835	88.58%	7289.26	2778.82	69.51
Dragon	4737	9090	416	95.42%	8450.79	4096.10	83.03

由表1可以看出, 摄像机在下一最佳观测方位下获取到的视觉目标上的点具有较高的新增率, 并且当






遮挡现象越严重, 即遮挡区域外接表面面积和 S_{Σ} 越大时, 本文所求下一最佳观测方位下获取到的视觉目标

点的新增率越高. 结合图 5 与表 1 进一步分析可知, 对于四边形法向量分布相对分散的视觉目标如 Duck、Mole、Bunny 和 Dragon 而言, 当遮挡现象越严重时, 本文所建模型计算出的 S_{NBV} 的值也越大, 说明摄像机在所求下一观测方位下的确可以获取到大量的当前观测方位下获取不到的信息. 同时, 对于四边形法向量分布相对集中的视觉目标 Rocker 而言, 虽然其遮挡区域外接表面面积小于视觉目标 Mole 的遮挡区域外接表面面积,

但本文所建模型计算出的 Rocker 的 S_{NBV} 的值却比 Mole 的 S_{NBV} 的值大, 这一现象反映了本文所建模型对遮挡区域分布较为集中的视觉目标具有更好的响应.

此外, 为了评估本文方法的时间消耗, 我们将其与同样基于深度图像且考虑了遮挡因素的文献[6]和文献[8]中的方法在相同环境下做了对比实验, 实验结果如表 2 所示.

表 2 不同方法确定下一最佳观测方位的时间消耗

方法名称	时间消耗/ms					平均时间消耗/ms
						
文献[6]	8246.35	8154.39	8165.29	8204.58	8296.31	8213.38
文献[8]	6153.48	952.89	5829.67	2103.39	1143.53	3236.59
本文	40.03	42.87	56.04	69.51	83.03	58.30

由表 2 的时间消耗可知, 图 5 中所有视觉目标计算下一最佳观测方位的平均耗时为 58.30ms, 而本文采用文献[13]方法检测遮挡过程的平均耗时为 47.41ms, 故即使将检测遮挡的耗时也考虑在内, 本文方法求解下一最佳观测方位的平均耗时也仅为 105.71ms, 与文献[6]、文献[8]方法求解下一最佳观测方位的耗时 8213.38ms、3236.59ms 相比, 本文方法明显具有较好的实时性.

5 结论

本文提出一种基于单幅深度图像利用遮挡信息确定下一最佳观测方位的方法. 主要贡献在于: (1) 提出一种对遮挡区域外接表面进行四边形剖分的遮挡区域建模方法, 该方法提供了一种基于可靠的已观测点对未知区域进行建模的思路; (2) 提出一种基于四边形法向量和面积信息构造下一最佳观测方位模型的方法, 该方法提供了一种根据摄像机观测距离量化观测损失的思路, 从而更利于获得合理的候选观测方位; (3) 针对发生遮挡现象的视觉目标, 给出了一种高效的下一最佳观测方位确定方法, 该方法采用梯度下降法求解下一最佳观测方位模型, 避免了寻求非凸模型全局最优解的复杂问题. 本文方法适用于不同型面的视觉目标, 具有较好的通用性和实时性, 且遮挡现象越严重, 本文方法效果越好. 未来我们将对如何自适应地计算遮挡现象的严重性进行进一步研究.

参考文献

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

126 - 132.

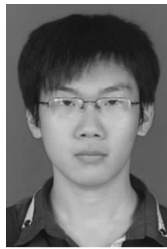
- [2] Haner S, Heyden A. Covariance propagation and next best view planning for 3D reconstruction [A]. Proceedings of the 12th European Conference on Computer Vision [C]. Heidelberg, Germany: Springer Verlag, 2012. 545 - 556.
- [3] Connolly C. The determination of next best views [A]. Proceedings of the IEEE International Conference on Robotics and Automation [C]. Washington, DC, USA: IEEE, 1985. 432 - 435.
- [4] Maver J, Bajcsy R. Occlusion as a guide for planning the next view [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(5): 417 - 432.
- [5] Pito R. Solution to the next best view problem for automated surface acquisition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(10): 1016 - 1030.
- [6] Banta J E, Wong L M, Dumont C, et al. Next-best-view system for autonomous 3-D object reconstruction [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 2000, 30(5): 589 - 598.
- [7] Chen S Y, Li Y F. Automatic sensor placement for model-based robot vision [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2004, 34(1): 393 - 408.
- [8] Li Y F, Liu Z G. Information entropy-based viewpoint planning for 3-D object reconstruction [J]. IEEE Transactions on Robotics, 2005, 21(3): 324 - 337.
- [9] Scott W R. Model-based view planning [J]. Machine Vision and Applications, 2009, 20(1): 47 - 69.
- [10] Vasquez-Gomez J I, Sucar L E, Murrieta-Cid R. Hierarchical ray tracing for fast volumetric next-best-view plan-

- ning[A]. Proceedings of the International Conference on Computer and Robot Vision[C]. Washington, DC, USA: IEEE, 2013. 181 – 187.
- [11] Zhang S, Gao F, Kong L. A self-occlusion detection approach based on range image of vision object[J]. ICIC Express Letters, 2011, 5(6): 2041 – 2046.
- [12] Zhang S, Yan S. Depth estimation and occlusion boundary recovery from a single outdoor image[J]. Optical Engineering, 2012, 51(8): 087003:1 – 11.
- [13] Zhang S, Liu J. A self-occlusion detection approach based on depth image using SVM[J]. International Journal of Advanced Robotic Systems, 2012, 9(12): 230:1 – 8.

作者简介



张世辉 男, 1973 年生于河北赞皇, 燕山大学信息科学与工程学院教授、博士生导师. 主要研究方向为视觉信息处理、模式识别、人工智能.
E-mail: sshzz@ysu.edu.cn



张钰程 男, 1990 年生于河北廊坊, 燕山大学硕士研究生. 主要研究方向为计算机视觉和模式识别.
E-mail: zycysucv@gmail.com