

基于深度图像利用随机森林实现遮挡检测

张世辉^{1,2} 刘建新¹ 孔令富^{1,2}

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 提出了一种新颖的利用随机森林检测深度图像中遮挡现象的方法。该方法从一幅深度图像中提取每个像素点的遮挡相关特征,利用随机森林分类器检测每个像素点是否为遮挡边界点,得到图像中的遮挡边界。主要贡献在于:提出了一种新的遮挡相关特征——深度值离散度特征,同时引入高斯曲率特征,并将它们与现有特征相结合来检测遮挡边界;以特征重要性和特征提取时间为衡量标准,对深度图像中的各遮挡相关特征进行了分析评估,在此基础上,选取平均深度差、最大深度差、平均曲率、高斯曲率和深度值离散度5种特征用于设计遮挡检测分类器;一种新的遮挡检测方法,利用随机森林解决深度图像的遮挡检测问题。实验结果表明,同已有方法相比,所提方法具有较高的准确性和较好的通用性。

关键词 图像处理;深度图像;遮挡检测;随机森林;遮挡相关特征;深度值离散度

中图分类号 TP391 **文献标识码** A **doi**: 10.3788/AOS201434.0915003

Using Random Forest for Occlusion Detection Based on Depth Image

Zhang Shihui^{1,2} Liu Jianxin¹ Kong Lingfu^{1,2}

¹*School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China*
²*Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Qinhuangdao, Hebei 066004, China*

Abstract A novel occlusion detection approach is proposed for depth image by using Random Forest. The occlusion related features of each pixel in the depth image are extracted, and then the Random Forest classifier is used for detecting whether each pixel is an occlusion boundary point or not. All the occlusion boundaries in the input depth image are obtained. This work is distinguished by three contributions. A new occlusion related feature named depth dispersion is proposed and the Gaussian curvature feature is introduced, and both of them are used for occlusion detection by combining with other features. All the occlusion related features in depth image are analyzed and evaluated by using the importance and extraction time as criterion. On this basis, five features such as average depth difference, maximal depth difference, mean curvature, Gaussian curvature and depth dispersion are selected for designing the occlusion detection classifier. A new occlusion detection approach takes the Random Forest to solve occlusion detection problem in depth image. The experimental results show that, compared with the existing methods, the proposed approach has higher accuracy and better generality.

Key words imaging processing; depth image; occlusion detection; Random Forest; occlusion related feature; depth dispersion

OCIS codes 150.5670; 150.1135; 330.1880

1 引言

视觉遮挡既是一个几何现象又是一个光学现

象。在大多数视觉研究领域,如物体识别、三维(3D)重建、目标跟踪、运动估计、视觉观测、场景绘

收稿日期: 2014-03-03; **收到修改稿日期**: 2014-05-13

基金项目: 国家自然科学基金(61379065)、河北省自然科学基金(F2014203119)、机器人技术与系统国家重点实验室开放研究项目(SKLR-S-2010-ZD-08)

作者简介: 张世辉(1973—),男,博士,教授,主要从事视觉信息处理、模式识别、智能机器人等方面的研究。

E-mail: sshhzz@ysu.edu.cn

制、机器人抓取、自动装配、航天器对接等领域都会涉及到遮挡现象。如果视觉系统无法采取正确的措施处理遮挡现象,其视觉技术将会因此而失效甚至出错,故遮挡现象的存在已成为相关课题深入研究的阻碍因素。遮挡问题的研究对大部分视觉技术的发展具有重要的基础作用,受到了学者们的广泛关注^[1-3]。

在视觉领域,解决遮挡问题的方法主要有两类:1)隐含的方法,该类方法并不直接针对遮挡现象进行处理,而是通过改善相关方法的鲁棒性以提高对遮挡现象的抗干扰能力^[4-6];2)明确的方法(遮挡检测方法),该类方法直接以遮挡现象作为研究对象,通过设计合理的方法检测出图像中的遮挡边界。已有的遮挡检测方法主要基于亮度图像(包括彩色图像和灰度图像)实现,代表性的成果有:Stein等^[7]以多幅彩色图像为依据,通过结合运动线索和局部表面观线索,提出了一种基于推理的遮挡检测方法;Hoiem等^[8]则以单幅彩色图像为依据,采用层次化区域分割方法,估计出场景中物体的深度(伪深度),然后依据深度值确定出遮挡边界;Ayvaci等^[9]提出了基于光流估计的遮挡检测方法,该方法通过将遮挡检测问题描述为变分优化问题,应用高效数值算法求解全局最优解来确定遮挡边界。相比之下,基于深度图像的遮挡检测研究由于可用信息相对较少,故成果较少。自2010年微软公司推出既可获得深度图像也可获得彩色图像的体感外围交互设备Kinect以来,深度图像更加体现出其价值所在,尤其是在图像获取与分析、人机自然交互、娱乐游戏等领域,深度图像及其相关设备越来越引起广大科技工作者及用户的关注,基于深度图像的图像处理技术已成为一个不可回避的问题。近年来,部分学者试图从视觉目标的深度图像信息入手研究遮挡情况下特定视觉任务的处理办法^[10-11],但并未给出明确的遮挡检测方法。

针对上述问题,文献[12-15]在基于深度图像的遮挡检测方法方面做了较深入的研究,取得了一定的进展。其中,文献[12]采用平均曲率特征、曲率阈值及深度阈值实现了遮挡检测;文献[13-14]基于平均深度差特征,分别利用最佳分割阈值迭代方法和简单图像统计方法计算适合的平均深度差阈值实现了遮挡检测;文献[15]采用最大深度差特征和夹角特征,利用支持向量机实现了遮挡检测。但分析发现:文献[12]中方法需人为设定阈值,即并未找到阈值选取的通用方法,具有局限性,同时,遮挡检

测结果的精度有待提高;文献[13-14]中尽管提出了自适应的阈值计算方法,但其检测结果的精度仍有待提高;文献[15]中方法虽然具有了一定的通用性且在某种程度上提高了检测精度,但仍需要改善。鉴于此,本文提出了一种利用随机森林检测深度图像中遮挡边界的方法。该方法同文献[12-15]方法一样,均利用所提遮挡相关特征并结合相关方法进行遮挡检测。不同的是,文献[12-14]方法基于阈值分割的思想检测遮挡,文献[15]方法和本文方法则基于机器学习的思想检测遮挡。虽然本文方法与文献[15]方法均基于机器学习思想进行遮挡检测,但是由于所提出的新特征及所采用的遮挡检测分类器模型不同,使得二者在方法设计和检测效果上均不相同。

2 方法概述

2.1 遮挡边界

遮挡边界是发生遮挡现象的不同物体(或同一物体上的不同部分)之间的分界线。如图1所示,在当前的观测方位下,由于B物体(遮挡物体)遮挡了A物体(被遮挡物体),导致A物体上与B物体重叠的区域不可见,图1中的红色边界即为遮挡边界。在图像中,遮挡边界由若干遮挡边界点(图像中的像素点)构成,且遮挡边界点均为遮挡物体上的像素点。所提方法的目的是检测出深度图像中的遮挡边界。深度图像是一种特殊的图像,其每个像素点的深度值为对应视觉目标上的点到参考平面(摄像机坐标系的 xoy 平面)的距离。

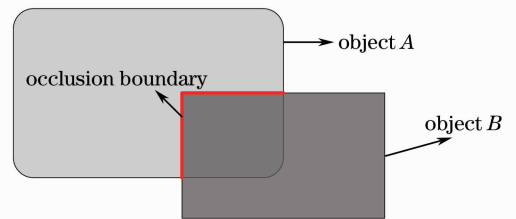


图1 遮挡边界示意图

Fig. 1 Sketch map of occlusion boundary

2.2 方法总体思想

检测遮挡边界的总体思想如下:从输入的深度图像中提取每个像素点的各类遮挡相关特征,将其归一化后得到各特征对应的特征矩阵;遍历输入的深度图像,每个像素点都被视为一个测试样本,从已计算出的特征矩阵中获取其对应的各类遮挡相关特征,将各特征值组合形成特征向量输入事先训练好的基于随机森林的遮挡检测分类器,判断该测试样

本类别(遮挡边界点或非遮挡边界点)并记录遮挡边界点在图像中位置;根据记录的遮挡边界点的位置

信息获得遮挡检测结果。所提方法的整体流程如图2所示。下面进行详细阐述。

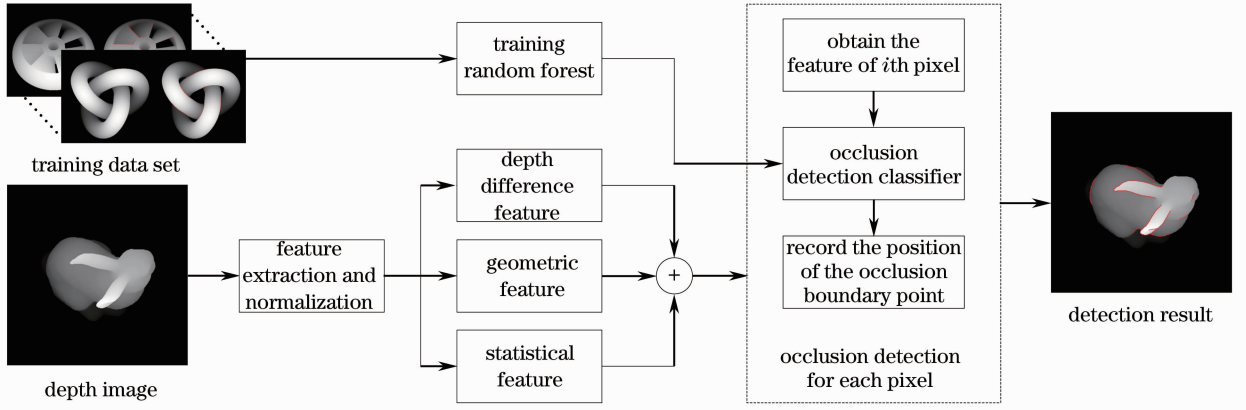


图2 遮挡检测方法整体流程

Fig. 2 Overall process of the proposed occlusion detection approach

3 遮挡相关特征描述

目前,已有的基于深度图像的遮挡相关特征^[12-15]主要包括两大类:深度差特征和几何特征。其中,深度差特征包括平均深度差特征和最大深度差特征,几何特征包括平均曲率特征和夹角特征。在对深度图像中遮挡边界点及其邻域点深度值进一步分析的基础上,又提出了深度值离散度特征。该特征不属于已有的深度差特征和几何特征,而属于统计特征。此外,还引入了高斯曲率作为遮挡相关特征。与平均曲率一样,高斯曲率也属于几何特征类。至此,基于深度图像的遮挡相关特征共有6种,分别是:平均深度差特征、最大深度差特征、平均曲率、高斯曲率、夹角特征以及深度值离散度特征。下

面分别针对各类遮挡相关特征进行详细描述。

3.1 深度差特征

在深度图像中,深度值的连续性是判断遮挡现象的重要依据之一。图3展示了深度值与视觉目标遮挡现象的关系。由图3可知,在深度图像中,遮挡边界上像素点与其8邻域内的像素点的深度值存在阶跃现象,即像素点深度值的不连续性较明显。深度差特征是一类直观反映深度图像中像素点之间深度值连续性的特征。通过分析深度图像中遮挡边界点及其8邻域内各像素点的深度值可知,遮挡边界点至少与其某一邻域点存在较大的深度值阶跃。因此,深度图像中各像素点与其邻域点的深度差可作为检测遮挡边界的依据。

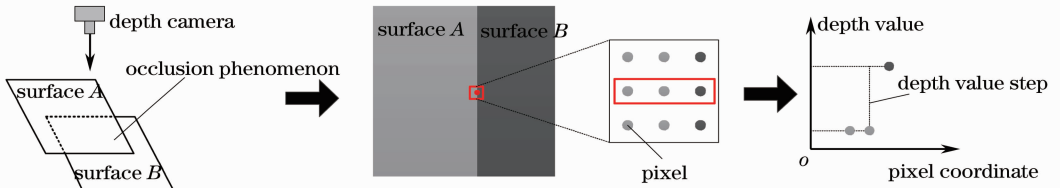


图3 深度值与遮挡现象关系示意图

Fig. 3 Sketch map between depth value and occlusion phenomenon

深度差特征包括平均深度差和最大深度差两种特征。平均深度差特征是指深度图像中任一像素点与其8邻域内各像素点深度差绝对值的平均值,而最大深度差特征是指深度图像中任一像素点与其8邻域内各像素点深度差的最大值。记深度图像 I 中任一像素点 $p(i, j)$ 的深度值为 $d(i, j)$,则其对应的平均深度差特征 $f_{ADD}(i, j)$ 和最大深度差特征 $f_{MDD}(i, j)$ 的计算公式分别定义为

$$f_{ADD}(i, j) = \frac{\sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} |d(x, y) - d(i, j)|}{8}, \quad (1)$$

$$f_{MDD}(i, j) = \max_{x, y} [d(x, y) - d(i, j)], \quad (2)$$

$$i-1 \leq x \leq i+1, j-1 \leq y \leq j+1.$$

图 4 展示了 Knot 深度图像对应的平均深度差特征及最大深度差特征,其中各特征对应的特征值均归一化到 $[0,1]$ 。由图 4 可以看出,深度图像中像

素点的平均深度差(或最大深度差)特征值越大,其越可能是遮挡边界点。

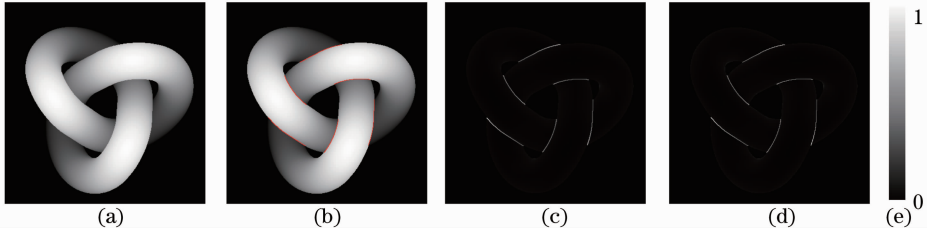


图 4 平均深度差特征和最大深度差特征可视化。(a) Knot 深度图像;(b)地面实况;(c)平均深度差特征;(d)最大深度差特征;(e)特征值标尺

Fig. 4 Visualization for average depth difference feature and maximal depth difference feature. (a) Depth image of Knot; (b) Ground Truth; (c) average depth difference feature; (d) maximal depth difference feature; (e) feature value scale

3.2 几何特征

深度图像是 2.5 维(2.5D)的图像,通过一定的方法(如反投影变换)可以求得深度图像中各像素点对应的三维空间点信息,此时,深度图像可视为三维空间中的曲面。几何特征是一类反映空间曲面弯曲程度的特征。深度图像中深度值的连续性在三维空间中则体现为其对应的空间曲面的弯曲程度,遮挡边界点对应的三维空间点处曲面的弯曲程度较大。因此,几何特征也可作为判断遮挡的重要依据。几何特征包括曲率和夹角两种特征。下面分别给出曲率特征和夹角特征的描述。

3.2.1 曲率特征

曲率特征包括平均曲率和高斯曲率。平均曲率与高斯曲率均为微分几何中反映曲面弯曲程度的内蕴几何量。曲面上一点处的主曲率为曲面在该点所有方向的法曲率中的最大值和最小值。设 κ_1 和 κ_2 为曲面上一点的两个主曲率,则该点的平均曲率为 κ_1 和 κ_2 的均值,记为 f_{MC} ,即 $f_{MC} = (\kappa_1 + \kappa_2)/2$;而该点的高斯曲率则为 κ_1 和 κ_2 的乘积,记为 f_{GC} ,即 $f_{GC} = \kappa_1 \kappa_2$ 。通过结合平均曲率和高斯曲率可以确定 8 种曲面类型^[16]:峰、脊、鞍形脊、最小面、平面、阱、谷、鞍形谷。深度图像中存在遮挡现象的局部曲面类型具有一定的规律,如图 5 给出了 Knot 深度图像中某处存在遮挡现象的局部曲面类型(此图仅列举了一种可能的情况)。这些规律可能不易通过明确的模型进行描述,但可以通过训练机器学习分类器的方式从训练数据中隐含地获得。由此可知,综合利用图像中各像素点对应的平均曲率和高斯曲率可以检测遮挡现象。此外,各曲率特征也可单独用于检测遮挡(如文献[12]利用平均曲率实现遮挡检测)。使用文献[16]中的方法计算平均曲率和高斯曲率,

该方法的优点是其不仅能够灵活地计算不同尺度下的曲率值,而且通过优化的卷积运算保证了计算的高效性。

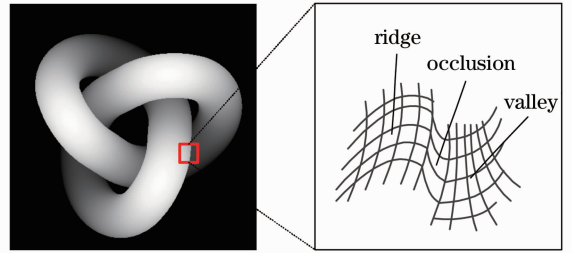


图 5 Knot 深度图像及其遮挡现象产生处的局部曲面类型

Fig. 5 Depth image of Knot and its local surface with occlusion phenomenon

3.2.2 夹角特征

图像中任一像素点对应的三维空间点与其 8 邻域内各像素点对应的三维空间点能够形成 8 个三维向量,它们分别与摄像机的观测方向(摄像机主光轴的方向向量)形成 8 个夹角,该像素点的夹角特征定义为上述 8 个夹角中的最小夹角。夹角特征也是一种反映曲面弯曲程度的特征,深度图像中任一像素点的夹角特征越小,其越有可能是遮挡边界点^[15]。记深度图像 I 中任一像素点 $p(i, j)$ 对应的三维空间点与其 8 邻域内任一像素点 $p(x, y)$ 对应的三维空间点所成的三维向量为 $\mathbf{V}_{x,y}$,摄像机的观测方向为 \mathbf{V}_c , $\mathbf{V}_{x,y}$ 与 \mathbf{V}_c 的夹角为 $\text{angle}(\mathbf{V}_{x,y}, \mathbf{V}_c)$,则 $p(i, j)$ 对应的夹角特征 $f_A(i, j)$ 的计算公式定义为

$$f_A(i, j) = \min_{x,y} \text{angle}(\mathbf{V}_{x,y}, \mathbf{V}_c), \quad i-1 \leq x \leq i+1, \quad j-1 \leq y \leq j+1, (x, y) \neq (i, j). \quad (3)$$

图 6 以某一遮挡边界点(红色像素点)为例给出了其对应的夹角特征的示意图。

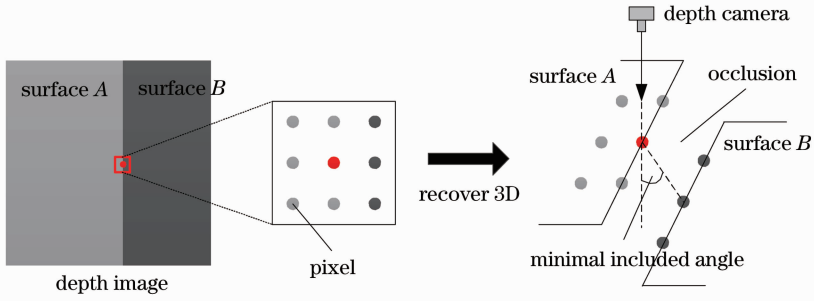


图 6 夹角特征示意图

Fig. 6 Sketch map of included angle feature

此处需要说明的是,虽然计算曲率特征和夹角特征都需要预先获得深度图像中各像素点对应的三维空间点信息,但两者所需的三维空间点类型却有所不同。其中,前者需要的是伪三维空间点,具体可直接将深度图像中各像素点的横纵坐标及其深度值分别作为其对应三维空间点的 x 轴、 y 轴和 z 轴的坐标。后者需要的则是真实三维空间点。对于一幅深度图像,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各像素点在世

界坐标系中的三维坐标。由此可见,与计算曲率特征相比,计算夹角特征所需信息较多,计算量较大。

3.3 统计特征

根据 3.1 节的描述,深度图像中深度值的连续性是判断遮挡的重要依据,图像中的遮挡边界点至少与其某一邻域点的深度值存在较大的深度值阶跃。因此,由遮挡边界点 8 邻域内的各像素点的深度值组成的一组数据应具有较大的波动。图 7 展示了遮挡边界点与其 8 邻域点的深度值变化情况,其中,红色像素点为任一遮挡边界点。

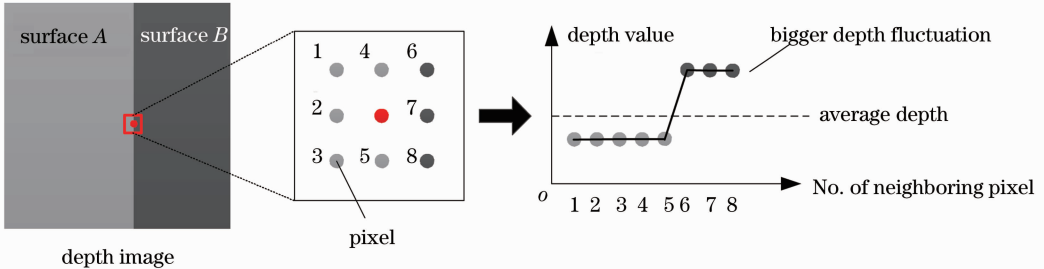


图 7 遮挡边界点与其 8 邻域点的深度值波动示意图

Fig. 7 Depth value fluctuation between occlusion boundary point and its eight neighbors

在统计学理论中,方差是能够反映一组数据波动大小的统计量。基于此,本文提出了深度值离散度特征。深度值离散度特征是指深度图像中任一像素点 8 邻域内各像素点深度值的方差。记深度图像

I 中任一像素点 $p(i, j)$ 的深度值为 $d(i, j)$, $p(i, j)$ 的 8 个邻域点深度值的平均值为 $E(i, j)$, 则其对应的深度值离散度特征 $f_{DD}(i, j)$ 的计算公式定义为

$$f_{DD}(i, j) = \frac{\sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} [d(x, y) - E(i, j)]^2 - [d(i, j) - E(i, j)]^2}{8}, \quad (4)$$

图 8 展示了 Knot 深度图像对应的深度值离散度特征,其中图像中各像素点对应的该特征值均归一化到 $[0, 1]$ 。由图 8 可以看出,深度图像中像素点的深度值离散度特征值越大,其越可能是遮挡边界点。

4 基于随机森林的遮挡检测方法

4.1 特征分析及选取

给出各遮挡相关特征的定义及计算方法后,下面对各遮挡相关特征进行分析,并根据分析结果选取合适的特征用于遮挡检测。特征分析主要涉及两

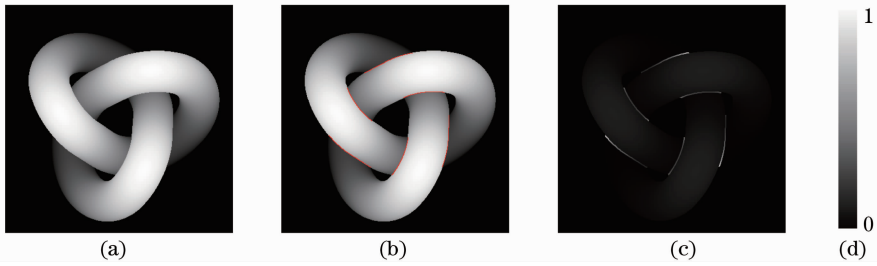


图 8 深度值离散度特征可视化。(a) Knot 深度图像;(b)地面实况;(c)深度值离散度特征;(d)特征值标尺

Fig. 8 Visualization for depth dispersion feature. (a) Depth image of Knot; (b) ground truth; (c) depth dispersion feature; (d) feature value scale

个方面的指标:特征重要性和特征提取时间。具体的特征分析方法如下所述。应用随机森林^[17]分析各遮挡相关特征的重要性,即采用所有的遮挡相关特征训练随机森林分类器,训练完成后即可得到各遮挡相关特征的重要性(此为随机森林的优点之一且计算出的各特征重要性大小关系与所选样本无依赖关系);通过计算 100 次取均值的方法得到提取一

幅 $400 \text{ pixel} \times 400 \text{ pixel}$ 的深度图像中每种遮挡相关特征所需要的时间,在此基础上,将计算出的各特征提取时间分别除以该幅深度图像中视觉目标点的个数得到提取一个视觉目标点的各遮挡相关特征的时间,并将其作为各特征最终的提取时间性能指标。通过上述分析过程,最终获得各遮挡相关特征的重要性及提取时间如图 9 所示。

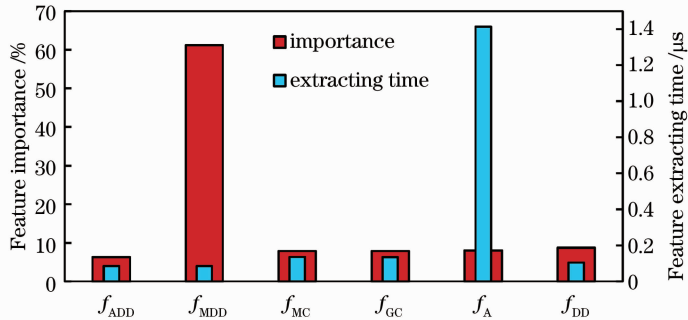


图 9 各遮挡相关特征的重要性及提取时间

Fig. 9 Importance and extracting time of all the occlusion-related features

图 9 中, f_{ADD} 为平均深度差特征, f_{MDD} 为最大深度差特征, f_{MC} 为平均曲率特征, f_{GC} 为高斯曲率特征, f_A 为夹角特征, f_{DD} 为深度值离散度特征。由图 9 可知,特征的重要性为 $f_{MDD} > f_{DD} > f_A > f_{MC} = f_{GC} > f_{ADD}$, 而特征的提取时间为 $f_{ADD} < f_{MDD} < f_{DD} < f_{GC} < f_{MC} < f_A$ 。此外,本文还兼顾了特征的通用性。基于各遮挡相关特征的定义可知,计算 f_A 时需要获得深度图像中各像素点对应的真实三维空间点信息以及摄像机的观测方向信息,而计算 f_{ADD} 、 f_{MDD} 、 f_{MC} 、 f_{GC} 和 f_{DD} 时仅需要深度图像中的深度信息,因此,与 f_{ADD} 、 f_{MDD} 、 f_{DD} 、 f_{GC} 和 f_{MC} 相比, f_A 的通用性较差。

基于上述特征分析结果,并综合考虑特征的重要性、提取时间及通用性,本文最终选取了平均深度差特征 f_{ADD} 、最大深度差特征 f_{MDD} 、平均曲率特征 f_{MC} 、高斯曲率特征 f_{GC} 、深度值离散度特征 f_{DD} 五种遮挡相关特征用于遮挡边界检测。没有选取夹角特

征的原因主要有以下两点:夹角特征的提取时间相对较长且该特征的重要性并不突出;由于夹角特征的通用性受到了限制,因此利用其设计出的遮挡检测分类器的通用性也将受到限制。至此,已确定出本文所需的全部遮挡相关特征,下面对所提遮挡检测方法进行描述。

4.2 遮挡检测方法

基于机器学习的思想,将深度图像中的遮挡检测问题视为二分类问题,并采用随机森林分类器作为遮挡检测分类器,通过检测图像中每个像素点的类别(遮挡边界点或非遮挡边界点),最终得到图像中的遮挡边界。随机森林是一种包含多个决策树的分类器,它利用 bootstrap 重抽样方法从原始样本中抽取多个样本,对每个 bootstrap 样本进行决策树建模,然后组合多棵决策树的预测,通过投票得出最终预测结果。它具有很高的预测准确率,对异常值和噪声具有很好的容忍度,且不容易出现过拟

合^[17]。在应用随机森林分类器进行遮挡检测之前,需要先对其进行训练,具体步骤如下:从训练数据集[由若干幅深度图像及其对应的地面实况(Ground Truth)组成,详见实验部分]中获取正负样本,将它们组成训练样本集,其中,正样本为遮挡边界点,负样本为非遮挡边界点;根据各遮挡相关特征的计算方法分别提取每个样本的平均深度差、最大深度差、平均曲率、高斯曲率和深度值离散度特征,并将每个样本的各特征值分别进行归一化;将每个样本的各归一化的特征值及其类别标签作为输入数据训练随机森林分类器。训练完成后,即可利用训练好的随机森林分类器检测深度图像中的遮挡边界点。下面给出遮挡检测算法的具体描述。

算法名称: Occlusion Detection Algorithm Based on RFC

Input: 深度图像 I

Output: 遮挡检测结果 I'

1) 遍历输入的深度图像 I , 分别提取各像素点的平均深度差、最大深度差、平均曲率、高斯曲率和深度值离散度五种特征, 并将获得的特征值保存到对应的特征矩阵中。

2) 应用最大最小值法对每种特征对应的特征矩阵进行归一化处理, 得到每种特征对应的归一化特征矩阵。

3) /* 检测遮挡边界 */

For $i := 1$ To Row // Row 为深度图像中像素点的总行数

For $j := 1$ To Column // Column 为深度图像中像素点的总列数

获取 I 中像素点 $p(i, j)$ 对应的所有遮挡相关特征归一化的特征值并组成特征向量 $V(i, j)$

将 $V(i, j)$ 输入到随机森林分类器(遮挡检测分类器)中得到其类别标签 $l(i, j)$

If $l(i, j) = 1$ // 遮挡边界点的类别标签为 1, 非遮挡边界点的类别标签为 0

记录 $p(i, j)$ 在图像 I 中的位置

End If

End For

End For

4) 复制图像 I 得到 I' , 根据记录的所有遮挡边界点位置将 I' 中的对应位置标记为红色。

5) 可视化 I' 得到遮挡检测结果。

5 实验及分析

5.1 实验环境及数据集

为了检验本文方法的效果,用斯图加特深度图像数据库(SRID)^[18]中的深度图像进行了实验。SRID 中深度图像的大小均为 400 pixel×400 pixel。实验硬件环境为 CPU Intel(R) Core(TM) i7-3770 3.40 GHz, 内存 8.0 G。遮挡检测程序采用 C++ 编程实现。实验过程中,随机森林的参数设置如下:单棵树的最大深度为 10,单棵树每个节点随机选择的特征数量为 2,随机森林中树的数量为 100。鉴于 SRID 中的图像数量较大且部分图像不存在遮挡现象,从中选取了 55 幅存在遮挡现象的深度图像作为实验数据集,并通过手工标记的方式得到每幅深度图像对应的 Ground Truth。手工标记依据的标准是结合深度图像中的视觉目标及其原始深度值数据,力求客观真实地反映视觉目标中的遮挡边界点。在标记过程中,采用了两步标记获取 Ground Truth 的方案。1) 对于每幅深度图像,基于最大深度差特征值利用阈值分割技术得到初始标记结果(由于最大深度差特征检测效果最佳,当设定的分割阈值较大时,可得到准确率极高的遮挡检测结果)。2) 将初始标记结果提供给 10 位标记人员分别进行手工标记,通过程序统计手工标记差异并采用投票的方式确定出最终的标记结果,即 Ground Truth。在此基础上,又从实验数据集中确定出 5 幅具有代表性的深度图像作为训练数据集,其余 50 幅作为测试数据集。具体确定过程中,要求训练数据集中的深度图像应能够尽可能多地涵盖不同情况下的遮挡边界点,而对测试数据集中的深度图像则无特殊要求。也就是说,由于设计的遮挡检测分类器判断的是深度图像中单个像素点的遮挡情况,所以其与图像中的视觉目标无依赖关系,因此,用于测试的深度图像即可以是训练图像中同一视觉目标在不同观测方向下的深度图像,也可以是其他视觉目标的深度图像,即任意一幅深度图像。图 10 展示了最终确定出的用于训练的深度图像及其 Ground Truth(即训练数据集),从左到右各视觉目标的名称分别为 Knot、Rocker、Plane、Bunny 和 Turbine。该训练数据集中遮挡边界点的总数为 3348,其中 Knot、Rocker、Plane、Bunny 和 Turbine 中遮挡边界点个数分别为 632、656、455、732 和 873 个。

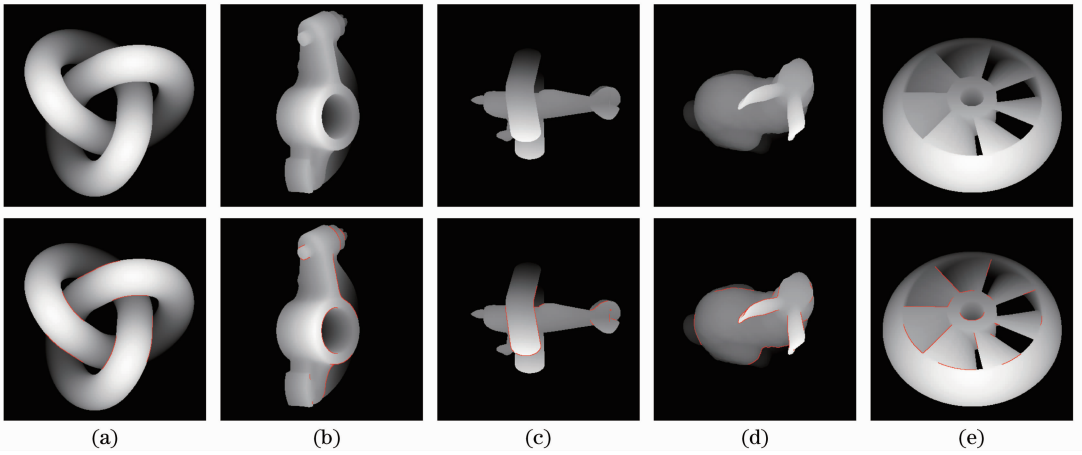


图 10 用于训练的深度图像及其地面实况。(a) Knot; (b) Rocker; (c) Plane; (d) Bunny; (e) Turbine

Fig. 10 Depth image and its Ground Truth for training. (a) Knot; (b) Rocker; (c) Plane; (d) Bunny; (e) Turbine

5.2 实验比较及分析

为了全面、合理地评估所提方法的遮挡检测效果,进行了两组对比实验:1)采用不同特征组合方式训练出的不同随机森林分类器之间的遮挡检测对比实验;2)是本文方法与其他已有方法之间的对比实验。实验中,首先选取了训练本文提出的随机森林遮挡检测分类器所需的正负样本集。在正负样本选取过程中,考虑到通过均衡的训练样本(正负样本数量相近)有利于获得性能较好的随机森林分类器,故将前述训练数据集中的全部 3348 个遮挡边界点作为正样本集,然后采用简单随机抽样的方法从前述训练数据集中选取 3452 个非遮挡边界点作为负样本集,并将它们组成训练样本集来训练随机森林。实验分析及评估时,记 n_{gt} 为 Ground Truth 中遮挡边界点的数量, n_d 为遮挡检测结果中遮挡边界点的数量, n_m 为检测结果中检测正确的遮挡边界点的数量,则查准率 $R_p = n_m/n_d$, 查全率 $R_r = n_m/n_{gt}$ 。平均查准率 R_{ap} 为测试数据集中所有深度图像对应的查准率的平均值,平均查全率 R_{ar} 为测试数据集中所有深度图像对应的查全率的平均值。下面分别给出两组对比实验的结果及分析。

5.2.1 不同随机森林分类器的比较

为了验证本文所提深度值离散度特征及所引入的高斯曲率特征的有效性,使用前述训练样本集分别基于 4 种不同的特征组合方式训练得到了 4 个随机森林分类器 RFC_1 、 RFC_2 、 RFC_3 和 RFC_4 。其中, RFC_1 对应的特征组合为 $\{f_{ADD}, f_{MDD}, f_{MC}\}$, RFC_2 对应的特征组合为 $\{f_{ADD}, f_{MDD}, f_{MC}\} \cup \{f_{GC}\}$, RFC_3 对应的特征组合为 $\{f_{ADD}, f_{MDD}, f_{MC}\} \cup \{f_{DD}\}$, RFC_4 对应的特征组合为 $\{f_{ADD}, f_{MDD}, f_{MC}\} \cup \{f_{DD}\} \cup \{f_{GC}\}$ 。在此基础上,将 RFC_1 、 RFC_2 、 RFC_3 和

RFC_4 在测试数据集上的平均查准率 R_{ap} 、平均查全率 R_{ar} 及 F_{score} 进行了比较,比较结果如表 1 所示。其中, $F_{score} = 2R_{ap}R_{ar}/(R_{ap}+R_{ar})$ 。

表 1 不同随机森林分类器的 R_{ap} 、 R_{ar} 和 F_{score}

Table 1 R_{ap} , R_{ar} and F_{score} for different random

forest classifiers

Classifier name	$R_{ap}/\%$	$R_{ar}/\%$	F_{score}
RFC_1	93.67	97.14	0.9537
RFC_2	93.01	98.05	0.9546
RFC_3	96.48	96.54	0.9651
RFC_4	96.64	97.41	0.9703

由表 1 可知, RFC_2 的平均查准率低于 RFC_1 , 但在平均查全率方面有了较大的提升,故 RFC_2 的 F_{score} 值较 RFC_1 有所提高,可见高斯曲率特征通过提高查全率改善了遮挡检测分类器的性能。 RFC_3 较 RFC_1 的 R_{ar} 略有降低,但在平均查准率方面有了较大的提升,故 RFC_3 的 F_{score} 值较 RFC_1 有了较大程度提高,可见深度值离散度特征主要通过提高查准率显著地提高了遮挡检测分类器的性能。与表 1 中 RFC_4 对应的结果进一步比较可知, RFC_4 的 R_{ar} 虽比 RFC_2 略有下降,但其 R_{ap} 却有了较为显著的提高,而其与 RFC_3 相比,在 R_{ap} 和 R_{ar} 方面均有所提高,同时 RFC_4 对应的 F_{score} 值也是四者中最高的。由此可见,通过同时结合高斯曲率特征和深度值离散度特征可进一步提升分类器的检测性能。上述比较验证了本文所提深度值离散度特征及引入的高斯曲率特征的有效性。

5.2.2 不同遮挡检测方法的比较

由于目前基于深度图像的遮挡检测成果较少,为了进一步评估本文所提遮挡检测方法的检测效果,将其与已有全部方法^[12-15]进行了比较。实验

中,根据文献[12-15]中的描述分别实现了它们对应的遮挡检测方法。其中,文献[12]方法的曲率阈值在本文中设为 0.005,深度阈值为 15.5;文献[13-14]的平均深度差阈值由程序自动计算;文献[15]所提的支持向量机遮挡检测分类器采用原文中提供的方法训练得到。本文则用前述已获得的训练样本集训练得到所提的随机森林遮挡检测分类器。在此基础上,在测试数据集上分别用不同的遮挡检测方法进行了实验。图 11 给出了部分实验对比结

果,从上到下 5 种不同复杂程度的视觉目标的名称分别为 Knot、Rocker、Plane、Duck 和 Mole。其中,图 11(a)为视觉目标的深度图像,图 11(b)为 Ground Truth,图 11(c)为文献[12]方法的遮挡检测结果,图 11(d)为文献[13]方法的遮挡检测结果,图 11(e)为文献[14]方法的遮挡检测结果,图 11(f)为文献[15]方法的遮挡检测结果,图 11(g)为本文所提方法的遮挡检测结果。

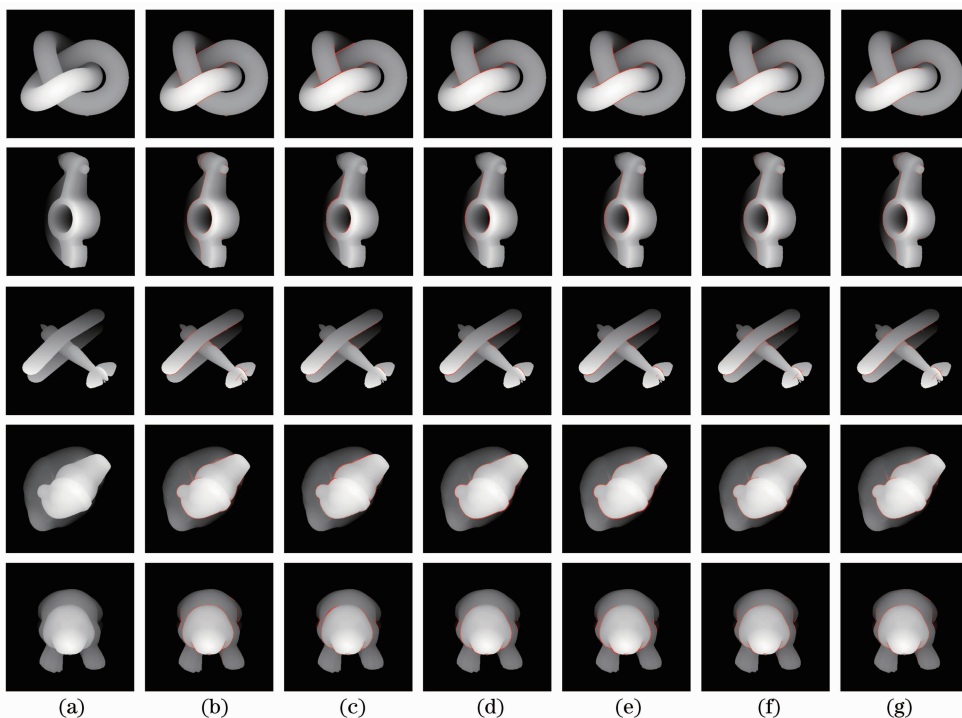


图 11 不同遮挡检测方法的实验结果比较。(a) 深度图像;(b) 地面实况;(c) 文献[12]方法;(d) 文献[13]方法;(e) 文献[14]方法;(f) 文献[15]方法;(g) 本文方法

Fig. 11 Comparison of occlusion detection results for different methods. (a) Depth image; (b) Ground Truth; (c) method in Ref. [12]; (d) method in Ref. [13]; (e) method in Ref. [14]; (f) method in Ref. [15]; (g) proposed

结合图 11 分析可知:文献[12]方法能够较为准确地检测出深度图像中比较明显的遮挡边界,但由于其仅采用了平均曲率特征作为检测遮挡的依据且阈值需人为设定,因此其准确性较差且遮挡边界的连续性不强。文献[13]和文献[14]方法尽管仅采用了平均深度差特征作为检测遮挡的依据,但由于它们能够针对不同的输入图像自动计算对应的最优阈值,具有自适应性,因此相对于文献[12]方法而言,它们在准确性和遮挡边界连续性方面均有一定程度的改善,其中文献[14]方法的准确性更好一些。文献[15]方法基于机器学习思想解决深度图像中的遮挡检测问题,并采用了夹角特征和重要性较高的最大深度差特征作为检测遮挡的依据,因此其在准确

性方面有了较大的提高且遮挡边界的连续性也较好。本文所提方法仍然基于机器学习的思想,但由于采用了优选出的遮挡相关特征,因此不仅在准确性方面较之前的方法又有了一定程度的提高,而且保证了遮挡边界较好的连续性。

为了更加合理地评估各遮挡检测方法的准确性,分别从查准率和查全率两个方面对上述 5 种遮挡检测方法进行了量化评估。图 12 给出了图 11 中遮挡检测结果对应的各遮挡检测方法的查准率和查全率。同时,为了更加全面地衡量本文方法的检测效果,表 2 给出了各遮挡检测方法针对测试数据集中所有深度图像的平均查准率和平均查全率。

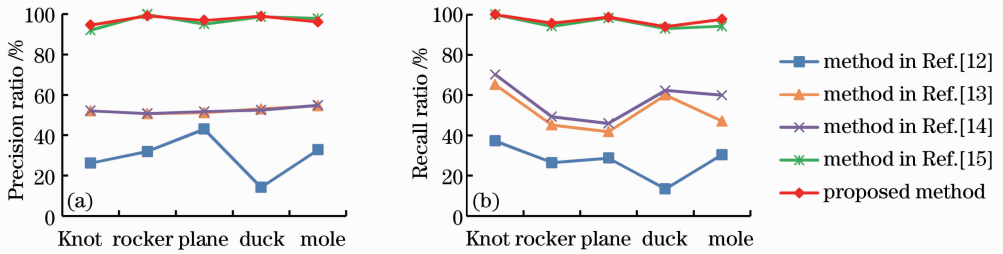


图 12 不同遮挡检测方法的查准率和查全率

Fig. 12 Precision ratio and recall ratio for different occlusion detection methods

表 2 各遮挡检测方法的 R_{ap} 和 R_{ar} Table 2 R_{ap} and R_{ar} for different occlusion detection methods

Method name	$R_{ap}/\%$	$R_{ar}/\%$
Method in Ref. [12]	27.39	26.38
Method in Ref. [13]	51.91	52.57
Method in Ref. [14]	52.28	55.98
Method in Ref. [15]	95.48	96.95
Proposed method	96.64	97.41

由图 12 和表 2 可以看出,所提方法的查准率和查全率较文献[12]、文献[13]及文献[14]方法均有较大幅度的提升,而同文献[15]方法相比虽然提升的幅度不是很大,但由于文献[15]方法的查准率和查全率已达到较高的水平(90%以上),可提升空间已比较小,因此如果按提升幅度占最大可提升幅度的比例作为评估标准进行分析,本文所提方法较文献[15]方法在查准率和查全率方面的提升幅度也是较大的。所以,本文所提方法在查准率和查全率两个方面均优于已有的方法。

此外,还对各方法检测遮挡的时间消耗进行了比较分析,主要包括特征提取和遮挡检测两个部分。

表 3 各遮挡检测方法的平均时间消耗(单位: ms)

Table 3 Average time-consumption for different occlusion detection methods (unit: ms)

Method name	Time-consumption for feature extraction	Time-consumption for occlusion detection	Total time-consumption
Method in Ref. [12]	20.94	1.25	22.19
Method in Ref. [13]	3.73	0.41	4.14
Method in Ref. [14]	3.73	0.40	4.12
Method in Ref. [15]	66.33	16.83	83.16
Proposed method	32.81	131.75	164.55

表 4 各遮挡检测方法概况

Table 4 Survey for different occlusion detection methods

Method name	Occlusion related features						Method type
	f_{ADD}	f_{MDD}	f_{MC}	f_{GC}	f_{Λ}	f_{DD}	
Method in Ref. [12]			✓				Threshold segmentation
Method in Ref. [13]	✓						Threshold segmentation
Method in Ref. [14]	✓						Threshold segmentation
Method in Ref. [15]		✓			✓		Machine learning
Proposed method	✓	✓	✓	✓		✓	Machine learning

由于各遮挡检测方法的核心是如何基于计算出的特征检测遮挡,且 4.1 节中已经对各特征的提取时间进行了分析比较,故下面主要针对各遮挡检测方法在遮挡检测部分的时间复杂度进行分析。从全局看,由于本文所提方法与文献[12—15]中的方法均以像素点为单位检测遮挡,对于一幅 $n \text{ pixel} \times n \text{ pixel}$ 的待测深度图像,各遮挡检测方法均需要遍历一遍输入图像才能够完成遮挡检测,因此各方法对应的时间复杂度均为 $o(n^2)$ 。从局部看,由于各遮挡检测方法具体检测一个像素点的遮挡情况时又有所不同,因此使得实际的耗时存在差异。综合来说,直接应用阈值检测的方法比应用分类器检测的方法耗时少,而且不同分类器检测的耗时也存在差异(主要是由于方法的原理及具体设计实现方式不同造成的)。表 3 中给出了在实验环境下不同遮挡检测方法检测一幅深度图像中遮挡现象所需要的时间信息,该时间信息为不同方法检测测试数据集中所有深度图像遮挡现象耗时信息的平均值。同时,为了便于比较分析,表 4 给出了各遮挡检测方法的概况,“✓”表示方法采用的特征。

结合表3、表4信息及4.1节给出的各特征提取时间分析可知：在特征提取阶段，由于文献[15]方法采用了耗时较长的夹角特征，故该方法的特征提取时间最长，而本文方法由于采用了较多的遮挡相关特征，故相对于文献[12-14]方法，本文方法的特征提取时间也较长。在遮挡检测阶段，由于文献[12-14]方法均直接利用阈值分割法，其计算简单、运算效率较高，因此相对耗时较少；而对于同为基于学习思想的文献[15]方法和本文方法而言，由于本文方法采用了较多的遮挡相关特征，导致其样本的特征向量维度较大，进而在生成样本特征向量以及将样本输入分类器判断样本类别时所需时间也较多，因此本文方法在遮挡检测阶段耗时最多。但是，从总体上看，所提遮挡检测方法所用的总时间仅为164.55 ms(约0.16 s)，仍具有较好的实时性。未来，将从两个方面进一步减少所提方法的耗时：1)从所提方法的流程入手，依据各遮挡相关特征的重要性和耗时选取适当特征设计一组滤波器(每个滤波器对应一个特征)，在待测样本输入基于随机森林的遮挡检测分类器之前先对其进行滤波处理，以100%查全率及尽可能高的查准率为前提，基于各滤波器对应特征的最优分割阈值设定其阈值，以有效减少输入分类器的负样本数量，从而减少样本分类耗时；2)从基于随机森林的遮挡检测分类器本身入手，研究如何能够在确保分类器准确性的前提下有效减少其中决策树的个数以及各决策树的深度，通过结合参数调优及具体实现细节优化等手段进一步减少分类器检测的耗时。

需进一步说明的是，虽然所提方法与文献[15]方法均基于机器学习思想，但由于文献[15]方法使用了通用性相对较差的夹角特征，而所提方法使用的特征均具有较好通用性，所以所提方法在通用性方面也要优于文献[15]方法。故综合考虑准确率、通用性、实时性等多方面因素，所提方法不失为一种较优的方法。

6 结 论

提出了一种利用随机森林进行遮挡检测的方法。该方法通过提取深度图像中与遮挡相关的平均深度差、最大深度差、平均曲率、高斯曲率和深度差离散度5种特征训练随机森林得到遮挡检测分类器，并利用该分类器实现对待测图像中遮挡边界的检测。主要创新点在于：1)提出一种新的遮挡相关特征(深度差离散度特征)，该特征不但重要性高而

且特征提取耗时少；同时，将高斯曲率作为特征引入遮挡检测领域。在基于深度信息的遮挡特征设计和研究上取得了突破性进展；2)以特征重要性和特征提取时间为衡量标准，对深度图像中的各遮挡相关特征进行了分析评估，其可为基于深度图像的遮挡检测方法研究提供较为全面的参考资料，便于相关研究人员结合已有的特征找准切入点和突破口进一步挖掘新特征；3)基于随机森林实现对深度图像中遮挡边界的检测，与已有方法相比，所提方法不但明显提高了遮挡检测结果的准确性，而且具有较好的通用性。

参 考 文 献

- 1 S J Lee, K R Park, J Kim. A SfM-based 3D face reconstruction method robust to self-occlusion by using a shape conversion matrix [J]. *Pattern Recognition*, 2011, 44(7): 1470-1486.
- 2 C Schmalz, B Rosenhahn, T Brox, *et al.*. Region-based pose tracking with occlusions using 3D models [J]. *Machine Vision and Applications*, 2012, 23(3): 557-577.
- 3 Wei Xinguo, Liu Tao, Liu Zhen. Three-dimensional measurement data registration based on planar circle target [J]. *Acta Optica Sinica*, 2013, 33(2): 0215002. 魏新国, 刘涛, 刘震. 基于平面圆靶标的三维数据拼接[J]. *光学学报*, 2013, 33(2): 0215002.
- 4 A Gupta, A Mittal, L S Davis. Constraint integration for efficient multiview pose estimation with self-occlusions [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(3): 493-506.
- 5 L Szirmay-Kalos, T Umenhoffer, B Tóth, *et al.*. Volumetric ambient occlusion for real-time rendering and games [J]. *IEEE Computer Graphics and Applications*, 2010, 30(1): 70-79.
- 6 J J McAuley, T S Caetano. Fast matching of large point sets under occlusions [J]. *Pattern Recognition*, 2012, 45(1): 563-569.
- 7 A Stein, M Hebert. Occlusion boundaries from motion: low-level detection and mid-level reasoning [J]. *International Journal of Computer Vision*, 2009, 82(3): 325-357.
- 8 D Hoiem, A Efros, M Hebert. Recovering occlusion boundaries from an image [J]. *International Journal of Computer Vision*, 2011, 91(3): 328-346.
- 9 A Ayvaci, M Raptis, S Soatto. Sparse occlusion detection with optical flow [J]. *International Journal of Computer Vision*, 2012, 97(3): 322-338.
- 10 Y Liu. Automatic range image registration in the Markov chain [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(1): 12-29.
- 11 I Y Jang, J H Cho, K H Lee. 3D human modeling from a single depth image dealing with self-occlusion [J]. *Multimedia Tools and Applications*, 2012, 58(1): 267-288.
- 12 Zhang Shihui, Zhang Yujie, Kong Lingfu. Self-occlusion detection approach based on depth image [J]. *Journal of Chinese Computer System*, 2010, 31(5): 964-968. 张世辉, 张煜婕, 孔令富. 一种基于深度图像的自遮挡检测方法[J]. *小型微型计算机系统*, 2010, 31(5): 964-968.
- 13 Zhang Shihui, Zhang Yujie, Kong Lingfu. Self-occlusion detection algorithm combining depth image and optimal segmentation threshold iteration [J]. *Chinese High Technology Letters*, 2010, 20(7): 754-757. 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[J]. 高技术通讯, 2010, 20(7): 754-757.
- 14 S Zhang, F Gao, L Kong. A self-occlusion detection approach based on range image of vision object [J]. ICIC Express Letters, 2011, 5(6): 2041-2046.
- 15 S Zhang, J Liu. A self-occlusion detection approach based on depth image using SVM [J]. International Journal of Advanced Robotic Systems, 2012, 9(12): 1-8.
- 16 P J Besl, R C Jain. Segmentation through variable-order surface fitting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1988, 10(2): 167-192.
- 17 L Breiman. Random forests [J]. Machine Learning, 2001, 45(1): 5-32.
- 18 G Hetzel, B Leibe, P Levi, *et al.*. 3D object recognition from range images using local feature histograms [C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, 2: II394-II399.

栏目编辑: 张浩佳