

# Kinect 下深度信息获取技术及其在三维目标识别中的应用综述

尹潘龙<sup>1</sup> 徐光柱<sup>1</sup> 雷帮军<sup>2</sup> 曹维华<sup>1</sup>

<sup>1</sup>(三峡大学计算机与信息学院 宜昌 443002)

<sup>2</sup>(三峡大学智能视觉与图像信息研究所 宜昌 443002)

**摘 要** 微软公司 2010 年推出的 Kinect 深度传感器能够同步提供场景深度和彩色信息, 其应用的一个关键领域就是目标识别。传统的目标识别大多限制在特殊的情形, 如: 手势识别、人脸识别, 而大规模的目标识别是近年来的研究趋势。通过 Kinect 得到的 RGB-D 数据集多为室内和办公环境下获取的多场景、多视角、分目标类型的数据集, 为大规模的目标识别算法设计提供了学习基础。同时, Kinect 获取的深度信息为目标识别提供了强有力的线索, 利用深度信息的识别方法较以前的方法具有无法比拟的优势, 大大地提高了识别的精度。文章首先对 Kinect 的深度获取技术做了详细介绍; 其次对现有的 3D 目标识别方法进行综述, 接着对已有的 3D 测试数据集进行分析和比较; 最后对文章进行小结以及对未来 3D 目标识别算法和 3D 测试数据集的发展趋势作了简单的阐述。

**关键词** Kinect; 3D 目标识别; RGB-D 数据集

## Review on the Technique to Obtain Depth Information Using Kinect and Its Application to Three Dimensional Object Recognition

YIN Panlong<sup>1</sup> XU Guangzhu<sup>1</sup> LEI Bangjun<sup>2</sup> CAO Weihua<sup>1</sup>

<sup>1</sup>(College of Computer Science and Information Technology, Three Gorges University, Yichang 443002, China)

<sup>2</sup>(Institute of Intelligent Vision and Image Information, Three Gorges University, Yichang 443002, China)

**Abstract** In 2010, Microsoft launched a depth camera named Kinect, which can synchronously provide the depth and RGB information of a scene. Using sensing technique, robust depth information can be achieved at real time. One key area of Kinect is object recognition. Most of the traditional object recognition techniques are limited in special situations, such as gesture recognition, face recognition. Recent trend in computer vision is large-scale recognition. We can get the RGB-D datasets from various scenes, categories, instances and viewpoints of real domestic and office environment by Kinect. The RGB-D datasets can meet the need of practical applications. The depth information provides a strong clue for object recognition. Compared with the previous methods, using depth information for object recognition has incomparable advantage, which can greatly improve the accuracy of recognition. The four major components of this paper are as follows: 1) a detailed introduction of the technique to obtain depth information using Kinect; 2) review of the existing 3D Object recognition methods; 3) an analysis and comparison of the existing 3D Datasets for testing; 4) summary of the article, and the future development trend of the 3D object recognition algorithms and the 3D datasets for testing.

**Keywords** Kinect; 3D object recognition; RGB-D dataset

**基金项目:** 湖北省教育厅基金项目(0720112090), 三峡大学楚天学者基金(KJ2012B001), 国家自然科学基金(61102155)。

**作者简介:** 尹潘龙, 硕士研究生, 研究方向为计算机视觉、图形图像处理; 徐光柱(通讯作者), 博士, 副教授, 研究方向为计算机视觉、图形图像处理、生物特征识别、模式识别、神经网络, E-mail: xgzhu@gmail.com; 雷帮军, 博士, 教授, 研究方向为计算机视觉、图像处理、模式识别; 曹维华, 硕士研究生, 研究方向为计算机视觉、图形图像处理。

## 1 引言

近年来, 由于深度传感技术取得了较大的进步, 三维深度传感器再次引发了人们的兴趣。这些深度传感技术包括: 基于激光雷达的方法<sup>[1]</sup>、飞行时间法<sup>[2]</sup> (Time-of-Flight)、基于投影纹理立体视觉的方法<sup>[3]</sup> (Projected Texture Stereo) 等。微软的 Kinect 传感器由于其价格适中, 且提供 Kinect 软件开发工具包供开发者使用, 从而得到了广泛的应用。Kinect 是目前世界上销售最快的消费电子设备, 具有广泛的应用前景。如何有效利用 Kinect 传感器得到的深度信息来解决计算机视觉中的关键问题已成为目前的研究热点, 参考文献[4-6]中借助 Kinect 传感器, 通过获取的深度信息对场景进行三维重建; 狄海进<sup>[6]</sup>利用 Kinect 同时获取的 RGB 和深度信息对手势进行检测和跟踪; 参考文献[7]中将 Kinect 传感器装在一个机器人上, 获取 RGB-D 数据, 对这些数据进行特征提取和机器学习, 然后对机器人输入一个目标名字, 让机器人从周围的环境中检测识别出输入目标物体。目前已有许多利用 Kinect 传感器获取深度信息进行 3D 目标识别的应用, 实验表明利用深度信息的 3D 目标识别系统都具有较好的鲁棒性, 并且都取得了较好的效果。

目前, 用于测试 3D 目标识别算法的大多数 3D 数据集都是在特定的场景下、单一视角、目标类别少、实例单一的情况下获取的。这些 3D 数据测试集仅使用于特定的场景, 缺乏广泛的实际应用。而利用 Kinect 传感器获取的 RGB-D 数据集<sup>[8,9]</sup>则可包含大量日常生活中常见的目标物体, 它是在室内或者办公环境下从多个视角拍摄获取的数据, 能够用于大规模的 3D 目标识别实际应用, 较以往的数据集具有更好的实用性。

本文首先将对目前流行的 Kinect 工作原理进行详细说明, 其次对目前已有的 3D 目标识别方法进行综述, 接着对现有的 3D 测试数据集进行比较, 最后给出小结以及对未来的 3D 目标识别数据测试集和识别方法的发展趋势进行展望。

## 2 Kinect 技术原理

### 2.1 Kinect

随着科技的不断进步, 在人机交互领域, 一些低成本的深度传感器陆续出现。在 2010 年 11 月, 微软在北美正式发布了一种新型的体感设备 Kinect<sup>[10,11]</sup>。

Kinect 是微软为了增加 Xbox 上游戏的可玩性, 实现自然用户交互方式而开发的一款设备。Kinect 充分利用了以色列 PrimeSense 公司的 PrimeSense 设备和感应芯片 PS1080, 以低廉的成本实现了深度摄像头功能。通过深度摄像头和 RGB 摄像头以及多阵列麦克风提供的信息, 使得 Xbox 可以感知和识别玩家及其运动, 进而操控游戏中的人物, 提高游戏的可玩性。Kinect 不仅可以玩游戏, 而且由于其提供了性价比很高的深度摄像头功能而得到了世界各地研究和开发人员关注, 因此 Kinect 被广泛地应用于各种研究领域中, 特别在计算机视觉和机器人学领域尤为活跃。

Kinect 的结构如图 1 所示, 它由一个多阵列麦克风、一个 RGB 摄像头、一个红外摄像头和一个红外发射器组成。它工作时首先通过 RGB 摄像头获取彩色信息, 利用红外摄像头感知红外线, 然后由 Kinect 内部硬件分析处理得到深度信息。Kinect 每秒可以处理 30 帧的深度信息, 每帧  $640 \times 480$  个像素, 深度探测范围 1.2~3.5 m。

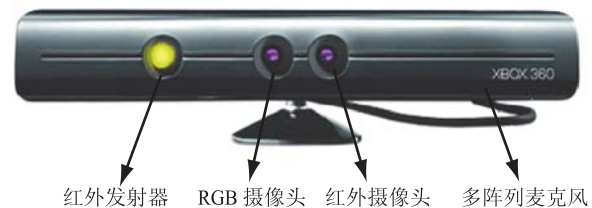


图 1 Kinect 结构图

### 2.2 Kinect 获取深度信息的原理

Kinect 传感器利用 PrimeSense 公司的 Lighting Coding (光编码) 技术来获取深度信息。Lighting Coding<sup>[11]</sup>用红外线给测量的空间进行编码, 经 CMOS 传感器读取编码的光线, 再通过芯片分析这些编码, 进行解码计算, 最后生成一幅表示深度信息的图像。图 2 示意了 Kinect 获取深度信息的原理。Light Coding 不同于传统的 TOF<sup>[2]</sup>和结构光测量技术, 它使用的是连续照明而非脉冲, 也不需要特制的感光芯片, 只需要普通的 CMOS 感光芯片, 这让方案的成

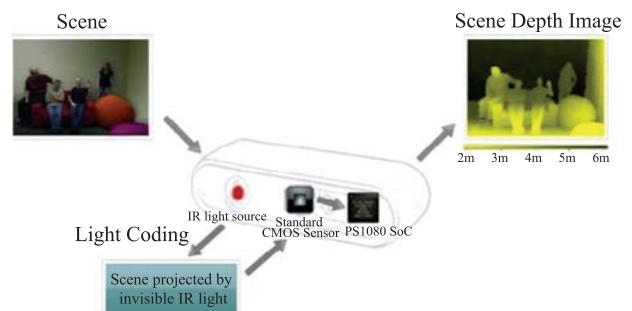


图 2 Kinect 获取深度信息的原理<sup>[12]</sup>图

本大量降低。Light Coding 就是用光源照明给需要测量的空间编上码,本质上还是结构光技术。但与传统的结构光方法不同,它发射出去的光线并不是一幅周期性变化的图像编码,而是一个具有三维纵深的“体编码”。

Lighting Coding 技术的关键是镭射光散斑(Laser Speckle),当激光照射到粗糙的物体或者穿透毛玻璃后,会形成随机的衍射斑点,即散斑。图3显示的是空间中放一本书的散斑图。散斑具有高度的随机性,而且当距离变化时会呈现不同的图案,也就是说空间中任意两点的散斑图案都是不同的。这样整个空间都加上了标记,所以当物体放进空间时,只要分析物体上面的散斑图案,就可以准确记录物体在什么位置。

Lighting Coding 发出激光对测量空间进行编码指的就是产生散斑。Kinect 利用红外线发出人们看不见的镭射光,然后透过镜头前的光栅将镭射光均匀的分布投射在测量空间中,再透过红外线接收器记录空间中的每个散斑,获取原始信息后,再通过芯片计算成具有

3D 深度的图像。

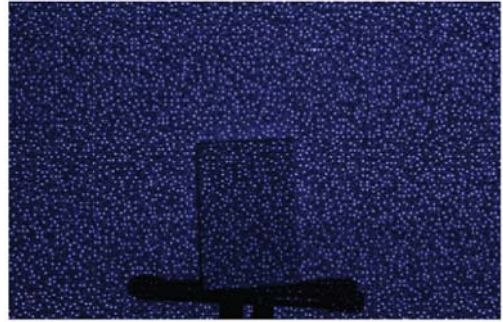


图3 空间中一本书的散斑图像<sup>[11]</sup>

### 3 3D 目标识别方法概述

3D 目标识别的核心问题就是如何利用 Kinect 传感器获取的深度信息进行目标识别,其关键就是从深度信息中提取目标的特征描述子,通过目标的特征描述子来识别物体,识别准确度依赖于所提取特征描述子的好坏。根据特征描述子的不同,总体上可以将

表1 3D 目标识别方法分析与比较

算法分类		算法描述	特点
基于 2D 图像的识别算法来实现 3D 目标识别的方法	基于深度图的 HOG 特征 <sup>[13]</sup>	用基于滑动窗口的检测方法,基于深度图的 HOG 特征用于检测器且深度信息估计物体大小,裁剪掉不可能是物体大小的目标用自定义的一个特征向量作为机器学习的输入进行学习训练,用训练的结果进行检测和识别	速度快
	机器学习 <sup>[7]</sup>		方法简单
	深度核描述子 <sup>[14]</sup>	基于核特征学习的 5 种深度核描述子,分别提供不同的目标识别线索	识别精度高,提高了系统的性能和鲁棒性
基于 3D 特征的方法	Spin Image (旋转图像) <sup>[16]</sup>	将 Spin Image 作为一个物体形状的特征描述子,来实现基于物体表面特征匹配的目标识别	在比较复杂和存在遮挡的场景下也适用,且能实现多目标的同时识别,识别系统具有很好的鲁棒性
	3D Shape Context (3D 形状上下文) <sup>[17]</sup>	基于 3D 形状上下文特征描述子的识别	在存在噪声和复杂的场景下也适用,且有较好的识别精度
	Viewpoint Feature Histogram (视点特征直方图) <sup>[18]</sup>	VFP 特征描述子用于编码目标的几何和视点信息	有较高的识别性能和快速的计算能力,能达到实时的效果,且可用于存在噪声和微笑几何形变的识别
	NARF(Normal Aligned Radial Feature) <sup>[19]</sup>	NARF 特征描述子较准确的表征了目标外部形状轮廓	为目标识别提供了强有力的线索

3D 目标识别的方法分为两类, 即基于 2D 图像的识别算法来实现 3D 目标识别的方法和基于 3D 特征的识别方法。如表 1 所示。这两类算法的共同之处, 基本上都是从获取的深度信息中提取目标的特征, 然后通过特征描述子来进行目标识别。

基于 2D 图像的识别算法来实现 3D 目标识别的方法包括基于深度图的 HOG 特征的方法、机器学习的方法和深度核描述子的方法等。这些方法都是在原有的 2D 图像识别算法的基础上, 基于深度图来提取特征进行识别。这类方法具有算法简单、速度快的特点。参考文献[13]中, 用标准的基于滑动窗口的检测, 但实验表明基于深度图的 HOG 特征检测结果较基于彩色图的 HOG 特征的差, 这是由于用 Kinect 获取的部分目标深度信息丢失了, 导致梯度特征不相称造成的。图 4 展示了分别基于彩色图和深度图的 HOG 特征检测器性能。为了提高检测性能, 利用深度信息得到物体近似大小, 在检测器中裁剪掉那些不可能是物体大小的候选目标; 同时, 用物体大小信息学习出一个用于检测的计分函数, 实验表明该方法取得了良好的效果, 提高了检测的性能。参考文献[7]中用机器学习的方法来进行检测与识别。首先分割出目标, 从中提取其特征向量, 将特征向量作为机器学习模块的输入, 进行学习训练, 最后利用训练的结果进行检测和识别。该识别方法简单高效, 但识别率比较低。参考文献[14]中, 受到基于核的特征学习<sup>[15]</sup>的启示, 该文献提出了 5 种深度核描述子(Depth Kernel Descriptors), 分别提供不同的识别线索(目标的大小、形状、边缘等), 将这些深度核描述子结合起来, 为目标识别提供了强有力的线索。表 2 显示了 5 种深度核描述子分别在类别识别(判别一个目标属于哪个类)和实例识别(识别不同的目标实例)下的识别率。实验表明, 基于深度核描述子的目标识别, 大大地提高了系统的识别精度, 提高了系统的鲁棒性。

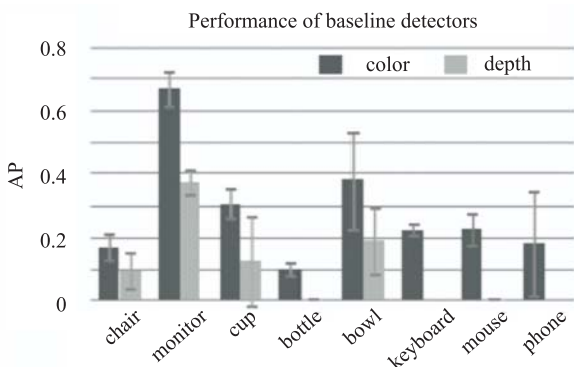


图 4 分别基于彩色图和深度图的 HOG 特征的检测器的性能分析<sup>[13]</sup>

表 2 各种深度核描述子在 RGB-D 测试数据集下识别率<sup>[14]</sup>

特征	实例识别	类别识别
大小特征	32.0	60.0±3.3
形状特征	29.5	50.2±2.9
边缘特征	28.8	64.4±3.1
梯度特征	39.8	69.0±2.3
局部二值模型特征	36.1	66.3±1.3
以上 5 种特征描述子结合	54.3	78.8±2.7

相比于基于 2D 图像的识别算法来实现 3D 目标识别的方法来说, 基于 3D 特征的识别方法主要思路是: 首先利用 Kinect 传感器得到深度信息, 再由深度信息得到 3D 点云模型, 接着从 3D 点云模型中提取 3D 特征描述子, 最后利用这些 3D 特征进行目标识别。这类方法不仅适用于复杂场景下的目标识别, 而且在存在遮挡和噪声的场景下也适用, 这类方法具有较高的识别精度和鲁棒性, 能实现多个目标的同时识别。参考文献[16]中, 提出了一种基于目标表面 3D 形状的识别方法。该方法是用 Spin Image(旋转图像)特征来进行物体表面点的匹配, 其中 Spin Image 是从 3D 点云模型中提取的特征描述子, 实验表明该算法能在比较复杂和存在遮挡的场景下实现多目标的同时识别。另外, 在 3D 计算机视觉领域中, 在有噪声和杂乱的场景下, 识别 3D 目标是一个具有挑战性的问题。参考文献[17]中, 提出了一种基于 3D 形状上下文(3D Shape Context)特征的描述子, 该方法和其他方法相比, 在存在噪声和比较杂乱的场景下具有较高的识别率。参考文献[18]中, 提出了一种新的 3D 特征描述子 VFH(Viewpoint Feature Histogram), VFH 用于编码目标的几何和视点信息。VFH 特征描述子对在目标表面存在有大量噪声和缺少深度信息的情况下, 具有很好的鲁棒性。该方法有较高的识别性能和快速的计算能力, 和其他方法相比, 识别系统可以达到实时的效果, 同时该方法可以用于存在微小几何变化的目标识别。参考文献[19]中, 也提出了一种新的感兴趣点的特征提取方法 NARF(Normal Aligned Radial Features)。NARF 特征描述子包含那些目标表面稳定且有边界信息的点, 即 NARF 能较准确地描述目标的外部形状轮廓, 为目标识别提供了强有力的线索。由此可见, 基于 3D 特征的方法能比较好地解决在复杂、存在噪声和遮挡场景下的目标识别问题。

## 4 3D 测试数据集

目前,有很多用于目标识别的 3D 测试数据集。3D 测试数据集包含目标的 RGB-D 信息,即彩色-深度信息。图 5 为一个碗的彩色图和深度图。

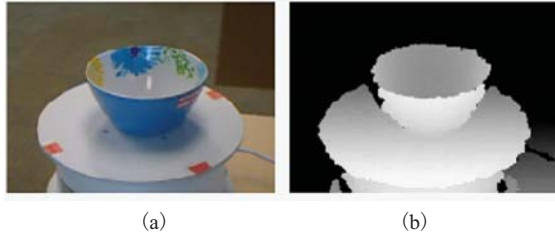


图 5 碗的彩色图(a),碗的深度图(b)<sup>[8]</sup>

为了方便进行 3D 目标识别的研究,不少科研机构提供了用于测试的数据集,目前主要使用的数据集包括 Solutions in Perception Challenge Dataset<sup>[20]</sup>、RGB-D Object Dataset<sup>[21]</sup>、Berkeley 3D Object Dataset<sup>[9]</sup>和其他用于特殊应用的 3D 数据集。下面对这些 3D 测试数据集作一个介绍,图 6 展示了以下各个 3D 测试数据集的部分数据。

(1) Solutions in Perception Challenge Dataset: 这个数据集来源于 Willow Garage,仅包含目标实例。该数据集是由在 27 个不同场景下获取的 35 个不同物体组成,如:品牌盒、家用清洁剂瓶等。

(2) RGB-D Object Dataset: 这些数据来源于国际研究机构和威斯康辛大学,包含 300 个常用的家用物体,这些物体分为 51 个类别,每个类别有 4 到 6 个实例物体。数据集是将物体放置于受控的转台上,用 Kinect 3D 摄像头以每秒 30 帧的速度同步拍摄彩色图像和深度图像,在此数据集中不存在明显的视角变化。

(3) Berkeley 3D Object Dataset: 这个数据集包含大量不同类别的物体,且每个类别都有许多不同的实例物体。这个数据集的大小不是固定不变的,而是在不断地增加,该数据集第一次发布就包含了从 75 个不同场景中获取的 849 张图片,超过 50 个具有代表性的不同类别的物体。拍摄时不是将物体放置于受控制的转台,而是在自然环境下随便拍摄,且在该数据集中,拍摄的视角和光照都有很明显的变化。这些数据均

为在室内或者办公环境下获取的,在日常生活中都很常见,因此这些数据具有很好的代表性和实用性。

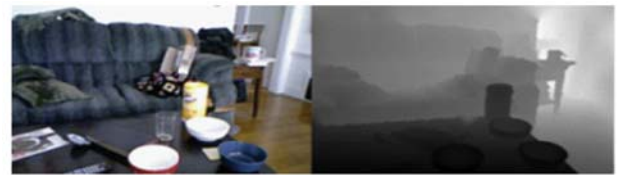
(4)其他用于特殊应用的 3D 数据集:除了上面介绍的这些数据集以外,还有一些数据集用于特殊的应用,如:自动驾驶<sup>[22]</sup>,行人检测<sup>[23]</sup>,辅助驾驶<sup>[24]</sup>等,它们的特殊性使得这些数据集不能用于大规模的目标识别应用。



(a) Solutions in Perception Challenge Dataset



(b) RGB-D Object Dataset



(c) Berkeley 3D Object Dataset

图 6 各个 3D 测试数据集的部分数据展示图

如表 3 所示,将各个 3D 测试数据集从拍摄场景、视角变化、物体类别等多个方面作了比较和分析。随着目标识别技术的不断发展和进步,在室内或者办公环境下,从多个场景、多个视角下获取的大量类别且每个类别物体实例多样化的实验 3D 测试数据集将是我們所需要的数据集。

目前,在已有的这些 3D 测试数据集的基础上测试不同 3D 目标识别算法的性能,其评价标准主要包括 4 方面的内容,即在复杂场景下算法识别率、在存在遮挡场景下算法的识别率、在有噪声的场景下算法的识别率、同时还包括算法的计算速度。通过这些评价标准可以很直观的看出不同算法之间的优缺点。

表 3 5 种 3D 测试数据集的对比

3D 数据集	场景数	物体类别	实例数	视角变化	受控拍摄
Solutions in perception Challenge	27	---	35	---	---
RGB-D Object Dataset	---	51	300	无明显的视角变化	转台
Berkeley 3D Object Dataset	$\geq 75$	$\geq 50$	---	视角变化明显	否

## 5 结束语

本文详细介绍了 Kinect 的技术原理, 并对现有的 3D 目标识别方法进行简单综述, 同时对已有的 3D 测试数据集特点进行了分析对比, 为读者提供了一个较为完整的基于 Kinect 技术的 3D 物体识别的现状分析。目前, 已经有越来越多的科研工作者意识到了 Kinect 的重要作用, 在学术界, 基于 Kinect 的研究在模式识别、计算机视觉、虚拟现实和人机交互等领域成为了近期热点的研究方向。2012 年中国计算机学会在天津举办了《学科前沿讲习班》基于 Kinect 研究与应用(本文作者亦应邀参加), 这说明国内基于 Kinect 的科学研究也正在不断发展和被认可。

在未来, 通过 Kinect 获取的 3D 测试数据集将包含从更广泛的场景下获取的、多个视角的、更多类别的、实例丰富多样化的、日常生活常见的目标实例, 可以适用于各种应用的测试。另外, 3D 目标识别的核心就是从获取的深度数据中提取特征描述子, 通过对 3D 目标识别方法的综述可以看出, 3D 数据能够提供许多强有力的目标线索, 如: 目标大小、形状、边界等, 从 3D 数据中提取各种新的 3D 特征描述子将是未来的发展趋势; 同时充分结合深度信息和 RGB 信息, 将各种特征描述子结合起来以进一步提高目标识别系统的鲁棒性和精确度也是进一步研究的热点问题。

### 参考文献

- [1] 卿慧玲, 蔡自兴. 移动机器人激光雷达数据的三维可视化方法 [J]. 吉林大学学报(信息科学版), 2004, 22(4): 411-414.
- [2] 沐光雨. 基于时间飞行深度传感器对三维物体的建模应用研究 [D]. 吉林: 吉林大学, 2011.
- [3] Konolige K, Garage W, Park M. Projected texture stereo [C] // Proceedings of IEEE International Robotics and Automation, 2010, 3: 148-155.
- [4] Lai K, Bo L, Ren X, et al. A large-scale hierarchical multi-view rgb-d object dataset [C] // IEEE International Conference on Robotics and Automation, 2011: 1817-1824.
- [5] 刘鑫, 许华荣, 胡占义. 基于 GPU 和 Kinect 的快速物体重建 [J]. 自动化学报, 2012, 38(8): 1288-1297.
- [6] 狄海进. 基于三维视觉的手势跟踪及人机交互中的应用 [D]. 南京: 南京大学, 2011.
- [7] Tian L, Putchakayala P, Wilson M. 3D Object Detection with Kinect [OL]. [http://www.cs.cornell.edu/courses/CS4758/2011sp/final\\_projects/spring\\_2011/Li\\_Putchakayala\\_Wilson.pdf](http://www.cs.cornell.edu/courses/CS4758/2011sp/final_projects/spring_2011/Li_Putchakayala_Wilson.pdf).
- [8] RGB-D Object Dataset [OL]. <http://www.cs.washington.edu/rgbd-dataset/index.html>.
- [9] B3DO: Berkeley 3-D Object Dataset [OL]. <http://kinectdata.com/>.
- [10] Kinect Introduction [OL]. <http://en.wikipedia.org/wiki/Kinect>.
- [11] Kinect 深度信息的获取原理 [OL]. <http://www.futurepict.org/?p=116>.
- [12] Kinect 高清拆机图文解析 [OL]. [http://tv.078.com/20101108/13625\\_3.shtml](http://tv.078.com/20101108/13625_3.shtml).
- [13] Janoch A, Karayev S, Jia YQ, et al. A category-level 3D object dataset: putting the kinect to work [C] // Proceedings of IEEE International Conference on Computer Vision Workshops, 2011: 1168-1174.
- [14] Bo L, Ren X, Fox D. Depth kernel descriptors for object recognition [C] // Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2011: 821-826.
- [15] Bo L, Ren X, Fox D. Kernel descriptors for visual recognition [C] // Advances in Neural Information Processing Systems, 2010.
- [16] Johnson AE, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(5): 433-449.
- [17] Frome A, Huber D, Malik J, et al. Recognizing objects in range data using regional point descriptors [C] // Proceedings of the European Conference on Computer Vision, 2004: 224-237.
- [18] Rusu RB, Bradski G, Hsu J, et al. Fast 3D recognition and pose using viewpoint feature histogram [C] // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and System, 2010: 2155-2162.
- [19] Steder B, Rusu RB, Konolige K, et al. Point feature extraction on 3D range scans taking into account object boundaries [C] // Proceedings of IEEE International Conference on Robotics and Automation, 2011: 2601-2608.
- [20] Solution in perception challenge [OL]. [http://opencv.willowgarage.com/wiki/http%3A/opencv.willowgarage.com/wiki/Solutions In Perception Challenge](http://opencv.willowgarage.com/wiki/http%3A/opencv.willowgarage.com/wiki/Solutions%20In%20Perception%20Challenge).
- [21] Lai K, Bo L, Ren X, et al. A large-scale hierarchical multi-view rgb-d object dataset [C] // Proceedings of IEEE International Conference on Robotics and Automation, 2011: 1817-1824.
- [22] Ford campus vision and lidar dataset [OL]. <http://robots.engin.umich.edu/Downloads>.
- [23] Walk S, Schindler K, Schiele B. Disparity statistics for pedestrian detection: combining appearance, motion and stereo [C] // Proceedings of European Conference on Computer Vision, 2010: 182-195.
- [24] Ess A, Leibe B, Gool LV, et al. Object detection and tracking for autonomous navigation in dynamic environments [J]. The International Journal on Robotics Research, 2010: 1707-1725.
- [25] Microsoft Kinect [OL]. <http://www.xbox.com/en-us/kinect>.
- [26] 李国镇. 基于 Kinect 的三维重建方法的研究与实现 [D]. 北京: 北京交通大学, 2012.
- [27] UBC Robot Vision Survey [OL]. <http://www.cs.ubc.ca/labs/lci/vrs/index.html>.
- [28] Helmer S, Little JJ, Lowe DG, et al. Multiple viewpoint recognition and localization [C] // Proceedings of Asian Conference on Computer Vision, 2010: 464-477.
- [29] Sun M, Xu BX, Bradski G, et al. Depth encoded hough voting for joint object detection and shape recovery [C] // Proceedings of European Conference on Computer Vision, 2010: 658-671.
- [30] Rusu RB, Blodow N, Beetz M. Fast point feature histograms (fpfh) for 3D registration [C] // Proceedings of IEEE International Conference on Robotics and Automation, 2009: 3212-3217.
- [31] Lai K, Bo L, Ren X, et al. A large-scale hierarchical multi-view RGB-D object dataset [C] // Proceedings of IEEE International Conference on Robotics and Automation, 2011: 1817-1824.