Benchmark Datasets for 3D Computer Vision

Yulan Guo, Jun Zhang, Min Lu, Jianwei Wan, Yanxin Ma
College of Electronic Science and Engineering
National University of Defense Technology
Changsha, Hunan, P. R. China 410073
yulan.guo@nudt.edu.cn

***Abstract*—With the rapid development of range image acquisition techniques, 3D computer vision has became a popular research area. It has numerous applications in various domains including robotics, biometrics, remote sensing, entertainment, civil construction, and medical treatment. Recently, a large number of algorithms have been proposed to address specific problems in the area of 3D computer vision. Meanwhile, several benchmark datasets have also been released to stimulate the research in this area. The availability of benchmark datasets plays an significant role in the process of technological progress. In this paper, we first introduce several major 3D acquisition techniques. We also present an overview on various popular topics in 3D computer vision including 3D object modeling, 3D model retrieval, 3D object recognition, 3D face recognition, RGB-D vision, and 3D remote sensing. Moreover, we present a contemporary summary of the existing benchmark datasets in 3D computer vision. This paper can therefore, serve as a handbook for those who are working in the related areas.**

I. INTRODUCTION

Computer vision in 2D images (including color and grey images) has been extensively investigated and a number of significant advances have been achieved [22]. However, 2D images cannot provide accurate geometrical (depth) information of a scene. They are also very sensitive to variations in scale, rotation and illumination [11]. In contrast, range images have the potential to overcome these limitations faced by 2D images. 3D computer vision in range images has became increasing popular due to the rapid development of 3D scanners and computing devices in the past few decades [2]. Moreover, the advent of low-cost commercial 3D scanners (e.g., Microsoft Kinect and Asus Xtion) and the opensource Point Cloud Library (PCL) has greatly boosted the research in 3D computer vision in recent years [2].

3D computer vision has a wide range of applications in many areas [14], [18], [10]. Some typical applications are listed below:

- Robotics: 3D object detection, 3D object recognition, scene perception, Simultaneous Localization and Mapping (SLAM), and navigation.
- Biometrics: face recognition, ear recognition, hand recognition, and face expression recognition.
- Remote Sensing: LiDAR survey and mapping, digital elevation model generation, change detection, object (e.g., tree, car) detection/recognition, and canopy estimation.
- Entertainment: human-machine interaction, 3D movie, 3D game, virtual 3D world, and augmented reality.
- Construction: as-built drawing, earthwork volume estimation, automatic construction work monitoring, and quality assessment.
- Medical Treatment: orthodontic diagnosis and treatment, maxillofacial surgery, orthopaedic and custom footwear applications, and assessment of oro-facial malformations.

A large number of algorithms have been proposed to address different technological aspects of these applications [1], [16], [5], [27], [7]. Meanwhile, a large body of datasets have been released to facilitate the research in 3D computer vision. Most of these algorithms and datasets in the literature focused on some specific research topics in 3D computer vision. Although different applications have several particular research issues which are different from the others, they also share a considerable number of techniques in common. That is, many ideas used in one research area can be modified or updated to another research area (e.g., 3D model retrieval *vs* 3D face/object recognition). A rich amount of research progress can already be found in each individual topic of 3D computer vision. It therefore, becomes very necessary to investigate the similarities and differences between the techniques used in different topics of 3D computer vision. On that basis, this paper presents a contemporary review on various topics in 3D computer vision.

The selection of benchmark datasets is the key to a rigorous evaluation of the state-of-the-art algorithms. Since the number of datasets in 3D computer vision has increased significantly, a timely review is in great demand. On that basis, this paper comprehensively compiles all major benchmark datasets in 3D computer vision with analysis on their characteristics. We also provide the link to each dataset, which can greatly benefit the researchers in related research areas.

The rest of this paper is organized as follows. Section II gives a review on 3D acquisition techniques. Sections III, IV, V, VI, and VII, VIII presents an overview on techniques and benchmark datasets for 3D object modeling, 3D model retrieval, 3D object recognition, 3D face recognition, RGB-D vision, and 3D remote sensing, respectively. Section IX concludes this paper.

II. 3D ACQUISITION TECHNIQUES

A number of techniques have been developed to acquire the 3D shape of an object, which can broadly classified into two categories: contact and non-contact 3D scanners. Several major 3D acquisition techniques are described in this section.

978-1-4799-4315-9/14/$31.00

(a) Faro Arm (b) Optech ILRIS-LR (c) Vivid 910

(d) Microsoft Kinect (e) Bumblebee XB3

Fig. 1: An illustrations of several 3D Scanners.

A. Contact 3D Scanners

A contact scanner acquires the 3D coordinates of the surface of an object through physical touch. The Faro Arm (as shown in Fig. 1 (a)) is an example of these contact scanners, it is commonly used the area of manufacturing. These scanners are very precise. However, they require contact with the objects being scanned and they also work relatively slowly compared to other types of scanners.

B. Non-Contact 3D Scanners

Non-contact 3D scanners can further classified into active scanners and passive scanners [8].

1) Active Scanners:

- Time-of-flight: the scanner emits a laser pulse and then counts the round-trip time of the pulse. The distance between a surface and the sensor is calculated by multiplying the half of the round-trip time by the speed of light. The Light Detection and Ranging (LiDAR) scanner (as shown in Fig. 1 (b)) is a typical example of this technique. It can operate over very long distance (e.g., several kilometers). It is widely used to scan large structures such as buildings, rock formations, and forests. Its accuracy is however, relatively low.
- Triangulation: the scanner first shines a laser spot on the surface of an object. A camera is then used to record an image of the spot. Once the center pixel of the spot is calculated, the location of the laser spot is finally determined by the triangle formed by the laser spot, the camera and the laser emitter. In order to scan the surface of an object efficiently, one approach is to scan the light spot over the whole surface of the object using mirrors, another approach is to use a plane rather than a beam of laser light (as shown in Fig. 2). Although triangulation has a limited range of several meters, its accuracy is relatively high. The popular Konica Minolta Vivid 910 (as shown in Fig. 1 (c)) and Cyberware 3030 are two examples of this technique.
- Structured Light: the scanner first projects a pattern of light onto the surface of an object using an LCD projector or other light source (as shown in Fig. 3). It then measures the deformations of the pattern on the surface by a camera. The distance of every point in the field of view is calculated based on the pattern deformation. A structured light scanner can scan multiple points at one time. Patterns of parallel stripes are frequently used in the literature. Microsoft Kinect camera (as shown in Fig. 1(d)) is the first consumer-grade product which uses a pattern of projected infrared points to generate a dense range image.

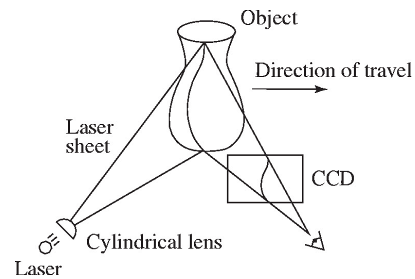


Fig. 2: An illustration of the triangulation technique.

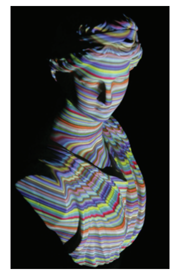


Fig. 3: A pattern of the structured light technique.

2) Passive Scanners:

- Stereoscopic: the scanner uses two cameras (which are slightly apart from each other) to look at the same surface of an object. The distance at each point in the images is determined by comparing the information of the two images. This process is quite similar to human binocular vision. Bumblebee XB3 (as shown in Fig. 1 (e)) is an example of this technique.

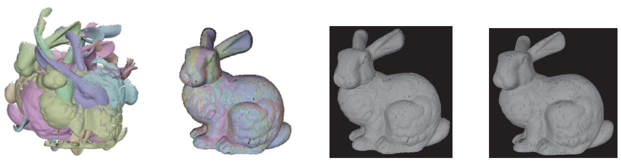
Other related techniques include conoscopic holography, modulated light, photometric stereo, and silhouette.

III. 3D Object Modeling

The task of a 3D object modeling system is to register and integrate a set of range images of an object that are acquired from different viewpoints to generate a single complete 3D model [12], [15]. A 3D modeling system usually consists of five modules: range image acquisition, correspondence generation, range image registration, integration, and surface reconstruction [20]. An illustration of the 3D modeling process is shown in Fig. 4. Range image registration plays an important role in the whole framework. It includes both coarse registration and fine registration [13]. Coarse registration can be achieved manually using a calibrated scanner and turntable (or

TABLE I: Datasets for 3D object modeling.

No.	Name and Reference	Data Type	Scanner	#Objects	Texture	Link
1	Stanford 3D Scanning Repository	Mesh	Cyberware 3030 MS, XYZ RGB	9	No	http://graphics.stanford.edu/data/3Dscanrep/
2	Stuttgart	Mesh	Synthetic	45	No	http://range.informatik.uni-stuttgart.de/
3	UWA	Mesh	Minolta Vivid 910	4	No	http://www.csse.uwa.edu.au/~ajmal/3Dmodeling.html
4	Bologna Space Time Views	Mesh	Space Time	2	Yes	http://vision.deis.unibo.it/research/78-cvlab/80-shot
5	Bologna Kinect Views	Mesh	Kinect	2	Yes	http://vision.deis.unibo.it/research/78-cvlab/80-shot



(a) Range Images (b) Registration (c) Integration (d) Reconstruction

Fig. 4: An iilustration of the 3D modeling process.

markers). It can also be achieved automatically based on the matching of local surface features. Fine registration is usually accomplished by the Iterative Closest Points (ICP) algorithm and its variants [4]. A review on automatic correspondence is given in [20], a survey and evaluation on range image registration methods in presented in [25]. The benchmark datasets for 3D object modeling are shown in Table I.

IV. 3D Model Retrieval

Given a query model, the task of a 3D model retrieval system is to search for similar models in a database using shape properties of the 3D models. An illustration of the 3D model retrieval process is shown in Fig. 5. A 3D model retrieval system usually consists of a database with an index structure and an online query engine [27]. During the phase of offline processing, each 3D model is represented with a global shape descriptor or multiple local shape descriptors. An indexing data structure is then used to organize these descriptors to enable efficient search. During the phase of online query, the descriptor of the query 3D model is calculated and then matched against these descriptors in the database. The models that are similar to the query model are retrieved. Two surveys on 3D model retrieval methods can be found in [27] and [6]. The benchmark datasets for 3D model retrieval are shown in Table II.

V. 3D Object Recognition

The task of 3D object recognition is to correctly identify the objects that are present in a scene and recover their poses (i.e., position and orientation) [11]. An illustration of the 3D object recognition process is shown in Fig. 6. Existing 3D object recognition methods can be divided into global feature based and local feature based methods. The global feature based methods process the object as a whole for recognition, and require a priori segmentation of the object from the scene. The local feature based methods usually consist of three main phases: 3D keypoint detection, local feature description, and surface matching. Several surveys on 3D object recognition

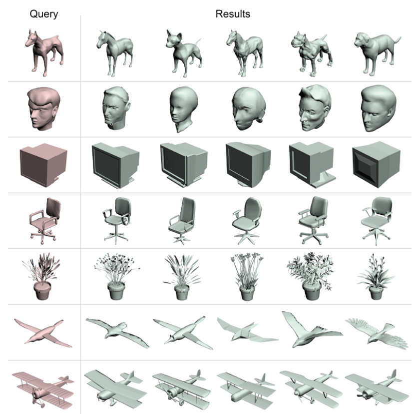


Fig. 5: An illustration of the 3D model retrieval process (Figure originally shown in [24]).

(a) Input Scene (b) Recognition Result

Fig. 6: An illustration of the 3D object recognition process.

methods can be found in [7], [19], [21], [9]. The benchmark datasets for 3D object recognition are shown in Table III.

VI. 3D Face Recognition

3D face recognition includes two application scenarios: face identification and face verification [5]. A 3D face recognition system usually consists of offline processing and online recognition phases. During the offline processing phase, 3D faces of known persons are initially enrolled into the system. This results in a gallery which contains a set of 3D faces (two sample faces are shown in Fig. 7). During the phase of online recognition, 3D faces (i.e., probes) of the gallery or other persons are match against the 3D faces in the gallery. In an identification scenario, a probe is matched against all faces in the gallery to find the most similar face. In a verification scenario, the probe is matched against the face with a claimed identity in the gallery. The claimed identity is considered to be verified if the quality of match is above a particular threshold. Several surveys on 3D face recognition methods can be found

TABLE II: Datasets for 3D model retrieval.

No.	Name and Reference	Description	Link
1	MPEG-7 database	1300 models; represented in VRML2.0 format;	
2	NTU 3D Model Database	10,911 models	http://3d.csie.ntu.edu.tw
3	Princeton Shape Benchmark (PSB)	Acquired from the World Wide Web; 1814 models in 161 classes; represented in .off format; no color/texture.	http://shape.cs.princeton.edu/benchmark/
4	Engineering Shape Benchmark (ESB)	3D shapes of mechanical parts; 801 models in 42 classes; represented in .stl and .obj formats.	http://purdue.edu/shapelab
5	Non-Rigid World	3D nonrigid shapes; 148 models; typical vertex count is about 3000; represented with MATLAB (.mat) and ASCII text files; no color/texture.	http://tosca.cs.technion.ac.il/book/resources_data.html
6	NIST Shape Benchmark (NSB)	800 models in 40 classes; represented in .off format; no color/texture.	http://www.itl.nist.gov/iad/vug/sharp/benchmark
7	TOSCA High-Resolution	High resolution 3D nonrigid shapes; 80 models; typical vertex count is about 50000; represented with MATLAB (.mat) and ASCII text files; no color/texture.	http://tosca.cs.technion.ac.il/book/resources_data.html
8	McGill 3D Shape Benchmark (MSB)	3D articulated objects; 255 models in 10 classes; represented in .im and .ply formats.	http://www.cim.mcgill.ca/~shape/benchMark/
9	3D Architecture Shape Benchmark (ASB)	3D architectural objects; 2257 models in 180 classes; represented in .off format.	
10	Toyohashi Shape	10000 models in 352 classes; represented with .off files; collected from NTU 3D Model Database and SHREC'10; no color/texture.	http://www.kde.cs.tut.ac.jp/benchmark/tsb/
11	SHREC'07	Includes several tracks in watertight models, partial matching, protein models, CAD models, relevance feedback, similarity measures, and 3D face models.	http://www.aimatshape.net/event/SHREC/shrec2007
12	SHREC'08	Includes several tracks in stability on watertight models, classification of watertight models, CAD models, generic 3D models, 3D face models.	http://www.aimatshape.net/event/SHREC/shrec2008
13	SHREC'09	Includes several tracks in generic retrieval on new benchmark, partial shape retrieval, 3D retrieval using machine learning, structural shape retrieval.	http://www.aimatshape.net/event/SHREC/shrec2009
14	SHREC'10	Includes several tracks in range scans, non-rigid shapes, generic 3D warehouse, protein models, correspondences, feature detection and description, robustness, face scans, large scale retrieval, architectural models.	http://www.aimatshape.net/event/SHREC/shrec2010
15	SHREC'11	Includes several tracks in range scans, non-rigid shapes, generic 3D warehouse, protein models, correspondences, feature detection and description, robustness, face scans, large scale retrieval, architectural models.	http://www.aimatshape.net/event/SHREC/shrec2011
16	SHREC'12	Includes several tracks in 3D mesh segmentation, stability on abstract shapes, sketch-based 3D shape retrieval, generic 3D model retrieval.	http://www.aimatshape.net/event/SHREC/shrec2012

TABLE III: Datasets for 3D object recognition.

No.	Name and Reference	Data Type	Scanner	#Objects	#Views	Link
1	OSU	Pointcloud	Minolta Vivid	NA	> 1000	http://sampl.eng.ohio-state.edu/~sampl/database.htm
2	Stuttgart	Mesh	Synthetic	45	11610	http://range.informatik.uni-stuttgart.de/
3	UWA	Mesh	Minolta Vivid	5	50	http://www.csse.uwa.edu.au/~ajmal/recognition.html
4	Bologna Stanford	Mesh	Synthetic	6	45	http://vision.deis.unibo.it/research/78-cvlab/80-shot
5	Bologna Spacetime Stereo	Mesh	Spacetime	8	15	http://vision.deis.unibo.it/research/78-cvlab/80-shot
6	Bologna Spacetime Stereo Texture	Mesh	Spacetime	8	16	http://vision.deis.unibo.it/research/78-cvlab/80-shot
7	Bologna Kinect	Mesh	Kinect	7	17	http://vision.deis.unibo.it/research/78-cvlab/80-shot
8	Washington Urban Scenes	Pointcloud		10	274	http://homes.cs.washington.edu/~kevinlai/datasets.html
9	UBC VRS	Pointcloud	-	603	60	http://www.cs.ubc.ca/labs/lci/vrs/index.html
10	3D Point Cloud People	Pointcloud	Velodyne HDL 64E S2		1400	http://www.informatik.uni-freiburg.de/~spinello/pcloud-dataset.html
11	Queen's LIDAR	Pointcloud	NextEngine	5	80	http://rcvlab.ece.queensu.ca/~qridb/
12	Queen's Stereo	Pointcloud	Stereo	5	365	http://rcvlab.ece.queensu.ca/~qridb/
13	Berkeley 3D Object	Depth Image	Kinect	50	850	http://kinectdata.com/
14	TUWIEN Kinect	Pointcloud	Kinect	35	50	
15	Ca' Foscari Venezia	Mesh	Synthetic	20	150	http://www.dsi.unive.it/~rodola/data.html
16	Bologna 3D Keypoints	Mesh				http://vision.deis.unibo.it/keypoints3d/?page_id=2

in [26], [5], [1], [16]. The benchmark datasets for 3D object recognition are shown in Table IV.

VII. RGB-D Vision

With the rapid development of low-cost commercial RGB-D cameras (e.g., Microsoft Kinect), RGB-D vision has became an extremely popular research area in recent years. A number of algorithms have been proposed in various research topics including scene mapping, object recognition, gesture recognition, face recognition, activity recognition, scene labelling, person tracking, and SLAM. Meanwhile, lots of RGB-D datasets have been made publicly available (as listed in Table V). Some example objects in the RGB-D Object Dataset [17]

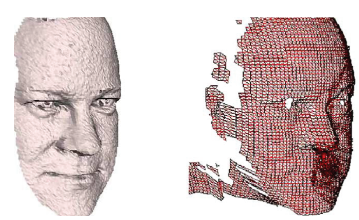


Fig. 7: Two 3D faces rendered as a shaded model and a wireframe, respectively (Figure originally shown in [5]).

TABLE IV: Datasets for 3D face recognition.

No.	Name and Reference	Scanner	#Subjects	#Views	Link
1	XM2VTS	Stereo	295	295	http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/
2	3D_RMA	Structured Light	120	720	http://www.sic.rma.ac.be/˜beumier/DB/3d_rma.html
3	GavabDB	Minolta VI-700	61	427	http://gavab.escet.urjc.es/recursos_en.html
4	University of York 1	3D Camera	97	970	http://www-users.cs.york.ac.uk/˜nep/research/3Dface/tomh/3DFaceDatabase.html
5	FRGC	Minolta Vivid	466	4950	http://www.nist.gov/itl/iad/ig/frgc.cfm
6	BU-3DFE	3DMD digitizer	100	2500	http://www.cs.binghamton.edu/˜lijun/Research/3DFE/3DFE_Analysis.html
7	Spacetime Faces	Spacetime Stereo			http://grail.cs.washington.edu/projects/stfaces/
8	BU-4DFE	Di3D	101	606	http://www.cs.binghamton.edu/˜lijun/Research/3DFE/3DFE_Analysis.html
9	ETH Face Pose	Stereo Enhanced Structured-Light	20	10545	http://www.vision.ee.ethz.ch/datasets/
10	Bosphorus	Inspeck Mega Capturor II 3D	105	4666	http://bosphorus.ee.boun.edu.tr/default.aspx
11	University of York 2	3D Camera	$>$350	$>$5000	http://www-users.cs.york.ac.uk/˜nep/research/3Dface/tomh/3DFaceDatabase.html
12	UHDB11	3dMD (TM) 2	23	1602	http://cbl.uh.edu/URxD/datasets/,2011
13	BIWI 3D Audiovisual	Stereo	14	1109	http://www.vision.ee.ethz.ch/datasets/
14	3D TEC	Minolta Vivid	214	428	http://www3.nd.edu/˜cvrl/CVRL/Data_Sets.html
15	Biwi Kinect Head Pose	Stereo	20	$>$15000	http://www.vision.ee.ethz.ch/datasets/
16	Human Face	Range Camera	1	15	http://tosca.cs.technion.ac.il/book/resources_data.html
17	NTU-CSP	Minolta Vivid	80	1280	http://eeeweba.ntu.edu.sg/csp-3dfdb/

TABLE V: Datasets for RGB-D vision.

No.	Name and Reference	Description	Application	Link
1	RGB-D Dataset 7-Scenes	Several sequences of 7 scenes, 640x480 resolution, 'ground truth' camera tracks and a dense 3D model	Dense tracking, mapping, relocalization	http://research.microsoft.com/en-us/projects/7-scenes/
2	Cornell-RGBD	24 labeled office scene and 28 home scene pointclouds, in pcd format	Semantic labeling	http://pr.cs.cornell.edu/sceneunderstanding/data/data.php
3	RGB-D Person Re-identification	4 different groups of data obtained by recording 79 people	Person re-identification, video surveillance	http://www.iit.it/en/datasets-and-code/datasets/rgbdid.html
4	RGB-D Object	300 common household objects in 51 categories, 3 video sequences for each object	Object instance/category recognition	http://www.cs.washington.edu/rgbd-dataset/
5	RGB-D Scenes	8 annotated video sequences of natural scenes	Object detection/labeling	http://www.cs.washington.edu/rgbd-dataset/
6	EURECOM Kinect Face	Multimodal facial images of 52 people, different facial expressions, lighting and occlusions	Face Recognition, face biometrics	http://rgb-d.eKinecturecom.fr/
7	Sheffield KInect Gesture (SKIG)	2160 hand gesture sequences collected from 6 subjects, 10 categories of hand gestures, 3 different backgrounds, 2 illumination conditions	Hand gesture recognition	http://lshao.staff.shef.ac.uk/data/SheffieldKinectGesture.htm
8	WorkoutSU-10 Exercise	1500 sequences of workout exercise actions represented by 3D positions of skeletal joints, 15 people performing 10 different exercises	Action recognition	http://vpa2.sabanciuniv.edu/databases/WorkoutSU-10/
9	NYU Depth V2	video sequences of 464 different indoor scenes, 26 scene types, 407024 unlabeled frames, 1449 densely labeled frames, object and instance labels	Scene segmentation	http://cs.nyu.edu/˜silberman/datasets/nyu_depth_v2.html
10	NYU Depth V1	64 different indoor scenes, 7 scene types, 108617 unlabeled frames, 2347 densely labeled frames	Scene segmentation	http://cs.nyu.edu/˜silberman/datasets/nyu_depth_v1.html
11	Cornell Activity CAD-60	60 video sequences of humans performing activities, 5 different environments, 12 activities, tracked skeletons	Activity detection, recognition, and anticipation	http://pr.cs.cornell.edu/humanactivities/data.php
12	Cornell Activity CAD-120	120 videos of long daily activities, 10 high-level activities, 10 sub-activity labels, 12 object affordance labels, tracked skeletons	Activity detection, recognition, and anticipation	http://pr.cs.cornell.edu/humanactivities/data.php
13	RGBDSLAM	9 sequences, groundtruth camera trajectory	SLAM	http://openslam.org/rgbdslam.html
14	Berkeley 3-D Object (B3DO)	849 images taken in 75 different scenes (with domestic and office settings), over 50 object classes	Category-level object recognition and localization	http://kinectdata.com/
15	RGB-D People	3000+ frames acquired in a university hall, mostly upright walking and standing persons	People detection and tracking	http://www.informatik.uni-freiburg.de/˜spinello/RGBD-dataset.html
16	BigBIRD	3600 Kinect-style RGB-D images, 600 high-resolution images	3D reconstruction, object recognition	http://rll.eecs.berkeley.edu/

TABLE VI: Datasets for 3D remote sensing.

No.	Name and Reference	Data Type	Scanner	Application	Link
1	Vaihingen/Enz Dataset	Pointcloud	Leica ALS50	Urban classification and 3D building reconstruction	http://www2.isprs.org/commissions/comm3/wg4/tests.html
2	Toronto Dataset	Pointcloud	Optech ALTM-ORION M	Urban classification and 3D building reconstruction	http://www2.isprs.org/commissions/comm3/wg4/tests.html
3	Bologna 3D Urban Scenes	Pointcloud	LIDAR	3D urban scenes	http://vision.deis.unibo.it/fede/3Dsegm.html
4	Ohio Statewide Imagery	Pointcloud	LIDAR	GIS	http://ogrip.oit.ohio.gov/ProjectsInitiatives/StatewideImagery.aspx
5	Oakland 3-D Point Cloud	Pointcloud	SICK LMS	Urban classification	http://www.cs.cmu.edu/˜vmr/datasets/oakland_3d/cvpr09/doc/

Fig. 8: Example objects in the RGB-D Object Dataset [17].

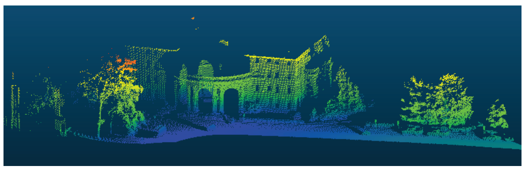


Fig. 9: A scene in the Oakland 3-D Point Cloud Dataset [23].

are shown in Fig. 8. A survey on the existing RGB-D datasets can also be found in [3].

VIII. 3D Remote Sensing

With the development of LiDAR scanners, an increasingly number of researchers work in the area of 3D remote sensing. The pointclouds used in this area were commonly acquired with a vehicle-borne or an airborne laser scanner. The research topics in this area include 3D building extraction, 3D building reconstruction, road extraction, urban object (e.g., buildings, trees, roads, cars) detection. An example pointcloud of urban landscape acquired with an optech LiDAR scanner is shown in Fig. 9. The benchmark datasets for 3D object modeling are listed in Table VI.

IX. Conclusion

In this paper, we presented several popular research topics in 3D computer vision including 3D object modeling, 3D model retrieval, 3D object recognition, 3D face recognition, RGB-D vision, and 3D remote sensing. The basic concept and the taxonomy of existing algorithms in each topic were discussed. The benchmark datasets for each research topic were also listed with their characteristics. With the flourishment of benchmark datasets in 3D computer vision, it is believed that the research in this area will be boosted in the next years.

References

[1] A.F. Abate, M. Nappi, D. Riccio, and G. Sabatino. 2D and 3D face recognition: A survey. *Pattern Recognition Letters*, 28(14):1885–1906, 2007.

[2] A. Aldoma, Z. Marton, F. Tombari, W. Wohlkinger, C. Potthast, B. Zeisl, R. Rusu, S. Gedikli, and M. Vincze. Tutorial: Point cloud library: Three-dimensional object recognition and 6 DOF pose estimation. *IEEE Robotics & Automation Magazine*, 19(3):80–91, 2012.

[3] K. Berger. The role of RGB-D benchmark datasets: an overview. *arXiv preprint arXiv:1310.2053*, 2013.

[4] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[5] K.W. Bowyer, K. Chang, and P. Flynn. A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. *Computer Vision and Image Understanding*, 101(1):1–15, 2006.

[6] B. Bustos, D.A. Keim, D. Saupe, T. Schreck, and D.V. Vranić. Feature-based similarity search in 3D object databases. *ACM Computing Surveys*, 37(4):345–387, 2005.

[7] R.J. Campbell and P.J. Flynn. A survey of free-form object representation and recognition techniques. *Computer Vision and Image Understanding*, 81(2):166–210, 2001.

[8] Brian Curless. From range scans to 3D models. *ACM SIGGRAPH Computer Graphics*, 33(4):38–41, 1999.

[9] Y. Guo, M. Bennamoun, F. Sohel, M. Lu, and J. Wan. 3D object recognition in cluttered scenes with local surface features: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence. In press*, 2014.

[10] Y. Guo, M. Bennamoun, F. Sohel, J. Wan, and M. Lu. 3D free form object recognition using rotational projection statistics. In *IEEE 14th Workshop on the Applications of Computer Vision*, pages 1–8, 2013.

[11] Y. Guo, F. Sohel, M. Bennamoun, M. Lu, and J. Wan. Rotational projection statistics for 3D local surface description and object recognition. *International Journal of Computer Vision*, 105(1):63–86, 2013.

[12] Y. Guo, F. Sohel, M. Bennamoun, M. Lu, and J. Wan. TriSI: A distinctive local surface descriptor for 3D modeling and object recognition. In *8th International Conference on Computer Graphics Theory and Applications*, pages 86–93, 2013.

[13] Y. Guo, F. Sohel, M. Bennamoun, J. Wan, and M. Lu. An accurate and robust range image registration algorithm for 3D object modeling. *IEEE Transactions on Multimedia. In press*, 2014.

[14] Y. Guo, J. Wan, M. Lu, and W. Niu. A parts-based method for articulated target recognition in laser radar data. *Optik - International Journal for Light and Electron Optics*, 124(17):2727–2733, 2013.

[15] Y. Guo, J. Wan, J. Zhang, K. Xu, and M. Lu. Efficient registration of multiple range images for fully automatic 3D modeling. In *9th International Conference on Computer Graphics Theory and Applications, in press*, 2014.

[16] S.M.S. Islam, M. Bennamoun, R. Owens, and R. Davies. A review of recent advances in 3D ear and expression invariant face biometrics. *ACM Computing Surveys*, 44(3):14, 2012.

[17] K. Lai, L. Bo, X. Ren, and D. Fox. A large-scale hierarchical multi-view RGB-D object dataset. In *IEEE International Conference on Robotics and Automation*, pages 1817–1824, 2011.

[18] Y. Lei, M. Bennamoun, M. Hayat, and Y. Guo. An efficient 3D face recognition approach using local geometrical signatures. *Pattern Recognition*, 47(2):509–524, 2014.

[19] G. Mamic and M. Bennamoun. Representation and recognition of 3D free-form objects. *Digital Signal Processing*, 12(1):47–76, 2002.

[20] A. Mian, M. Bennamoun, and R. Owens. Automatic correspondence for 3D modeling: an extensive review. *International Journal of Shape Modeling*, 11(02):253–291, 2005.

[21] A.S. Mian, M. Bennamoun, and R. Owens. Automated 3D model-based free-form object recognition. *Sensor Review*, 24(2):206–215, 2004.

[22] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, 2005.

[23] D. Munoz, J. A. Bagnell, N. Vandapel, and M. Hebert. Contextual classification with functional max-margin markov networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 975–982, 2009.

[24] P. Papadakis, I. Pratikakis, T. Theoharis, and S. Perantonis. Panorama: a 3D shape descriptor based on panoramic views for unsupervised 3d object retrieval. *International Journal of Computer Vision*, 89(2-3):177–192, 2010.

[25] J. Salvi, C. Matabosch, D. Fofi, and J. Forest. A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, 25(5):578–596, 2007.

[26] A. Scheenstra, A. Ruifrok, and R. Veltkamp. A survey of 3D face recognition methods. In *Audio-and Video-Based Biometric Person Authentication*, pages 325–345. Springer, 2005.

[27] J.W.H. Tangelder and R.C. Veltkamp. A survey of content based 3D shape retrieval methods. In *IEEE International Conference on Shape Modeling and Applications*, pages 145–156, 2004.