

Point Cloud Recognition Algorithm based on the Fusion of VFH and SHOT

Wanqiu Yuan^{1,*}, Bin Lu¹

¹School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, Hebei Province, China

Abstract

Point cloud recognition is a current research hotspot in computer vision, which is widely used in aerospace, medical and health, and unmanned driving. The existing point cloud recognition algorithms are mainly divided into: methods based on local features and methods based on global features. The former is more robust to the occlusion of the scene, but the problem of mismatching of feature points is serious; the latter has a higher description of the scene. However, there are defects in local recognition. To solve the above problems, a point cloud recognition algorithm is proposed that combines the statistical histogram (SHOT) algorithm of geometric distribution information and the viewpoint feature histogram (VFH) algorithm. This article takes the steps of point cloud recognition-key point extraction, feature description, feature matching, and hypothesis verification as the main objects. First, the VFH feature and FLANN library are used to realize the initial recognition of the point cloud target, and the candidate can be quickly obtained from the constructed point cloud library. model. Secondly, use the SHOT feature for accurate recognition, and finally choose the Hough voting algorithm to eliminate the mismatched points, calculate the rotation and translation matrix of the target point cloud and the scene point cloud, and obtain a good matching effect. By testing indoor scene targets with different degrees of defect, the algorithm recognition rate in this paper has reached more than 90%, which can meet the needs of point cloud recognition.

Keywords

VFH descriptor; SHOT descriptor; Point cloud recognition; Feature matching.

1. Introduction

1.1. Research Background and Purpose

In recent years, as an important research direction in the field of computer vision, target recognition technology has received close attention from all walks of life. Target recognition is the process of identifying targets by extracting the characteristic information of the target. It is image processing, pattern recognition and artificial The core content of the intelligent field has a wide range of applications in the fields of aerospace, medicine and health, road monitoring and industrial inspection.

With the increasingly low price of 3D scanning equipment and the development of point cloud processing technology, target recognition objects gradually transition from the traditional two-dimensional image direction to three-dimensional point cloud data. The acquisition technology of two-dimensional image-based recognition methods is mature, and the upsurge of machine learning has developed into a relatively mature field, but it still has obvious limitations. First of all, the image acquired by a traditional camera is a projection of a three-dimensional object on a two-dimensional plane. In the actual scene, there are situations such as shooting illumination differences and camera viewing angle changes. This process will inevitably cause information

loss, and it is very vulnerable to various Kind of interference. The 3D point cloud is mainly used by a scanner such as Microsoft Kinect to laser scan the outer surface of the model at different viewing angles, and record the position, color and even texture information of the target according to the content obtained by the reflection. Compared with two-dimensional images, point cloud data contains the three-dimensional coordinates of the sampling points of the scene, providing more geometric information, and the features extracted from it are not affected by rotation, scale and illumination, and are closer to the real environment. Recognition in complex scenes shows a more reliable processing effect.

In addition, point cloud recognition technology is also applied to other fields such as robotics, biomedical technology, unmanned driving technology, smart home, etc., and has an indispensable position in industrial production and daily life. However, there are a lot of nuisance phenomena in the point cloud data obtained directly through the scanning device. How to achieve accurate target recognition in the presence of noise and point cloud density changes has been a hot spot in recent years.

1.2. Research Status of Point Cloud Recognition

Along with the movement of the target, the sink node timely notifies the sensor nodes in the relevant detection area to join in the process of target tracking. Figure 1 is the flow chart of the moving target tracking process.

The existing point cloud recognition methods are mainly divided into two categories: methods based on local features and methods based on global features.

The local feature-based method first extracts the key points of the scene and the model, and calculates the local feature descriptors after selecting appropriate neighborhoods for the key points, and matches the feature descriptors of the model with the feature descriptors of the scene, and matches certain features The rule finds topologically consistent matching clusters, and finally the final recognition result is obtained through hypothesis verification on the matching clusters.

Forme [1] proposed the shape context (3DSC) method, which divides the key point neighborhood into spherical grids, and completes object recognition by comparing the number of point clouds in each grid. Rusu [2] proposed the Point Feature Histogram (PFH) descriptor. This method counts the normal vector angle and distance between any two vertices, and maps it to the histogram as a local feature. This method is computationally expensive and compares time consuming. Rusu later improved it and proposed a fast point feature histogram (FPFH) descriptor. This method only records the relationship information between a point normal vector and its k-neighbor point normal vector, which greatly improves the computational efficiency. Zhang Zhe [3] and others combined the FPFH descriptor with the traditional iterative closest point method and proposed a new point cloud registration method, which improved the registration efficiency. Tombari [4] and others proposed the histogram signature method (SHOT), which uses the feature point as the center of the spherical grid to support the domain, and the normal vector of each point and the angle information of the local reference coordinate system are used as local features. Rotation is invariant and robust to noise. Hao Wen [5] et al. summarized and compared typical local 3D descriptors, and SHOT descriptors show excellent performance.

The method based on global features first needs to segment the point cloud scene, segment the scene point cloud containing multiple objects into several point cloud clusters, and then encode the three-dimensional geometric information of each point cloud cluster and express it as a group Global features, based on this for identification and classification. Rusu proposed the viewpoint feature histogram [7] (VFH). This method adds viewpoint orientation information to the calculation of the angle between the relative normals, which can distinguish the pose of the object while maintaining the invariance of rotation and scaling. Guo Baoqing [6] and others

used the extraction of VFH features to classify and identify the point clouds of three typical objects in railway scenes, and fulfill the requirements of foreign body intrusion detection in key railway sections. Aldoma [8] proposed the clustered viewpoint feature histogram (CVFH) based on the viewpoint feature histogram. After removing noise points and boundary points, the method uses the area growth method to divide the point cloud into multiple point cloud clusters, and then Calculate VFH for each part. This method is more robust to local noise and surface occlusion.

The method based on global features has fast calculation speed, better descriptiveness on point cloud clustering with simpler shape, higher recognition rate, but requires pre-segmentation of the scene and is more sensitive to occlusion. The recognition method based on local features can calculate the posture of the object more accurately in the presence of occlusion, and the recognition effect is better. In order to improve the accuracy of the point cloud recognition algorithm when the point cloud target is missing, this paper proposes a recognition algorithm that combines the global feature VFH descriptor and the local feature SHOT descriptor.

2. Algorithm Flow

The point cloud recognition process includes: data preprocessing, global feature matching, local feature matching, and Hough voting. The process flow is shown in Figure 1. In the offline training stage, the SHOT feature descriptor and the VFH feature descriptor of the point cloud model are calculated, and the index containing the point cloud model category and VFH feature is stored in the kd tree structure, thus establishing Point cloud model library for later point cloud recognition process.

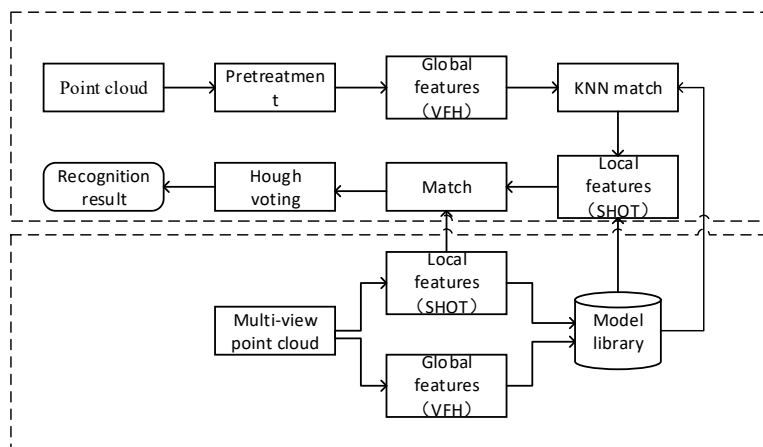


Figure 1. Algorithm flow framework

The implementation steps of the point cloud recognition algorithm are:

1. The input point cloud is preprocessed to obtain the scene point cloud P to be recognized.
2. Calculate the VFH feature of the point cloud P target, based on the FLANN library and the kd tree structure, perform a KNN search on the calculated VFH feature descriptor of the point cloud target to be identified and all the VHF feature descriptors in the model library. According to the set chi-square distance threshold d , N nearest neighbor point cloud models with chi-square distance less than d are obtained, and the N models are sorted according to the chi-square distance.
3. If the minimum chi-square distance of the N models is much smaller than the set chi-square threshold d , that is, $d_{\min} \ll d$, then the model point cloud corresponding to the minimum value is used as the target point cloud to complete the recognition; otherwise, local features are performed Identify, go to ④-⑤.

4. Calculate the SHOT feature of the point cloud target P to be identified, match it with the SHOT feature descriptors of the N model point clouds obtained from the global feature matching result, and record the feature matching logarithm.
5. Using the Hough voting algorithm [9] to remove the mismatched feature point pairs, if the maximum value of the matching logarithms in the N model point clouds is greater than the set threshold, the corresponding model point cloud is considered to be the target point cloud, and n pairs are returned Best matching point pair, output rotation and translation matrix.
6. The point cloud to be recognized and the point cloud of the model are matched and displayed in a connected pair to complete the recognition.

3. Point Cloud Processing

3.1. Point Cloud Preprocessing

Currently, statistical analysis techniques are often used to remove outliers from point cloud data. It mainly calculates the distance distribution from points to neighboring points in the point cloud data, and considers points that are far from the subject's point cloud far greater than the average distance between neighboring points around the subject as outliers.

The main idea of the statistical outlier elimination filtering algorithm [10] is: for each point in the point cloud, calculate the average distance from all its k neighboring points, and get the Gaussian distribution shape of the result according to the average distance expectation and standard deviation And the distance threshold, the points whose average distance is greater than the threshold are regarded as outliers and removed.

3.2. Extract Key Points

The key point is the continuous and stable structure points set on the point cloud. 3D point cloud key point extraction and feature description are the most critical part of point cloud processing. The robustness of most algorithms such as point cloud recognition, segmentation, and registration Both depend on the results of key point extraction and feature descriptor processing.

The simplest key point extraction algorithm is point cloud grid simplification and uniform sampling, but the key points obtained by these methods contain less overall information. Literature [11] compares and analyzes 3D key point detection operators, and comprehensively considers the local The ISS algorithm [12] shows obvious advantages in terms of the number of key points, efficiency, etc. This paper chooses this algorithm to extract the key points of the 3D point cloud.

3.3. Feature Description

The viewpoint feature histogram (VFH) [7] adds the component of the viewpoint direction on the basis of the fast point feature histogram (FPFH), and the constructed features are scale-invariant. A VFH feature is composed of the component describing the surface shape of the extended FPFH and the component related to the direction of the viewpoint. The former uses the point pair between the centroid of the point cloud object and other points on the surface as the calculation unit, and the latter uses the viewpoint and other points on the surface. The point pair is the calculation unit. Its calculation speed is fast, the recognition ability is strong, and the target feature description effect with simpler shape is better, but its robustness is slightly worse in the case of missing point clouds and more noise.

The components of the same type of object point cloud describing the surface shape are similar, while the components of different types of object point clouds describing the surface shape are different. Therefore, VFH features can be used to identify and classify different objects.

Directional histogram feature descriptor (SHOT) counts the geometric distribution information of the surface in the histogram, which is rotation invariant and has good descriptive and robustness.

The implementation steps of the SHOT descriptor are as follows: Construct a spherical area with the feature point p as the center and r as the radius, and divide the area into 32 grid areas along the radial direction twice, azimuth eight times, and pitch twice; in each grid area, calculate the cosine of the angle between the normal line of the midpoint and the normal line of the center point p , and count it as a histogram of 11 small cells, the final SHOT The length of the descriptor is $32 \times 11 = 352$.

4. Point Cloud Target Recognition Experiment

4.1. Experiment Platform

This system uses 8GB of Windows 10 64-bit operating system, combined with the open source point cloud library PCL 1.9.0, and is programmed in the Visual Studio 2017 environment.

200 sets of data recorded by Kevin Lai [13] and others using a Kinect-style 3D camera, including nine types of common household objects such as bowls, hats, cereal boxes, coffee cups, chairs, tables, etc., each type of object Contains several objects of different styles, each style selects a set of the most complete point cloud as the standard sample to build the model library, and the rest of the point cloud data with different degrees of defects are used as the samples to be tested, which are set according to different degrees of certainty The Chi-square distance threshold and the Hough voting threshold are used to verify the performance of the point cloud recognition algorithm when the point cloud to be recognized is defective. Part of the model object point cloud is shown in Figure 2.

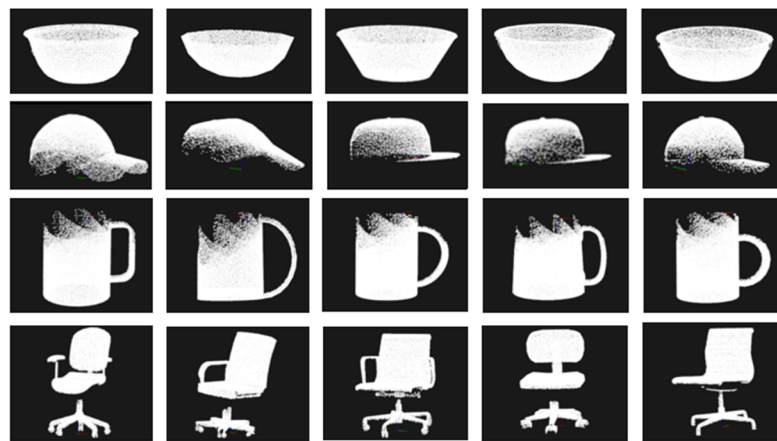


Figure 2. Part of the model object point cloud

4.2. Experimental Results and Analysis

In the process of 3D point cloud recognition, first compare the VFH feature descriptor of the target to be recognized with the VFH feature descriptor in the model library to obtain a candidate model with a smaller gap. Since the feature descriptor has been stored locally, the calculation speed is Soon, through the initial matching of VFH feature descriptors, the recognition efficiency is also high. Local feature matching is performed on the initial recognition results, which improves the recognition rate and saves time.

4.2.1 Global feature matching

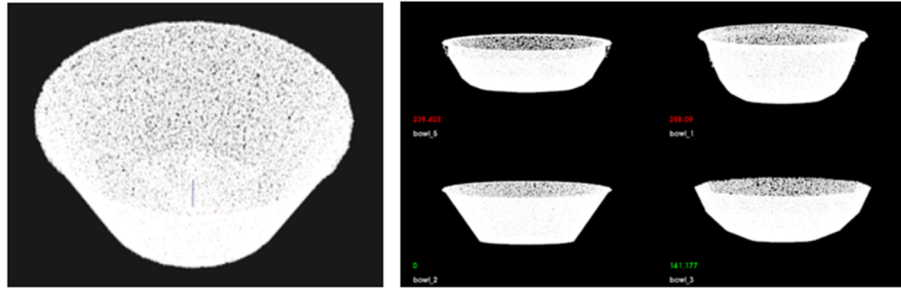


Figure 3. Initial recognition result

Figure 3 shows the results of the global feature VFH matching when the bowl_2 model is intact and the chi-square threshold is set to 200. The green value in the lower left corner indicates the distance within the chi-square threshold, and the red value indicates that it is outside the chi-square threshold. the distance. Since the minimum distance value obtained is much smaller than another matching result, the final recognition result can be directly obtained.

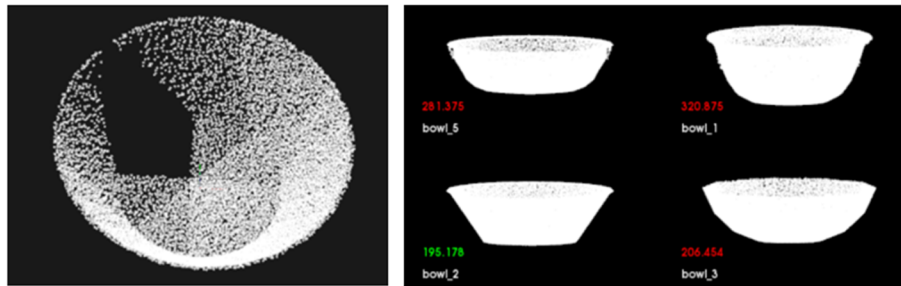


Figure 4. Initial recognition result

Figure 4 shows the results of the global feature VFH matching when the bowl_2 model is missing the point cloud and the chi-square threshold is set to 200. There are some missing points in the point cloud to be identified. The results found by VFH feature matching are all similar to the bowl_2 category. Models with large differences in the shape of the bowl, such as chairs, sofas, etc., do not appear within the threshold set for initial recognition.

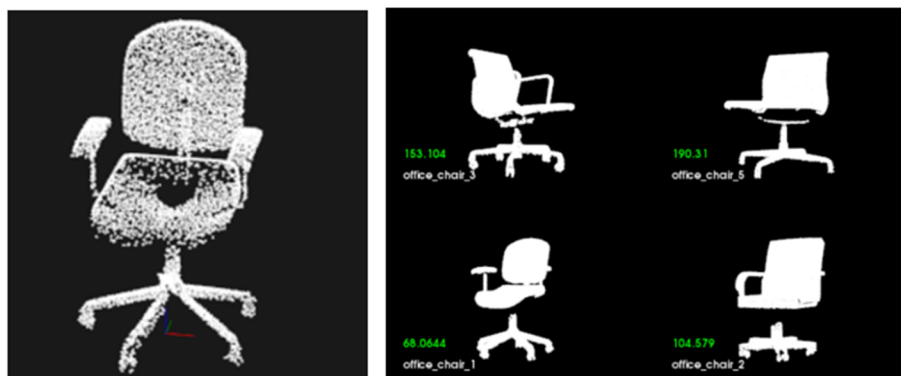


Figure 5. Initial recognition result

Experiments show that the VFH algorithm can query models that are similar in shape to the target point cloud to be recognized. The recognition efficiency is higher when the target point cloud is complete, but the recognition rate is significantly reduced when the point cloud is missing. Therefore, in this paper, when the point cloud target is missing, the candidate target is screened through VFH, and then the local feature descriptor SHOT is used for identification.

4.2.2 Local feature matching

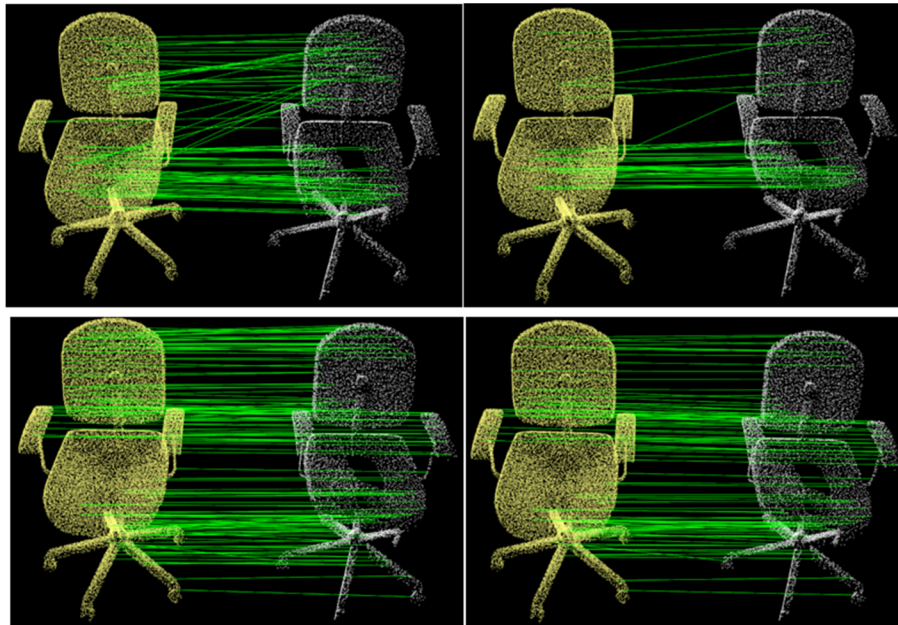


Figure 6. Effect picture of multiple matching methods

The above figure respectively shows the feature matching results of FPFH and SHOT descriptors based on the nearest neighbor and two-way nearest neighbor methods after the key points are obtained using the ISS algorithm. The left is the model point cloud, the right is the point cloud to be identified, and the green line is the connection effect of the feature matching pair.

According to the experimental results, the FPFH descriptor has certain defects. Compared with the SHOT descriptor, the description effect is worse, and there are more mismatches, and the obtained feature matching pairs are more concentrated; while the matching pairs obtained by the SHOT descriptor are more Scattered and less crossed, the description effect of point cloud is better than FPFH descriptor. At the same time, it can be seen that the feature matching pairs found by the nearest neighbor matching method include a large number of mismatched pairs, while the two-way nearest neighbor matching method eliminates a small number of correct matching pairs while removing a large number of mismatched pairs. While improving the accuracy of matching pairs, it also improves the time efficiency of subsequent algorithms.

4.2.3 Summary

Table 1. Identification of different defect percentages

Percentage defect	0%	0%—5%	5%-10%
Recognized/unrecognized	58/1	136/4	26/3
Recognition accuracy rate	98.3%	97.1%	90%

Table 1 shows the results of target recognition when the point cloud data is defective. Experiments show that when the target is complete, the recognition process can be directly completed through the VFH feature and KNN algorithm, and the calculation speed is faster, and the recognition efficiency can reach 98.3% ; When there are fewer defects in the target, the VFH feature and KNN algorithm can be used to quickly screen to obtain candidate models with

similar shapes, and then use the SHOT descriptor, improved feature matching principles and Hough voting to complete the final recognition, with a recognition rate of 90 %the above.

5. Conclusion

In order to improve the recognition rate of point cloud defects in indoor scenes, this paper proposes an improved point cloud recognition algorithm that combines local and global features. It uses global feature VFH for rapid object recognition and local feature SHOT for target positioning. , It is proved by experiments that the recognition rate of the algorithm in this paper is above 90%, which can meet the needs of scene target recognition.

References

- [1] Andrea Frome,Daniel Huber,Ravi Kolluri, et al.Recognizing Objects in Range Data Using Regional Point Descriptors[C].//8th European Conference on Computer Vision (ECCV 2004), part 3.2004:224-237.
- [2] Radu Bogdan Rusu,Nico Blodow,Zoltan Csaba Marton, et al.Aligning Point Cloud Views using Persistent Feature Histograms[C].//2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Pages 3133-3660.2008:3384-3391.
- [3] Zhang Zhe,Xu Hongli,Yin Hui.A fast point cloud registration algorithm based on key point selection[J].Progress in Laser and Optoelectronics,2017,54(12):149-157. DOI:10.3788/LOP54.121002 .
- [4] Samuele Salti,Federico Tombari,Luigi Di Stefano.SHOT: Unique signatures of histograms for surface and texture description[J].Computer vision and image understanding: CVIU,2014,125:251-264.
- [5] Hao Wen,Wang Yinghui,Ning Xiaojuan, et al. Overview of point cloud-oriented 3D object recognition methods[J].Computer Science,2017,44(9):11-16. DOI:10.11896/j.issn.1002-137X. 2017.09.002.
- [6] Guo Baoqing, Yu Zujun, Zhang Nan, et al. Three-dimensional point cloud segmentation and classification recognition algorithm for railway scene[J].Chinese Journal of Scientific Instrument,2017,38(9):2103-2111. DOI:10.3969/j.issn.0254-3087.2017. 09.002.
- [7] Rusu R B , Bradski G R , Thibaux R , et al. Fast 3D recognition and pose using the Viewpoint Feature Histogram[C]// 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 18-22, 2010, Taipei, Taiwan. IEEE, 2010.
- [8] Aldoma, Aitor,Vincze, Markus,Blodow, Nico, et al.CAD-model recognition and 6DOF pose estimation using 3D cues[C].//2011 IEEE International Conference on Computer Vision Workshops. [v.1].2011:585-592.
- [9] Hough Voting for 3D Object Recognition under Occlusion and Clutter[J].IP SJ Transactions on Computer Vision and Applications,2012,4(0):20-29. DOI:10.2197/ipsjtcva.4.20.
- [10] Xi'an Engineering University. 3D point cloud denoising method based on statistical outliers and adaptive bilateral hybrid filtering: CN201910506642.1[P].2019-10-18.
- [11] Wu Fan. Research on 3D object recognition and understanding algorithm based on local surface features in multi-target scenes[D]. Hubei: Wuhan University of Science and Technology, 2016. DOI:10.7666/d.Y3021969.
- [12] Zhong Y . Intrinsic shape signatures: A shape descriptor for 3D object recognition[C]// IEEE International Conference on Computer Vision Workshops. IEEE, 2010.
- [13] Kevin Lai, Liefeng Bo, Xiaofeng Ren, et al. A Large-Scale Hierarchical Multi-View RGB-D Object Dataset[C].//China Association of Automation. 2011 IEEE International Conference on Robotics and Automation (2011 IEEE International Conference on Robotics and Automation Conference on Automation ICRA 2011) Proceedings. 2011:1817-1824.