Shape Matching of Repeatable Interest Segments in 3D Point Clouds

Joseph Lam[1] Michael Greenspan[1,2]

[1]Dept. Electrical & Computer Engineering, Queen's University, Kingston, Canada
[2]School of Computing, Queen's University, Kingston, Canada

Abstract

A novel approach to object recognition based on shape matching of repeatable segments is presented. The motivation is to increase the recognition system robustness in handling problems such as noise corruption at a local level, featureless surfaces, and variations in 3D data sources. Inspired by the detection of repeatable interest points, interest segments were extracted through region growing and the reconstruction of piece-wise boundary curves from connected interest points. An object pose is automatically estimated if only one of the repeatable scene segments can be matched and aligned correctly with a model segment. To demonstrate this capability, shape matching of selected segments, filtered by size, were registered using the 4 points congruent sets (4PCS) algorithm and compared with an overlap metric. Three different free-form objects were evaluated against nine different occluded and cluttered 2.5D scenes. It was found that on average 1.4 ± 0.8 *scene segments can be matched correctly to a model segment in the database, indicating that a highly robust object recognition system will result.*

1. Introduction

Two common approaches to the problem of object recognition in 3D data are Local Shape Descriptors (LSDs) and global shape matching. The LSD approach relies on encoding geometric characteristics within a local neighbourhood into multi-dimensional vectors (i.e. *features*), and establishing one-to-one point correspondences by comparing scene features to a database of model features. LSDs can be quite effective at recognizing objects in cluttered or occluded scenes [6] [11] [16] [17], but a lack of distinctive local features, noise corruption of the 3D data, or differences between the resolutions of the scene and model data can adversely affect the performance of LSD-based recognition. Alternately, global shape matching methods are more resilient to changes at the local level since the entire shape is considered, but typically global methods require the target object to be first isolated from the scene [1] [12]. They therefore become less effective in cluttered and occluded 2.5D scenes.

In light of the limitations of each of these approaches, the present study investigates the extraction of interest segments that can be used to facilitate shape matching of 3D data in possibly cluttered and occluded scenes. Segments have been demonstrated to be a strong candidate to generate object recognition events in 2D images such as texton boosting [13]. In our previous work [9], it was shown that interest points can be used to robustly generate repeatable 3D segments, resulting in the same (or similar) segments under a variety of different conditions (e.g. viewpoint, noise levels, sensor mode, resolution, etc.). The segments tend to be repeatable follows from the characteristic that the interest points upon which they are based are highly repeatable. Indeed, while interest points are not attributed like features, and therefore lack the distinctiveness that is essential for generating meaningful correspondences, they do tend to be highly repeatable, which is one of the main characteristics that supports their utility.

In this work, we demonstrate that repeatable segments can be coupled with a shape matching method to effectively recognize 3D objects in cluttered and possibly occluded 2.5D scenes. The interest segments can either be used to drive part labelling and object detection using shape indexing, or they can alternately be used to generate features for establishing one-to-one correspondences. Indeed, if the segments are sufficiently distinct, then only a small number (minimally just one) of segments need be repeatably extracted for a given object viewpoint, in order to yield a successful recognition event.

An algorithm was previously presented for extracting interest segments from free-form objects in 3D point clouds [9]. The technique is a combination of a sequence of processes that includes: a) the extraction of interest points, also referred to as keypoints; b) piece-wise linear curve reconstruction, and; c) region growing. Interest points, often selected as a sparse subset of points around which to construct features in LSD-based object recognition, are ex-

978-1-4673-1612-5/12/$31.00 ©2012 IEEE

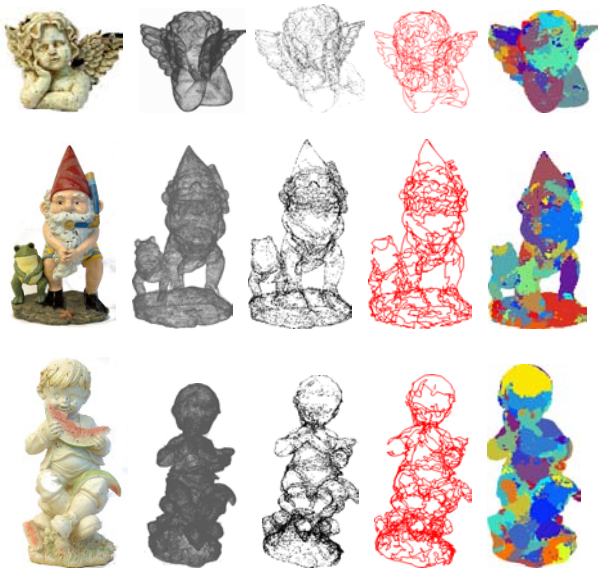


Figure 1: Segmentation based on interest points: Original models (1^{st} column); Corresponding 3D point clouds (2^{nd} column); DoN interest points (3^{rd} column); Piece-wise linear curve reconstruction (4^{th} column); Interest segments (5^{th} column)

tracted from the data using interest operators based on a saliency measure of the 3D surface [10] [14] [15] [17] [18] . Interest points are then joined into piece-wise linear curves that act as boundaries for region growing, thus achieving the segmentation of free-form surfaces.

The contribution of this work is twofold. First, a segment merging method is introduced which improves upon shape matching when the scene data is over-segmented compared to the model data (which is the desired condition). The second contribution is a set of quantitative experiments conducted to investigate the shape matching ability of repeatable interest segments extracted from 3D data under various conditions. These experiments include comparisons between the segmentation of complete 3D models with various 2.5D scenes, including cluttered and occluded scenes containing multiple objects.

2. Related Work

Object recognition that relies on matching point correspondences using LSDs have become a common accepted object recognition paradigm in recent years. Many 3D LSDs have been proposed, including Spin Images [6], Point Fingerprint [16], Shape Context [11], and recently Variable Dimensional LSDs [17]. Object recognition that is based on matching LSDs can be broken down into three stages: 1) The selection of points (interest points) to generate the LSDs; 2) Generation of LSDs, and; 3) Comparison and matching of LSDs.

To compensate for the usually high dimensionality of the LSDs, and the relative expense of generating them, interest operators are often utilized to select a small subset of repeatable points for the generation of only a small number of distinctive LSDs. A few existing 3D interest operators include the Eigenvector-based operator [10], Laplacian-based operator (also a interest region operator) [18], and the 3D Harris operator [14]. Both Laplacian and Harris operators are 3D extensions to classical 2D operators used widely within the computer vision community. In most cases, a non-maxima suppression step is applied to filter out 3D points that are not a local maximum, so that the extracted interest points are sufficiently separated spatially.

The segmentation of 3D data in early work mainly deals with range images or depth maps, as opposed to unordered point clouds. Most developed techniques usually treat objects as a combination of different surface patches and edges that can be described in parametric space, which include methods such as clustering points based on surface properties [3], robust variable-order surface fitting using the H-K mapping [2], and extraction of primitives with the General Hough Transform (GHT) [19]. These methods are usually divided into edge-based, surface-based, and hybrid edge-surface-based approaches. The major limitation of these techniques is that objects with irregular shapes, such as free-form surfaces, cannot be properly segmented.

The 3D mesh is another format for representing 3D objects, and a popular approach to mesh segmentation is applying graph cuts and energy minimization [8]. Recently, [7] considered the problem of identification and labelling of 3D mesh segments from isolated objects using probabilistic learning. The conditional random field (CRF) was employed to survey a collection of common 3D features that include both single features (e.g. PCA singular values) and pairwise features (e.g. dihedral angles). It can be concluded that previous work usually treats segmentation and recognition as two separate problems in the 3D domain, unlike the present study in which segmented results are directly used as the input for an object recognition system.

3. Interest Segment Extraction

The methodology proposed here is based on the observation that interest points can be localized repeatably using interest operators. By applying techniques such as RANSAC [4] and nearest neighbour search, boundaries in the form of piece-wise linear curves are constructed from sets of adjacent interest points. Interest segment extraction is then achieved by applying region growing, where segments are localized by finding points within a region that is bounded by the interest curve boundaries. Fig. 1 demonstrates the intermediate results for various 3D objects at different stages of the segmentation process. The processes described in Sec. 3.1, Sec. 3.2, and Sec. 3.3 have been previously pre-

sented in our recent work [9], but are repeated here for completeness. The techniques described in section Sec. 3.4 and the remainder of the paper are novel contributions.

3.1. Detection of Interest Points

Unlike other approaches that use interest operators, non-maximum suppression is omitted here in the extraction of interest points. This is because it is advantageous to include all interest points, including the non-maxima, to facilitate the estimation of piece-wise linear interest curves (Sec. 3.2).

The Difference of Normal (DoN) operator [5], a scale-based operator, was selected as our interest operator for the proposed segmentation technique. DoN functions similarly to the Difference of Gaussian (DoG) operator in 2D images, where two normals $\{\vec{n_1}, \vec{n_2}\}$ based on two different respective neighbourhood radii $\{r_1, r_2\}$ are computed for each point in the point cloud. A point is declared as an interest point if the solid angle between $\vec{n_1}$ and $\vec{n_2}$ exceeds a predetermined threshold. The DoN operator is effective at identifying those points whose surface geometry changes significantly as the neighbourhood radius increases from r_1 to r_2.

Since the DoN operator extracts interest points by measuring differences based on neighbourhood scale rather than some saliency measure at a fixed scale, the DoN is a good choice for comparisons between scans of different resolutions, such as those used in this study. Column 3 of Fig. 1 shows the extracted interest points of various 3D objects using the DoN operator. It can be seen that the interest points are selected at locations where there is significant geometric variation on the object, rather than smoothly varying regions.

3.2. Boundary Curve Reconstruction

To achieve region growing for 3D free-form surface segmentation, the second and crucial step is to form continuous and closed boundary curves.

Let p_j be the j^{th} interest point, and let $\mathbf{P} = \{p_j\}_1^n$ be the set of all n such extracted points. The p_j are joined into sets of interest curves, and then connected to form continuous curves, as follows.

For each interest point p_j a tangent direction $\vec{d_j}$ is estimated by applying RANSAC to the neighbours of p_j that are elements of **P**.

Each point p_j is only allowed to be linked with a maximum of two other points p_i and p_k. A point p_k is selected by finding the point that is closest to $\vec{d_j}$ within **P** of p_j . Similarly, the point p_i is connected to p_j, if p_j is the nearest point to $\vec{d_i}$, and is within **P** of p_i. These three joined points define the incoming vector $\vec{v_{ij}}$ and outgoing vector $\vec{v_{jk}}$ at p_j. To restrict abrupt changes (i.e. *kinks*) in the direction of the interest curves, the solid angle between $\vec{v_{ij}}$ and $\vec{v_{jk}}$ is limited to a maximum threshold value (45^0 in this work).

Tree traversal is used next to search for the longest path along the linked interest points. Each longest path can be defined as an independent interest curve, represented in the form of a piece-wise linear curve.

Redundant representation of a single boundary line can occur if there are sections where two interest curves are within close proximity of each other. These sections are eliminated and the two interest curves are merged into a single interest curve.

To ensure that a closed segment is formed, the end of all interest curves are always joined to the end of another interest curve. The formation of continuous and closed boundary lines has no limitation on the solid angle between the ends of two interest curves, for example, the tip of the gnome's head. Each end of an interest curve is connected to the end of another by selecting the one that gives the shortest Euclidean distance between them.

The boundary curves reconstructed for various 3D objects can be seen in Fig. 1, column 4.

3.3. Interest Segments from Interest Curves

The final step of the segmentation method is to label each point within an enclosed segment defined by a boundary curve with a unique label. Let **B** denote the set of all points that are geometrically close to any boundary curve, and let the set of all other points that are not on a boundary (and which are therefore within some region) be denoted as **R**. Points within **R** are randomly chosen as seed points to start the process of region growing. A seed point is assigned a new region label if it does not belong to any previously labelled region, and all of its adjacent neighbours within **R** that are not yet labelled are added to this same region. If an adjacent neighbour point has already been associated with another distinct region, then the two regions are merged together into one. A region is allowed to grow until the neighbours of all points in the region are labelled as associated with that region, or are within **B**.

To refine the results, regions with only one point are allowed to concatenate with the largest regions nearby, which is particularly useful for sparse data. Finally, all points within **B** are labelled with the region associated with the closest neighbouring point that is within **R**. The fifth column of Fig. 1 illustrates the resulting region segmentation for various 3D objects.

3.4. Merging Segments

Interest segments in an unknown input scene will often be over-segmented as a result of outliers caused by noise, data resolution, differences in source data, etc. It is therefore desirable that scene segments can be merged in such a fashion so as to increase the shape matching rate.

One approach to segment merging is the n choose k factorial solution, where n is the total number of neighbour segments plus the current segment, and $1 <= k <= n$. Starting will all scene segments, the set **S** of all possible n choose k merged segments are generated. The set **S** also includes all of the original scene segments, so there will be overlap between the elements of **S**. Segments that are small in size can possibly be eliminated as they offer little distinction that is essential to facilitate successful shape matching. All the segments in **S** are subsequently used in the shape matching process. In the present study, for experimentation purposes, a supervised merge method based on the known ground-truth transformation is used, as described in Sec. 4.1.

4. Shape Matching of Interest Segments

The objective of performing shape matching on interest segments is to determine the extent to which segments extracted from a 3D model in the database can be matched correctly with segments extracted from a 2.5D input scene, under various realistic conditions. The motivation behind this test is the understanding that if even only a small number (minimally one) of correct correspondences can be established between scene and model segments, then database objects can be matched, and successful recognition events will result.

4.1. Supervised Merge Based on Repeatability Measures

To demonstrate that the merging of interest segments increases the rate of successful shape matching, which ultimately increases the ability to establish correct correspondences between scene and model segments, a supervised merge of interest segments is applied in this work.

The ground truth transformations between the models and the scenes were obtained by manually selecting a minimum of 3 correspondence points and then applying the Iterative Closest Point (ICP) algorithm. Thus, for each pair of data (3D model vs. 2.5D scene), a one-to-one correspondence can be established between a point p_M on the model and a point p_S in the scene. Each interest segment in a scene is first assigned the most likely corresponding model segment according to the repeatability measure Q, based on the known ground truth transformation. The scene segments are then allowed to merge with neighbouring segments, as the value of Q approaches the maximum of 1.

Let the set of established correspondence points between a model and a scene be denoted as $\mathbf{P}_{MS}$. Given that both the model and the scene are segmented, each point $p_{MS} \in \mathbf{P}_{MS}$ is associated with both a ground truth 3D model segment label R_M and a 2.5D scene segment label R_S. For each segment R_S represented by a scene point within the set $\mathbf{P}_{MS}$, the most likely corresponding R_M can be determined by finding that model segment that has the largest degree of overlap with R_S.

Let N_S denote the number of points in scene segment R_S, N_M the number of points in model segment R_M, and $N_{S\cap M}$ the number of points within the intersection $R_S \cap R_M$. An expression for the quality of the match Q between two overlapping segments R_S and R_M is:

$$Q = \begin{cases} N_{S\cap M}/N_S & \text{if } N_S > N_M \\ N_{S\cap M}/N_M & \text{otherwise} \end{cases} \quad (1)$$

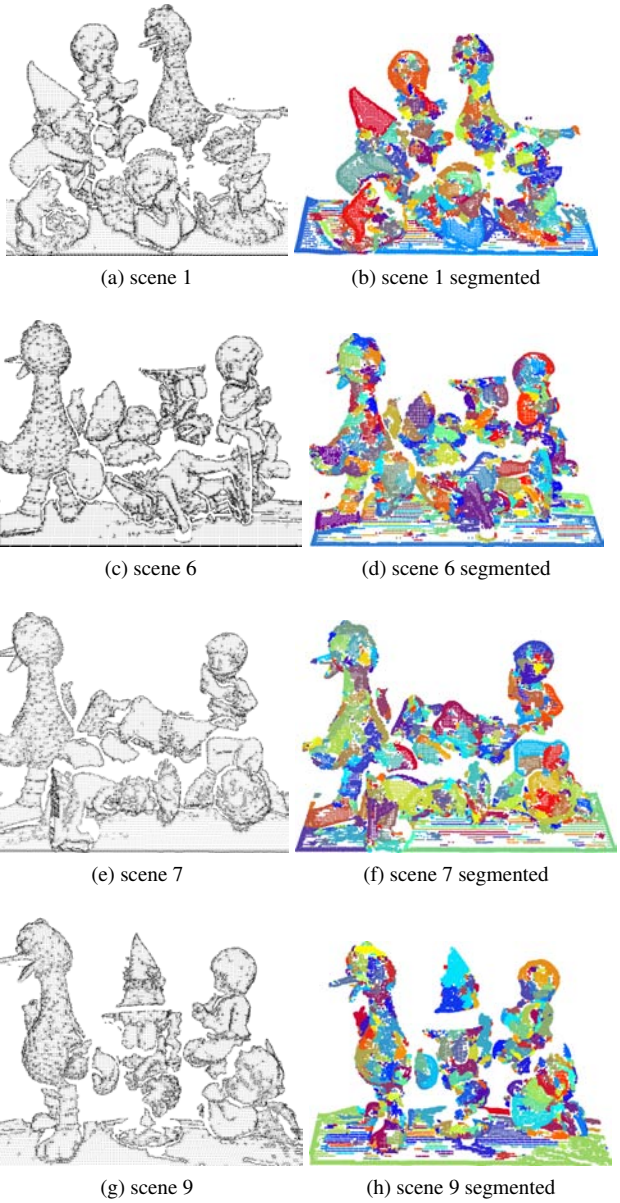
The interest segments in the scene are then ranked against Q, and the segments are only allowed to merge with neighbour segments with lower Q. This is to prevent more repeatable segments from dropping the Q score. To select the optimal combination of segments that generate the best Q, the n choose k strategy described in Sec.3.4 is deployed to select all possible combinations of the current segment and its neighbour segments. The combination that yields the best Q score is the optimal merge for the current segment.

4.2. Shape Registration by 4PCS

A number of methods are available for registering scans. One option is to first align two segments using their centroids and estimated normals. To resolve the rotation ambiguity, the scans can be rotated by a small increment, and then apply ICP. Unfortunately, scenes in this study are often occluded by itself or other objects, which drastically changes the shapes of two segments even though they may be perfectly segmented. Hence, this method is not practical for the current experimental setup.

Instead, the 4-Points Congruent Sets (4PCS) [1] was applied. 4PCS is an algorithm originally developed for robust pairwise surface registration, which according to the claims in the original paper, is able to register two scans with minimum overlap under noisy conditions. Taking advantage of this property, 4PCS is an ideal algorithm to determine the percentage of overlap between a scene segment and a model segment, which can be defined as the matching score $\mathbf{M}= N_{S\cap M}/N_S$.

The 4PCS algorithm works by extracting a congruent base composed of 4 coplanar points in two different sets of data. Two coplanar sets $X = a, b, c, d$ across different sets are efficiently extracted by taking advantage of the preserved ratios $r_1 = ||a - e||/||a - b||$ and $r_2 = ||c - e||/||c - d||$ under affine transformation, where e is the intersection point between the two lines ab and cd. The base X that yields the most aligned points is the best recovered pose between the two data sets. Overall, a number of parameters control the performance of 4PCS, including: δ - the tolerance for the intermediate point e; ρ - the tolerance for approximate congruent set; R - the nearest neighbour search radius, and; L - the number of RANSAC cycles.



(a) scene 1 (b) scene 1 segmented

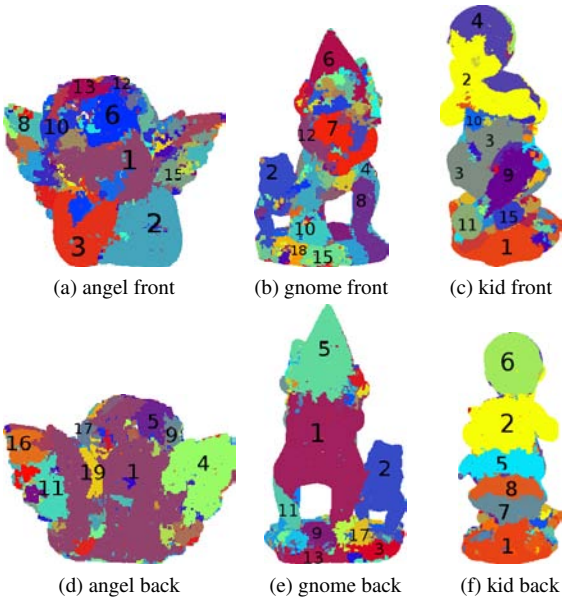
(c) scene 6 (d) scene 6 segmented

(e) scene 7 (f) scene 7 segmented

(g) scene 9 (h) scene 9 segmented

Figure 2: Selected 2.5D scenes acquired from the Vivid LiDAR scanner used in this work

The shape matching of interest segments is based on the essence of a single object detection test, in which the prior information is that the model to be localized is present in an unknown scene. In other words, the defined problem is to retrieve the 3D pose of a known object in an unknown scene. As mentioned, a minimal number of just one correctly matched segment is sufficient to retrieve the 3D pose, which can be verified subsequently by finding the 3D pose that produces the smallest Euclidean distance error by reprojecting the model back into the scene. In the case of the presence of two or even more correctly matched segments,



(a) angel front (b) gnome front (c) kid front

(d) angel back (e) gnome back (f) kid back

Figure 3: Segments for each model labelled by size

the similarity between the recovered 3D poses further increases both the processing speed in matching and the likelihood of an object being localized. The shape matching process is to be carried out for each scene interest segment, compared exhaustively with m model segments, and the best match is chosen based on the score **M**. Subsequently, a verification step consists of first reprojecting the 3D model onto the 2.5D scene using the recovered 3D pose, followed by measuring the average nearest neighbour distance, can be used to verify if the segment with the highest **M** is a true positive or a false positive segement (This step is currently not implemented in the experiment, instead, visual verification is used).

5. Experiments

Experiments based on the evaluation scheme outlined in Sec. 4 were conducted using three free-form objects (Fig. 1) in nine different realistic 2.5D realistic scenes with cluttered and occluded objects (Fig. 2 shows a few sample scenes and the segmented counterparts) Both the 3D models and the 2.5D scenes were acquired using the Vivid LiDAR scanner, although the resolution of the scanned models and the scenes are not the same (e.g. due to the difference in the vantage of the scanner). The resolution of the 3D models were also reduced to improve the efficiency of the algorithm. All the code was written in C++ on a Linux platform, running on a computer with 8 GB of RAM and a quad-core 2.26 GHz processor. The 4PCS algorithm is currently re-implemented without any code optimization, thus the process of shape matching by considering all scene and model segments is exhaustive. To deal with this issue and

	s1	s2	s3	s4	s5	s6	s7	s8	s9
Segment 2	70	70		82	87		76		
Segment 3		81		72	85				
Segment 4			89			97			
Segment 8	89							88	

Table 1: Angel Result: **M** score for true positive segments detected

	s1	s2	s3	s4	s5	s6	s7	s8	s9
Segment 1				85	77		100		93
segment 2	91	86	97	99			97		100
Segment 4		82							
Segment 5									92
Segment 6								91	
Segment 7								90	

Table 2: Gnome Result: **M** score for true positive segments detected

simplify the experiment, only the largest 20 segments from each model and scene were considered (Fig. 3).

5.1. Parameters Selection

Since segments are first filtered based on size, it is ideal that the selection of the DoN angle generates a set of large and descriptive segments. Intuitively, increasing the DoN angle threshold reduces the number of interest points, leading to fewer interest curves and larger interest segments. For the entire experiment, the DoN angle threshold was set at 13^0 for the Angel model, 12^0 for the Gnome and 14^0 for the Watermelonkid model. Other parameters included: The nearest neighbour search, which was set at $5mm$; The DoN neighbourhood scale difference, which was set at $2\times$, and; The voxel size that was used for partitioning the data for efficient nearest neighbourhood search was set at $1cm^3$. The segmentation for 2.5D scenes uses the same set of parameters for all models under comparison during the test. For the 4PCS parameters: $\delta = 10mm$, $R = 4mm$, $\rho = 0.5mm$ and $L = 30$. It is worth mentioning that ρ has a significant impact on the efficiency of the shape matching process: Because the resolution of the scene and model data are different, ρ cannot be so small such that the correct 4 point congruent sets cannot be chosen. The trade-off of increasing ρ is the increase of number of candidate coplanar planes, reducing the processing efficiency.

5.2. Results and Discussion

Figs. 4, 5, and 6 show the resulting scene segments for the three models in the nine different scenes tested in this experiment. The respective matching results based on supervised merged segments is shown in the second row of the same figures, where the true positive segments that record the highest **M** score are coloured in green. A visual confirmation is provided by re-projecting the segments onto the model using the recovered 3D pose from 4PCS. The score **M** for each true positive segment for each model found in each scene are also provided in Tables 1, 2, and 3. On average, each true positive segment is matched with an **M** score of 82, 91, and 88 for the Angel, Gnome, and Watermelonkid models respectively, yielding a total average of **M** = 87. It can be observed that many true positive segments have higher ranks. This is because in this test, only the top 20 scene models are considered, which eliminated any comparisons between lower rank segments based on size. Based only on the top 20 segments in each model and scene, Table 5 provides a summary of the number of true positive segments for each model in each scene. Only one scene for the angel and the gnome model recorded zero true positive segments (possibly because of the selected segments in the current setup). Based on these results, an average of 1.4 ± 0.8 true positive segments can be retrieved from each scene. For completeness, Table 4 shows the same statistic for segments before the merge is applied, using the same set of parameters previoulsy selected. This shows that merging does increase the repeatability and recognition rate of the system.

	s1	s2	s3	s4	s5	s6	s7	s8	s9
Segment 2		97	94	80		81	91	70	85
Segment 3	93					90			
Segment 4	93								
Segment 5		93							
Segment 6			100						
Segment 7					83				

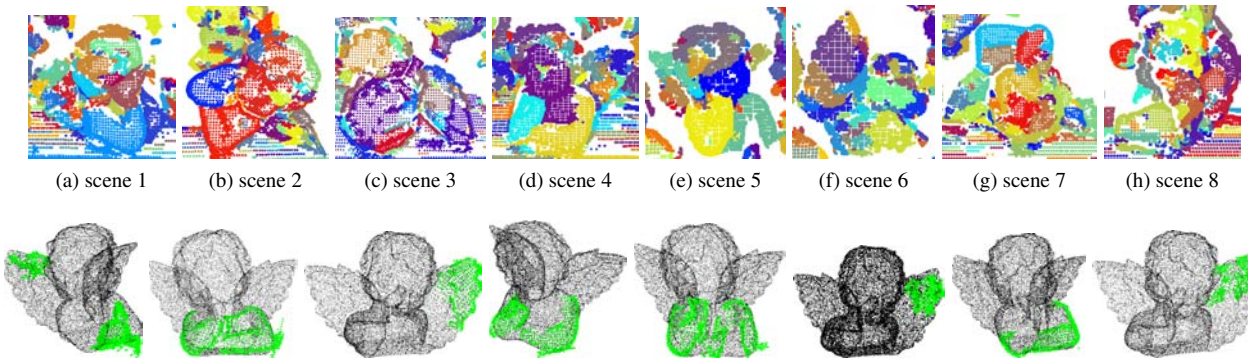
Table 3: Watermelonkid Result: **M** score for true positive segments detected

	s1	s2	s3	s4	s5	s6	s7	s8	s9
Angel	0	2	0	1	2	0	1	1	0
Gnome	1	2	1	1	1	0	2	0	2
Kid	1	1	1	1	0	0	0	0	0

Table 4: Number of true positive segments in each scene using original segments from all models

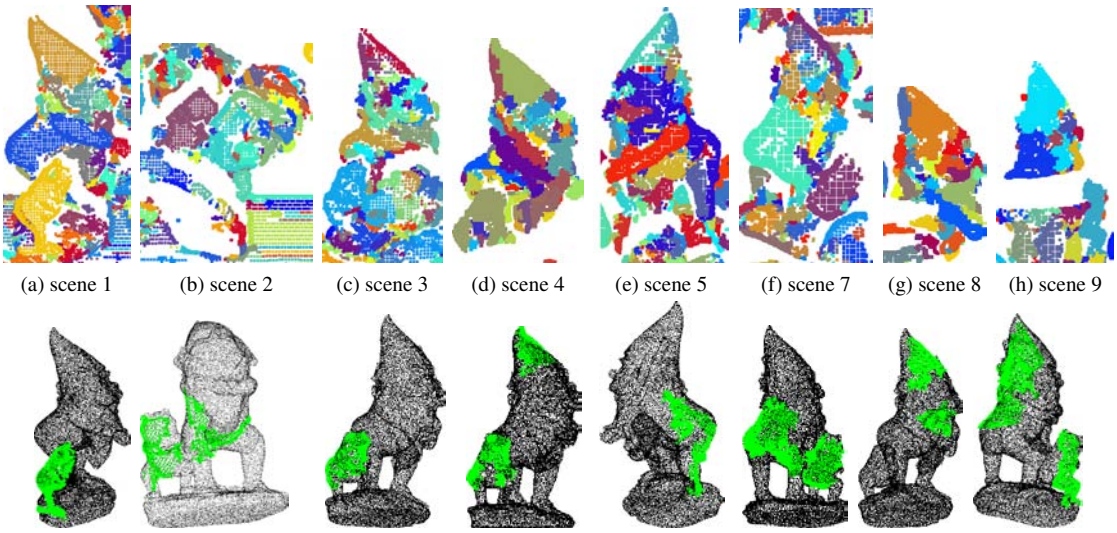
	s1	s2	s3	s4	s5	s6	s7	s8	s9
Angel	2	2	1	2	2	1	1	1	0
Gnome	1	2	2	2	1	0	2	1	3
Kid	2	3	1	1	1	2	1	1	1

Table 5: Number of true positive segments in each scene using merged segments from all models



(a) scene 1 (b) scene 2 (c) scene 3 (d) scene 4 (e) scene 5 (f) scene 6 (g) scene 7 (h) scene 8

Figure 4: Results for Angel: Segmentation (1^{st} row); True positive segments matched with model segments (2^{nd} row)



(a) scene 1 (b) scene 2 (c) scene 3 (d) scene 4 (e) scene 5 (f) scene 7 (g) scene 8 (h) scene 9

Figure 5: Results for Gnome: Segmentation (1^{st} row); True positive segments matched with model segments (2^{nd} row)

6. Conclusions and Future Work

This work demonstrates a different use for 3D interest operators for the goal of object recognition. In previous work, interest operators were executed as a prelude to the extraction of highly distinctive features, which were then used for matching. Only a small number of extracted interest points were passed to the feature extraction phase, with the vast majority being filtered out using non-maxima suppression. In contrast, in this work, the high repeatability of interest operators is leveraged to generate highly repeatable segments, which are then matched in place of features. The success of this approach in the current experiments demonstrates that the high repeatability of interest operators may be as important as the high descriptiveness of features, which has been the focus of most recent work in this area.

A limitation for the current experimental setup is the exhaustive comparisons of segments using 4PCS, which inevitability increases the processing time. In the future, a more efficient registration algorithm, possibly using fast indexing and shape hashing, will be developed. An unsupervised merge will also replace the supervised merge in a real object recognition system. It is also desirable to extend the current test to include a larger database of objects with various geometric properties, and compare with state-of-the art LSDs approaches.

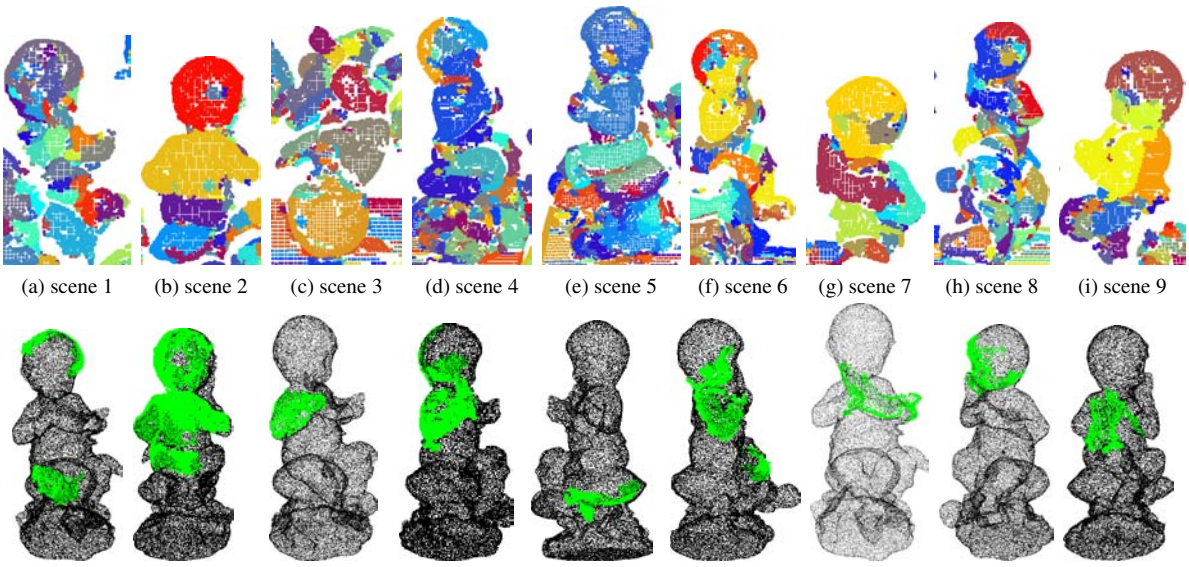
References

[1] D. Aiger, N. J. Mitra, and D. Cohen-Or. 4-points congruent sets for robust surface registration. ACM Transactions on Graphics, 27(3):#85, 1–10, 2008.

[2] P. Besl and R. Jain. Segmentation through variable-order surface fitting. IEEE Trans. Pattern Analysis and Machine Intelligence, 10(2):167–192, 1988.

[3] S. Bhandarkar and A. Siebert. Integra-an integrated approach to range image understanding. IEEE Trans. Pattern Recognition and Artifical Intelligence, 6(5):913 – 953, 1992.

[4] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image



(a) scene 1 (b) scene 2 (c) scene 3 (d) scene 4 (e) scene 5 (f) scene 6 (g) scene 7 (h) scene 8 (i) scene 9

Figure 6: Results for Watermelonkid: Segmentation (1^{st} row); True positive segments matched with model segments (2^{nd} row)

analysis and automated cartography. Communications of the ACM, 24:381–395, 1981.

[5] Y. Ioannou. Automatic urban modelling using mobile urban lidar data. Master's thesis, Queen's University, 2010.

[6] A. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. IEEE Trans. Pattern Analysis and Machine Intelligence, 21(3):433–449, May 1999.

[7] E. Kalogerakis, A. Hertzmann, and K. Singh. Learning 3d mesh segmentation and labeling. ACM Transactions on Graphics, 29(4):259–289, July 2010.

[8] V. Kolmogorov and R. Zabih. What energy function can be minimized via graph cut. IEEE Trans. Pattern Analysis and Machine Intelligence, 26(2):147 – 159, 2004.

[9] J. Lam and M. Greenspan. On the repeatability of 3d point cloud segmentation based on interest points. In CRV 2012: 9^{th} Canadian Conference on Computer and Robot Vision, Toronto, Ontario, Canada, May 2012.

[10] A. Mian, M. Bennamoun, and R. Owens. On the repeatability and quality of keypoints for local feature-based 3d object retrieval from cluttered scenes. Int. Journal of Computer Vision, 2009.

[11] G. Mori, S. Belongie, and J. Malik. Efficient shape matching using shape contexts. IEEE Trans. Pattern Analysis and Machine Intelligence, 27(11):1832–1837, Nov 2005.

[12] L. Shang and M. Greenspan. Real-time object recognition in sparse range images using error surface embedding. In International Journal of Computer Vision. Springer Netherlands, August 2009. ISSN: 09205691 (Print) 15731405 (Online).

[13] J. Shotton. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. International Journal of Computer Vision, 81(1):2–23, 2007.

[14] I. Sipiran and B.Bustos. A robust 3d interest points detector based on harris operator. In Eurographics workshop on 3D object retrieval (3DOR), 2010.

[15] S.Salti, F. Tombari, and L. Stefano. A performance evaluation of 3d keypoint detectors. In 3DIMPVT 2011: The First Joint 3DIM/3DPVT Conference on 3D Imaging, Modeling, Processing, Visualization, Transmission, Hangzhou, China, 2011.

[16] Y. Sun, J. Paik, A. Koschan, D. Page, and M. Abidi. Point fingerprint: A new 3-D object representation scheme. IEEE Trans. Systems, Man, and Cybernetics, 33(4):712–717, Aug 2003.

[17] B. Taati and M. Greenspan. Local shape descriptor selection for object recognition in range data. Computer Vision and Image Understanding: Special issue on 3D Imaging and Modelling, 115(5):681 – 694, 2011.

[18] R. Unnikrishnan and M. Hebert. Multi-scale interest regions from unorganized point clouds. In CVPR 2008, 2008.

[19] X. Yu, T. Bui, and A. Krzyzak. Robust estimation for range image segmentation and reconstruction. IEEE Trans. Pattern Analysis and Machine Intelligence, 16(5):530–538, May 1994.