CUBOIDS DETECTION IN RGB-D IMAGES VIA MAXIMUM WEIGHTED CLIQUE

Han Zhang[1], *Xiaowu Chen*[1*], *Yu Zhang*[1], *Jia Li*[1,†], *Qing Li*[2], *Xiaogang Wang*[1]

[1]State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University

[†]International Research Institute for Multidisciplinary Science, Beihang University

[2]Beijing Key Laboratory of Information Service Engineering, Beijing Union University

ABSTRACT

Cuboid detection is an essential step for understanding 3D structure of scenes. As most of indoor scene cuboids are actually objects, we propose in this paper an object-based approach to detect 3D cuboids in indoor RGB-D images. The proposed approach is learning-free and can handle general object classes rather than a limited pre-defined category set. In our approach, we first apply an extended version of the CPMC framework to generate a set of segment hypotheses, and fit a set of cuboid candidates. Given the candidate set, we select several cuboids that can provide plausible interpretations of the images by solving a Maximum Weighted Clique (MWC) problem. With this formulation, a set of ranked mid-level representations of the input image is obtained, and are further re-ranked by Maximal Marginal Relevance (MMR) measure to improve their diversity. Experimental results on NYU-V2 dataset shows that our method significantly outperforms the state-of-the-art, and shows impressive results.

Index Terms— Cuboid detection, scene understanding, depth image, maximum weighted clique.

1. INTRODUCTION

Identifying 3D structure of objects in images is an important task in computer vision. These structures, sometimes in the form of 3D cuboids, are crucial to be understood for indoor robots to interact with the environment successfully. In the last years, the literature focuses on how to reliably detect such 3D cuboid structures for monocular imagery [1, 2, 3, 4, 5]. However, it is extremely difficult, which is partly due to certain unique challenges that the monocular images present, including severe object occlusion, unapparent boundaries, diversity in scene type and a lack of distinctive features [6].

To address this problem, the research community resorts to utilizing depth sensors to complement the ambiguous 2D information, making detection of 3D cuboid structures in indoor images more tractable. In the relevant studies, 3D cuboid candidates are first fitted from RGB-D superpixel pairs [7] or depth-sensitive segment proposals [8], and then selected considering their physical and contextual relationships.

*corresponding author (email: chen@buaa.edu.cn)

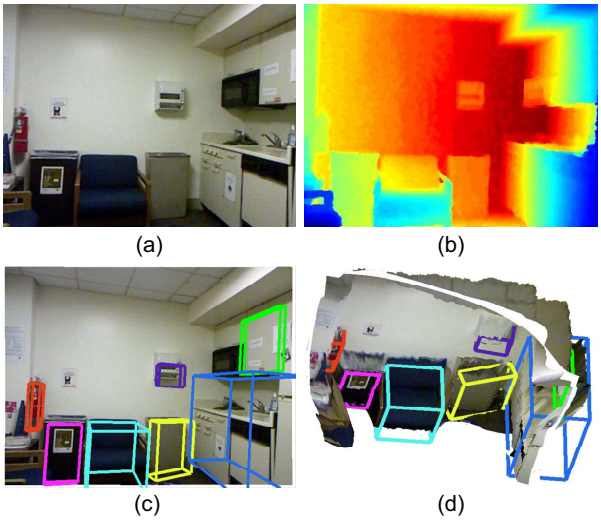

Fig. 1. The motivation of our method. Given an image (a) with its depth (b), our method detects cuboid structures (c) in the image. Results from another perspective are shown in (d).

In general, a good cuboid detector for indoor scene images should exhibit the following two merits: 1) it is intuitively to be object-based, namely, the detected 3D cuboids should capture indoor objects but not any regions with orthogonal surfaces; 2) it is required to handle general object classes instead of a limited pre-defined category set. Although existing studies [7, 8] have achieved promising advance in one aspect, they might neglect the other. Moreover, these approaches both produce a single interpretation of the indoor scene, which may occasionally be imperfect and miss some important objects.

In this paper, we propose an object-based approach to detect 3D cuboids in RGB-D images, as shown in Fig. 1. The proposed approach is learning-free and is able to generate a set of various mid-level representations of the input image. The pipeline of our framework is summarized in Fig. 2. In our framework, we first apply an extended version of the CPMC framework [9] on the input image to generate a set of object hypotheses, based on which cuboid candidates are fitted. The selection of cuboid candidates which produces plausible interpretations of the images is formulated as a *Maximum*

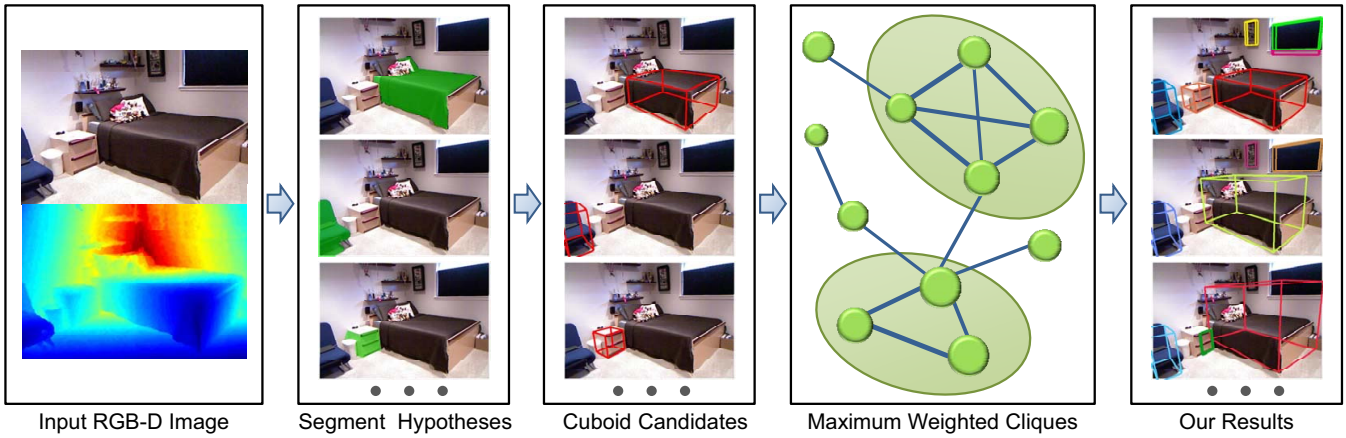

Fig. 2. The overview of our method (best viewed in color). Starting from segment hypotheses extraction, our method fits a series of candidate cuboids for the extracted segments. After that, a ranked set of cuboid representation of the input scene is obtained through solving a Maximum Weighted Clique (MWC) problem.

Weighted Clique (MWC) problem. As the obtained ranking may put similar results in adjacent positions, we further adopt *Maximal Marginal Relevance (MMR)* technique to diversify the ranking. Extensive experiments conducted on the challenging NYU-V2 RGB-D dataset [10] suggest, that such a strategy can achieve impressive improvement compared with the state-of-the-art general cuboid detector [7], and obtain similar results with those of the learning-based algorithm [8] that is optimized on the dataset with training images. In summary, the main contributions of this paper are:

1) We propose an object-based framework for detecting 3D cuboids of objects in RGB-D image. It is learning free, and boosts performance of cuboid detection significantly on challenging public dataset.

2) We formulate selection of the cuboid candidates as a *maximum weighted clique* problem, which can be efficiently solved and has the flexibility to produce a set of mid-level representations of the RGB-D image.

2. RELATED WORK

Understanding 3D scene structures in images is an established topic for both monocular and RGB-D data. In the rest of this section, we briefly review the literature in both aspects.

A series of approaches on 3D cuboid detection has been proposed for RGB images [3, 2, 5, 11]. Hedau et al. [3] generated cuboid hypotheses by sliding a 3D cuboid detector. Lee et al. [2] utilized volumetric constraints of the physical world to generate 3D parametric cuboid representation of indoor scenes. Zhao et al. [5] proposed a stochastic grammar model over the function-geometry-appearance (FGA) hierarchy for scene parsing. Choi at al. [11] developed a model which combines object detection, layout estimation and scene classification. They further introduced a 3D Geometric Phrase Model to enhance the semantic and geometric relationships between objects that frequently co-occur. However, due to the ambiguity of monocular setting, these methods rely on knowledge of indoor layouts and require complicated inference to obtain reliable results.

Recently, depth information is explored by a line of works on indoor scene analysis such as scene labeling and object detection [12, 7, 8]. For example, Silberman et al. [13] incorporated appearance potential and statistical location priors into a CRF model to infer pixel-wise labeling. Zhang et al. [12] jointly estimated the layout of rooms as well as clutters present in the scene using both depth and appearance information. Jiang et al. [7] constructed cuboid candidates using superpixel pairs extracted on the RGB-D image, and then formulated the multiple cuboid matching as a linear integer program. Lin et al. [8] adopted an extended version of CPMC [9] to 3D space to generate candidate cuboids, and trained a random field model that integrated information from different sources for classification. In these methods, Jiang et al. [7] is the most similar work to ours. Compared with [7], our method detects object-like cuboid candidates instead of the two facet structures extracted on over-segmented RGB-D images. Moreover, by solving a maximum weighted clique problem, our method provides various interpretations of the input image, in contrast to the single representation of [7].

3. CUBOID CANDIDATE GENERATION

As shown in Fig. 2, our method starts from generating a set of object cuboid candidates for the input RGB-D image. This is achieved by fitting the segment hypotheses produced by an extended version of [9] incorporating depth information.

Formally, given the input image with a set of pixels $\mathcal{V}$,

CPMC attempts to minimize the following objective on the binary pixel labels $\mathbf{x} = \{x_i\}_{i=1}^{|\mathcal{V}|} \in \mathbb{B}^{|\mathcal{V}|}$:

$$\min_{\mathbf{x}} E(\mathbf{x}, \lambda) = \sum_{i \in \mathcal{V}} D(x_i, \lambda) + \sum_{(i,j) \in \mathcal{E}} V(x_i, x_j), \quad (1)$$

where λ is an uniform offset that controls the foreground bias, the unary term $D(x_i, \lambda)$ defines the cost of assigning the ith pixel to foreground, and the pairwise term $V(x_i, x_j)$ penalizes assignments of neighboring pixels, with $\mathcal{E}$ denoting the set of neighboring pixel pairs. Definition of the unary term D is similar with that of [8], which incorporates the depth information. In addition, to explore depth boundaries, we augment the pairwise term V in the objective (1) by

$$V(x_i, x_j) = \begin{cases} g(i,j), \text{ if } x_i \neq x_j \\ 0, \text{ otherwise} \end{cases}, \text{ where}$$
$$g(i,j) = e^{-\frac{1}{\sigma^2} \max(B_I(i), B_I(j), B_D(i), B_D(j))}, \quad (2)$$

where $B_I(i)$ and $B_D(i)$ are the ith pixel's boundary confidence produced by the contour detector [14] computed on the color image and the depth image, respectively.

Varying the offset λ generates a set of candidate object segments. Using the ranking mechanism of CPMC detector, we select the top $K = 80$ candidates and discard the rest. After that, we utilize the method of [8] to fit a set of 3D cuboid candidates, denoted as $\mathbb{S} = \{\mathcal{C}_i\}_{i=1}^{K}$.

4. CUBOID CANDIDATE SELECTION

4.1. Formulation as Maximum Weighted Clique

Given the set of cuboid candidates $\mathbb{S}$, the target of our approach is to find a subset of cuboids to best interpret the configurations of the objects in the RGB-D image. Denote a set of binary variables $\mathbf{a} = \{a_i\}_{i=1}^{K}$, where a_i indicates the selection of the ith cuboid candidate. Intuitively, the selected cuboids should be physically reliable themselves and exhibit plausible mutual relationships. Thus, we propose the following optimization problem:

$$\max_{\mathbf{a}} a_i \boldsymbol{\lambda}_u^{\mathrm{T}} \sum_{i=1}^{K} \Phi(\mathcal{C}_i) + a_i a_j \boldsymbol{\lambda}_p^{\mathrm{T}} \sum_{i=1}^{K} \sum_{j \in \mathbb{N}_i} \Psi(\mathcal{C}_i, \mathcal{C}_j),$$
$$\text{s.t. } a_i \in \{0,1\}, \forall i \in \{1,2,...K\}, \quad (3)$$

where Φ and Ψ are unary and pairwise potentials of the candidate cuboids, respectively, the set $\mathbb{N}_i$ indexes the neighboring cuboids of the ith cuboid. Note that the unary and pairwise potentials are both column vectors that concatenate various features, weighted by parameters $\boldsymbol{\lambda}_u$ and $\boldsymbol{\lambda}_p$.

The objective (3) is a binary quadratic optimization problem over a few hundreds of variables, which are typically handled by mixed integer solvers. However, since the problem is NP-hard, the general branch-and-bound solution can be very slow in some cases. In fact, in practice we find that when the potentials of cuboid candidates are similar to each other, the running time may exceed an hour for a single input image using the state-of-the-art solver Gurobi [15]. Thus, inspired by [16], we adopt a heuristic algorithm to solve (3). Specifically, consider a graph where the vertices are the candidate cuboids, and the edges are connected with every two cuboids. In such case, the solution of (3) can be treated as a *Maximum Weighted Clique (MWC)* sampled from the graph. Our formulation is the most general case of MWC with weights defined both on vertices and edges. Many heuristic algorithms are proposed for solving a general MWC, among which we adapt the FGTiling procedure proposed in [16], due to its efficiency and the flexibility to generate a set of ranked cliques.

The key idea of FGTilling is to extract many ranked weighted maximal cliques seeded from different vertexes. We briefly describe the clique extraction procedure here and refer the reader to [16] for more details. To sum up, it consists of the initialization step and refinement step.

Initialization. For the ith vertex, we keep a set S_i^0 which represents a potential maximal clique containing the vertex. The set grows greedily by adding another vertex whose represented cuboid has less than 10% volume overlapped with cuboids in the current set at one time. The set keeps growing until inclusion of any new vertex is infeasible. At last, we get a maximal clique starting at the given vertex. After initialization, we have N node sets in total, one for each vertex.

Refinement. Given an initial set S_i^0, We refine it through the following steps. For each vertex absence in S_i^0 that has low overlap with the ith vertex, we try to add it into the set while removing the vertexes already included by the set that have more than 10% volume overlapped with the added vertex. The modified set is further expanded to a maximal clique using the greedy algorithm of the initialization step. If the modified set S_i' increases the objective compared with the initial set S_i^0, we set $S_i^0 = S_i'$, otherwise we keep S_i^0 unchanged. For each initial set, we repeat the step until convergence.

Diversify ranking. In order to prevent minor variations between adjacently ranked results, we adopt *Maximal Margin Relevance (MMR)* to diversify the ranking. Specifically, denote $\mathcal{S}_1$ and $\mathcal{S}_2$ the two different cliques, $\mathcal{P}_1$ and $\mathcal{P}_2$ the projected regions of the cuboids in $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, and $\mathbf{a}_1$ and $\mathbf{a}_2$ the corresponding binary selection vectors, the redundancy measure is defined as follows

$$sim(\mathcal{S}_1, \mathcal{S}_2) = \frac{|\mathcal{P}_1 \cap \mathcal{P}_2|}{|\mathcal{P}_1 \cup \mathcal{P}_2|} + \frac{\|\mathbf{a}_1 \cdot \mathbf{a}_2\|_0}{\|\mathbf{a}_1 + \mathbf{a}_2\|_0}, \quad (4)$$

where $|\mathcal{P}|$ denotes the area of region $\mathcal{P}$, $\|\cdot\|_0$ is the ℓ_0-norm. We refer the reader to [17] for more details of the ranking procedure in MMR.

On a single-core 3.1GHz machine, the extraction procedure costs 30 seconds on average. Note that it can be easily paralleled since each clique is extracted independently.

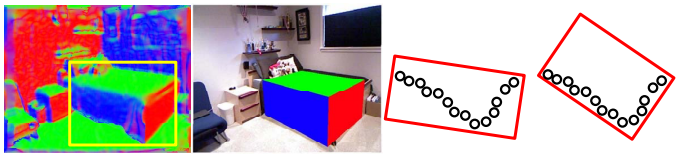

Fig. 3. (a): Surface normal computed with least square plane.(red, blue and green denotes three orthogonal directions) (b): Surface normal computed with the recovered cuboid. (c): An recovered cuboid with minimum volume. (d): A fitted cuboid that better aligns the surface.

4.2. Unary Potential

In the unary term, we exploit features that encode the quality of a cuboid, including four image-based features and four depth-related features. Four each cuboid, the first set of image features consist of the following ones:

The objectness of cuboid on its original image segment predicted by [9]. This feature favors choosing cuboids covering object-like image regions.

The covered area computed as the intersection area of the cuboid's CPMC segment with its projected region on the image. It favors choosing the cuboid with a larger projecting area, since small cuboids are often noisy.

The tightness which measures how well a fitted cuboid bounds its segment. This is given by the overlap of its segment area and its projected region on the image under the *intersection-over-union overlap* metric.

The boundary distance which encodes how well the projected region of the cuboid matches the image boundaries. It is calculated as the average distance between the projected boundary and the edges detected by Canny detector [18].

The depth-related features include:

The credibility of surface. We compute 3D surface normals at each pixel by sampling its surrounding pixels within a depth threshold and then fitting a least square plane, as shown in Fig. 3 (a). The normal is denoted as (N_x, N_y, N_z) at the location (x, y, z). Similarly, we also extract the surface normal of visible pixels of the recovered cuboid candidate, denoted as $(N_x^{'}, N_y^{'}, N_z^{'})$ shown in Fig. 3 (b). The credibility of the pixel is computed as $N_x \cdot N_x^{'} + N_y \cdot N_y^{'} + N_z \cdot N_z^{'}$. We average over the credibility of all the pixels shared by both the CPMC segment and the projected region of the cuboid as the surface credibility feature.

The minimum surface distance. The cuboid fitted with minimum volume may not have correct orientation (Fig. 3 (c)). A good fitting should be the case that most pixels of a segment locate on the surface of the recovered cuboid in 3D space (Fig. 3 (d)). Therefore, we compute the distance of each segment pixel to six surfaces of the recovered cuboid, and assign each pixel the minimum distance. The minimum surface distance is computed by averaging per-pixel distances.

The orientation of the cuboid. According to the Manhattan world assumption, visible surfaces should be along one of three orthogonal. We detect walls and floor through 3D Hough transform [19], which produces the three orthogonal. Denote the minimum angle between the given cuboid and detected walls as θ, we use e^{θ} to evaluate the orientation.

The cuboid volume, which is the cuboid's coverage in 3D space. It encourages choosing cuboid of large volume.

The unary potential concatenates the 8 features and is normalized with zero-mean and standard deviation.

4.3. Pairwise Potential

The pairwise potential considers the mutual region properties (e.g. intersection, occlusion, support) in both rgb and depth domain among cuboids, including:

The cuboid intersection, defined as the larger one of the two possible intersection ratios between two given cuboids.

The cuboid occlusion. From the camera view, when an object is mostly occluded by another object, the cuboid of the occluded object is unlikely to be a true detection and should be considered unreliable. Concerning this, we introduce the occlusion term given by $S(H \bigcap Q)/S(Q)$, where H and Q are the projected regions of the closer and farther cuboid to the camera, respectively. $A(\cdot)$ is the area (i.e. the number of pixels) of a region. We define the distance of the cuboid to the camera following [7].

The minimum support distance is used to exploit the supporting interactions between the cuboids and the environment. It is computed by top-down support reasoning on simple geometric rules of the recovered cuboids, e.g. the rough size. Specifically, given two cuboids $\mathcal{C}_1$ and $\mathcal{C}_2$, if $\mathcal{C}_1$ is on top of $\mathcal{C}_2$ and the ground projection of $\mathcal{C}_1$ is contained within that of $\mathcal{C}_2$, we identify that $\mathcal{C}_1$ is on top of $\mathcal{C}_2$, and define the support distance from $\mathcal{C}_1$ to $\mathcal{C}_2$ as the distance from the lower surface of $\mathcal{C}_1$ to the upper surface of $\mathcal{C}_2$. The minimum support distance of $\mathcal{C}_1$ is defined as the minimum of its distance to walls, floors and that to other cuboids supporting it.

Like the unary potential, the pairwise potential concatenates the three features and is normalized.

5. EXPERIMENTS

We perform the evaluation of the proposed framework on the testing set of the NYU-V2 RGB-D dataset [10], which consists of 645 image scenes. We use the ground-truth cuboids by [8]. On this dataset, we directly compare with the state-of-the-art cuboid detection methods [7] and [8]. We also analyze the influence of the used potentials with respect to the detection accuracy. We fix the potential weights $\boldsymbol{\lambda}_u^{\mathrm{T}} = (10, 15, 2, -0.25, 1.8, -1, -0.7, 30)$ and $\boldsymbol{\lambda}_p^{\mathrm{T}} = (-20, -15, -25)$ across all the experiments. These parameters are set through grid search and fine-tuned empirically on a validation set of 50 images.

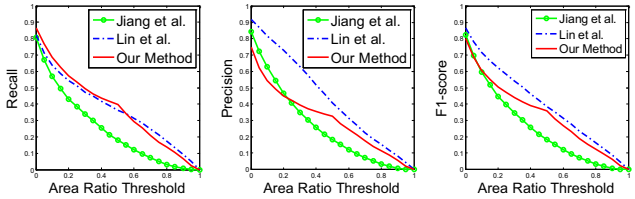

Fig. 4. Comparison with Jiang et al. [7] and Lin et al. [8] in terms of overlap ratio threshold vs. F1-score curve.

Table 1. F1-scores on on different scenes of NYU-V2 dataset.

	Jiang et al. [7]	Lin et al. [8]	Ours
kitchen	0.178	0.359	0.375
office	0.169	0.357	0.375
bathroom	0.162	0.366	0.315
livingroom	0.154	0.406	0.339
bedroom	0.244	0.442	0.384
bookstore	0.0	0.0	0.0
classroom	0.083	0.323	0.363
homeoffice	0.162	0.387	0.357
playroom	0.239	0.308	0.306
receptionroom	0.158	0.546	0.350
study	0.108	0.270	0.302
diningroom	0.171	0.363	0.331
other	0.094	0.295	0.340
overall	0.182	0.385	0.357

Detection performance. We evaluate the recall, precision and F1-score of all the competitors. Note that a detected cuboid is said to be true if there exists a ground-truth cuboid that overlaps with it over a threshold. We vary the threshold to produce the curves in Fig. 4. The curves show that our method significantly outperforms [7] on all indicators. Particularly, our method achieves the highest recall, proving that the object-based assumption of the cuboid candidates can retrieve more true detections. Compared with [8], our method performs slightly worse. However, it is predictable since [8] utilizes class-specific segment features, geometric properties as well as contextual relations, which are learned from a training set consisting of hundreds of training images. Conversely, our method is category-independent and has no training phase. To exploit the efficacy of all the methods on different scene categories, we show the F1-scores (using the overlap ratio 50% by convention) of the three methods on 13 scene categories in Table 1. The comparison shows that our method consistently outperforms [7], and even surpasses [8] on *office*, *classroom*, *study* and *kitchen* and *other*. Looking into these categories, we see that the objects in these scenes have large intra-class variations, thus the trained potentials in [8] may not characterize the categories well. In contrast, our method is generic and can obtain higher recall on these object categories.

Table 2. Detection performance of our method under different combinations of potentials.

	Recall	Precision	F1-score
objectness only	0.308	0.167	0.217
2D unaries	0.261	0.213	0.235
2D unaries + normal	0.339	0.225	0.271
2D + 3D unaries	0.306	0.293	0.299
all combined	0.397	0.325	0.357

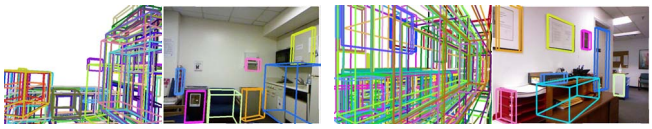

Fig. 5. Top 100 cuboid candidates and the cuboids detected by the proposed framework shown on the rgb image.

Fig. 6. Visual comparison of our method (top row) with Jiang et al. [7] (bottom row).

Influence of potentials. The table 2 summarizes the detection performance of our method by isolating different features in the used potentials. We observe that even using objectness only, we can already get a higher F1-score (0.217 in Table 2) than that of [7] (0.182 of [7] in Table 1). Moreover, combining the 2D features and the 3D features can improve the performance significantly, demonstrating that the two sets of features are complementary. By further adding into pairwise potentials, we reach the highest accuracy.

Qualitative results. We show two examples of the generated cuboid candidates and detection results in Fig. 5. The proposed method generates about 7 cuboids per scene on average. We also show the visual comparisons of our method compared with [7] in Fig. 6. In general, we find that [7] performs well when the two facets of objects are salient in the images. However, when dealing with occlusion and background clutters, our method can obtain more reliable detections.

With the MWC framework, our method is able to generate a series of plausible interpretations for the image, as shown in Fig. 7. More representative results are shown in Fig. 8.

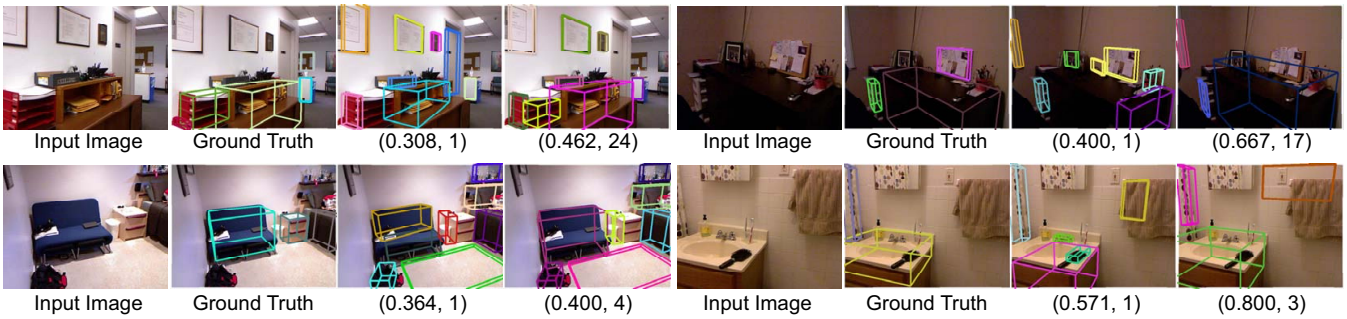

Fig. 7. Various interpretations of the input images with corresponding F1-scores and ranks generated by our method.

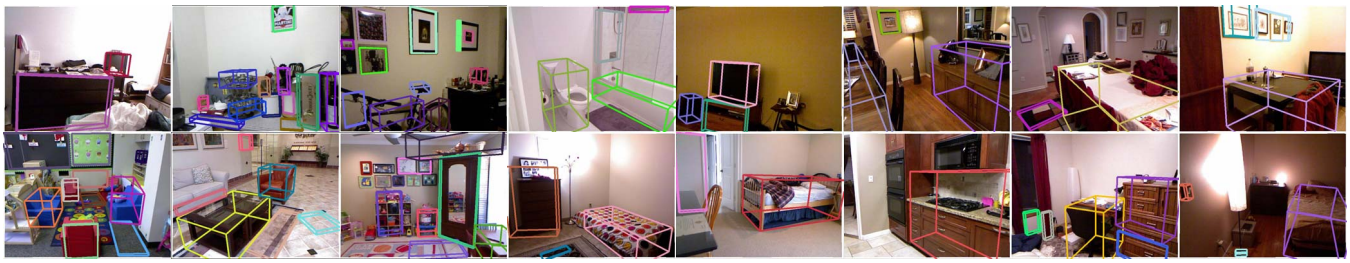

Fig. 8. More representative results generated by our method on NYU-V2 dataset.

6. CONCLUSION AND DISCUSSION

In this paper, we present an object-based approach to detect cuboids in RGB-D images. In this approach, cuboid candidates are fitted from object segment proposals. The selection of cuboid candidates is formulated as a Maximum Weighted Clique problem and solved heuristically, which generates a set of ranked interpretations for the input image. Experimental results demonstrate that the proposed method boosts cuboid detection performance on NYU-V2 dataset. In the future, we will try to extend the framework to generate consistent cuboid detection results for RGB-D videos.

Acknowledgement. We would like to thank the reviewers for their valuable feedback. This work is partly funded by NSFC (61325011) & (61421003), 863 program (2013AA013801), SRFDP (20131102130002), and State Key Laboratory of Virtual Reality Technology and Systems, Beihang University.

7. REFERENCES

[1] A. Gupta, A. A. Efros, and M. Hebert, "Blocks world revisited: Image understanding using qualitative geometry and mechanics," in *ECCV*. 2010.

[2] A. Gupta, M. Hebert, T. Kanade, and D. M. Blei, "Estimating spatial layout of rooms using volumetric reasoning about objects and surfaces," in *NIPS*, 2010.

[3] V. Hedau, D. Hoiem, and D. Forsyth, "Thinking inside the box: Using appearance models and context based on room geometry," in *ECCV*. 2010.

[4] V. Hedau, D. Hoiem, and D. Forsyth, "Recovering free space of indoor scenes from a single image," in *CVPR*, 2012.

[5] Y. Zhao and S. Zhu, "Scene parsing by integrating function, geometry and appearance models," in *CVPR*, 2013.

[6] X. Ren, L. Bo, and D. Fox, "Rgb-(d) scene labeling: Features and algorithms," in *CVPR*, 2012.

[7] H. Jiang and J. Xiao, "A linear approach to matching cuboids in rgbd images," in *CVPR*, 2013.

[8] D. Lin, S. Fidler, and R. Urtasun, "Holistic scene understanding for 3d object detection with rgbd cameras," in *ICCV*, 2013.

[9] J. Carreira and C. Sminchisescu, "CPMC: Automatic object segmentation using constrained parametric min-cuts," *PAMI*, vol. 34, no. 7, pp. 1312–1328, 2012.

[10] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from rgbd images," in *ECCV*. 2012.

[11] W. Choi, Y. Chao, C. Pantofaru, and S. Savarese, "Understanding indoor scenes using 3d geometric phrases," in *CVPR*, 2013.

[12] J. Zhang, C. Kan, A. G. Schwing, and R. Urtasun, "Estimating the 3d layout of indoor scenes and its clutter from depth sensors," in *ICCV*, 2013.

[13] N. Silberman and R. Fergus, "Indoor scene segmentation using a structured light sensor," in *ICCV Workshops*, 2011.

[14] M. Leordeanu, R. Sukthankar, and C. Sminchisescu, "Efficient closed-form solution to generalized boundary detection," in *ECCV*. 2012.

[15] Inc. Gurobi Optimization, "Gurobi optimizer reference manual," 2014.

[16] A. Ion, J. Carreira, and C. Sminchisescu, "Image segmentation by figure-ground composition into maximal cliques," in *ICCV*, 2011.

[17] J.Carbonell and J.Goldstein, "The use of mmr, diversity-based reranking for reordering documents and producing summaries," in *SIGIR*, 1998.

[18] J. Canny, "A computational approach to edge detection," *PAMI*, , no. 6, pp. 679–698, 1986.

[19] D. Borrmann, J. Elseberg, and K. Lingemannand A. Nüchter, "The 3d hough transform for plane detection in point clouds: A review and a new accumulator design," *3D Research*, vol. 2, no. 2, pp. 1–13, 2011.