

Using a Visual Cuboid Classifier to Match Cuboids in RGBD images

Julia Deeb | Dr. Abhinav Gupta

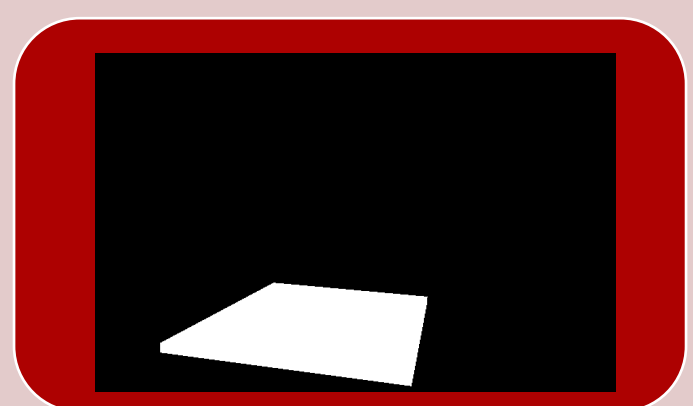
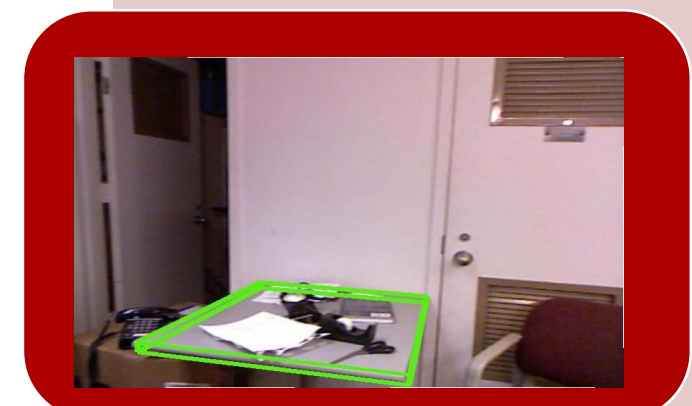
Problem

The problem can be divided into two parts:

- 1) Can we build a visual cuboid classifier?
- 2) Can such a classifier be used to successfully match cuboids in RGBD images?

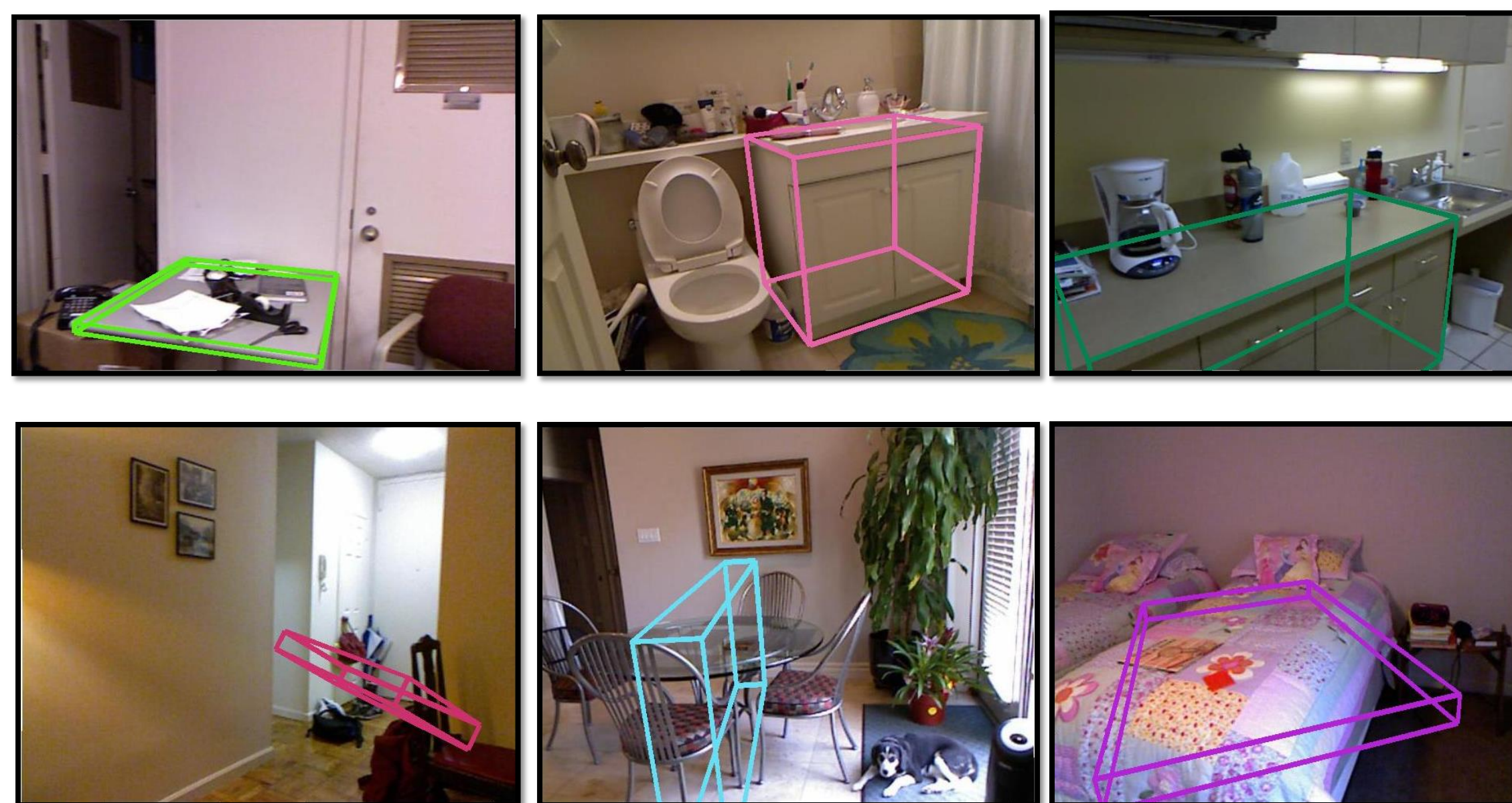
Constructing the Data Set

The data set was constructed from about **500 positive** ground truth cuboids and about **1500 negative** cuboids. Each cuboid was projected onto the image and a mask was extracted. The mask was then applied back onto the image and features were extracted from that region. The features extracted for training were **color** and **visual texture**.



Extract color
and texture
features

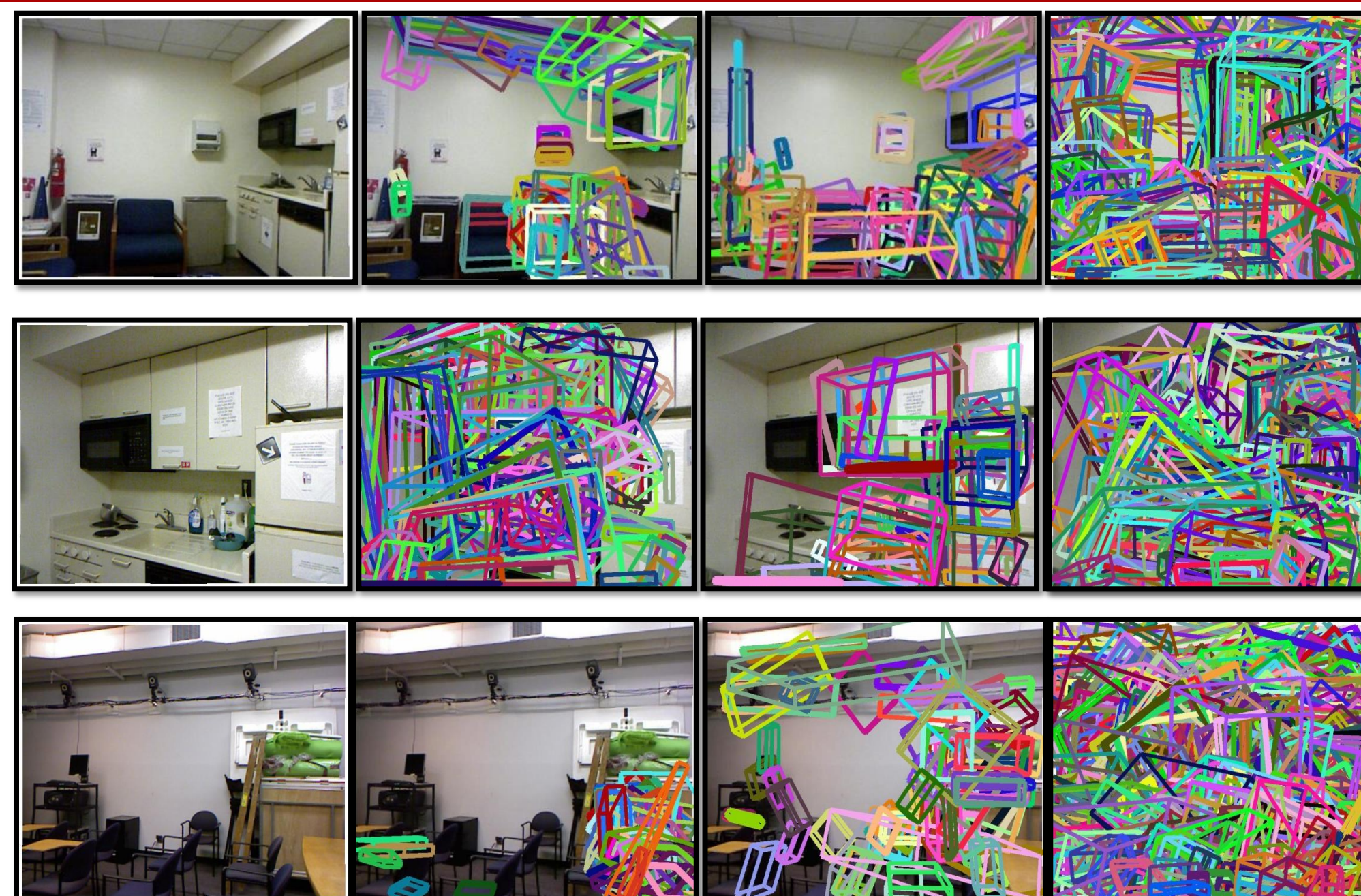
Positive VS Negative Cuboids



TOP ROW: Ground truth positive cuboids, BOTTOM ROW: negative cuboids

The negative cuboids are often visually inconsistent and geometrically skewed. Whereas, the positive cuboids tend to be consistent in color and/or texture.

SVM Cuboid Results



FROM LEFT TO RIGHT: original RGB image, SVM positive results, top 100 results from [Jiang '13], SVM negative results

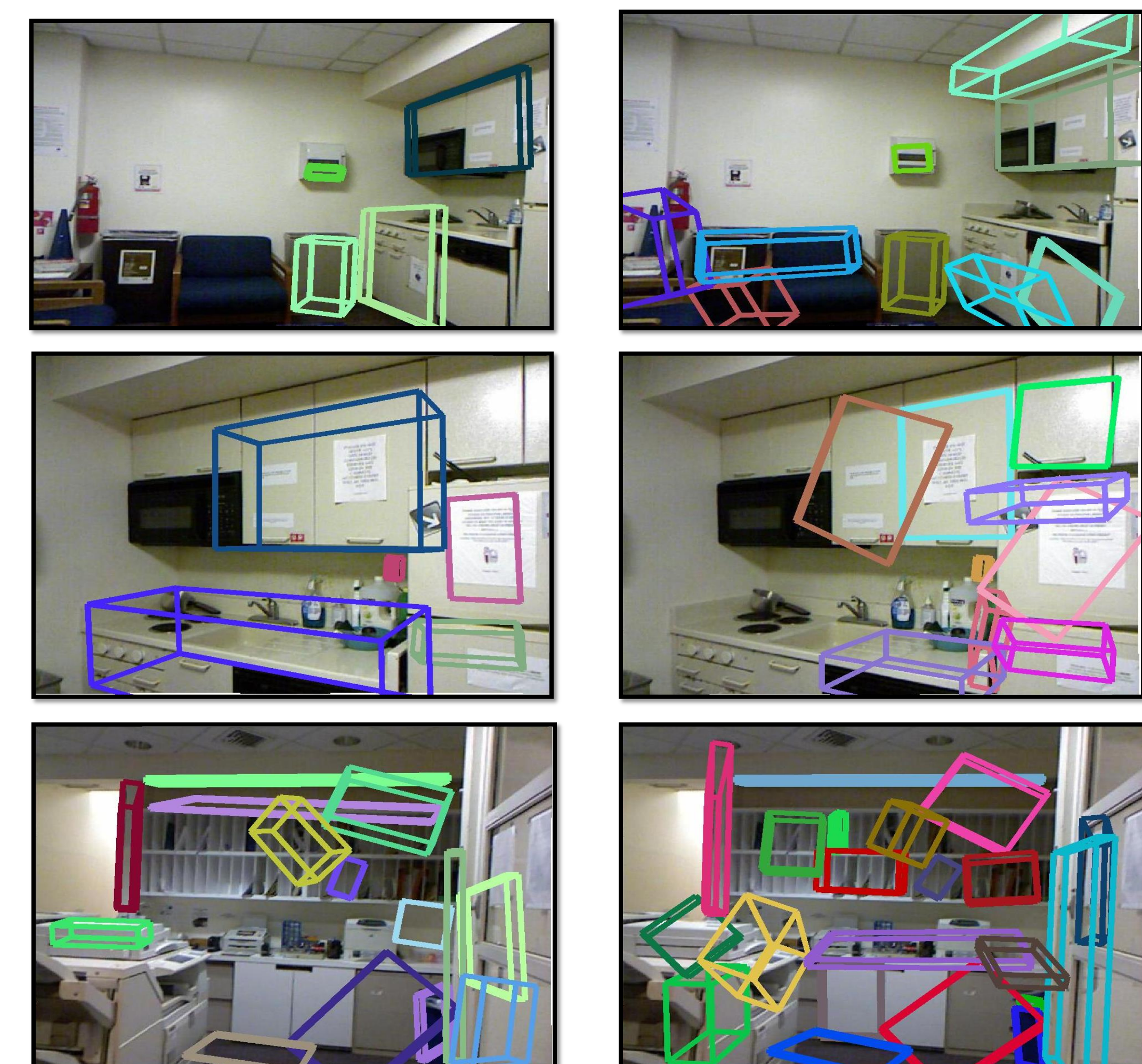
Matching Cuboids

To match the actual cuboids in the 2D image, we used the analytical method described in [Jiang '13]. However, instead of ranking the cuboids based on volume exclusion and surface cover and then optimizing over the top 100, we are instead **optimizing over the positive results** returned by the SVM cuboid classifier.

Selected Results



Comparison of Selected Results



LEFT COLUMN: our results , RIGHT COLUMN: results from [Jiang '13]

We can compare our results to the results of [Jiang '13]. Our results typically have less cuboids to explain the picture, but often more of the cuboids are 'correct'.

Conclusions

The visually-based SVM classifier does appear to be quite successful on many of the images. However, there is a quite number of false positive and false negatives, especially in images with high occlusions and/or high clutter. This is perhaps due to the fact that the cuboids in this scenario are likely to not be visually consistent and therefore are indistinguishable from the non-cuboids. Perhaps, in these scenarios, accounting for geometric features like in [Xiao '12] may increase the number of 'correct' cuboids identified.

References

- Lee, David C., et al. "Estimating Spatial Layout of Rooms using Volumetric Reasoning about Objects and Surfaces." *NIPS*. Vol. 1. No. 2. 2010.
- Jiang, Hao, and Jianxiong Xiao. "A Linear Approach to Matching Cuboids in RGBD Images." *CVPR*, 2013.
- Xiao, Jianxiong, Bryan Russell, and Antonio Torralba. "Localizing 3D cuboids in single-view images." *Advances in Neural Information Processing Systems*. 2012.