Incremental Model Selection for Detection and Tracking of Planar Surfaces

Johann Prankl[1]
prankl@acin.tuwien.ac.at
Michael Zillich[1]
zillich@acin.tuwien.ac.at
Bastian Leibe[2]
leibe@umic.rwth-aachen.de
Markus Vincze[1]
vincze@acin.tuwien.ac.at

[1] Automation and Control Institute
Vienna University of Technology
Austria
[2] UMIC Research Centre
RWTH Aachen
Germany

Abstract

Man-made environments are abundant with planar surfaces which have attractive properties and are a prerequisite for a variety of vision tasks. This paper presents an incremental model selection method to detect piecewise planar surfaces, where planes once detected are tracked and serve as priors in subsequent images. The novelty of this approach is to formalize model selection for plane detection with Minimal Description Length (MDL) in an incremental manner. In each iteration tracked planes and new planes computed from randomly sampled interest points are evaluated, the hypotheses which best explain the scene are retained, and their supporting points are marked so that in the next iteration random sampling is guided to unexplained points. Hence, the remaining finer scene details can be represented. We show in a quantitative evaluation that this new method competes with state of the art algorithms while it is more flexible to incorporate prior knowledge from tracking.

1 Introduction

Regardless whether indoor or outdoor, man-made environments are abundant with planar structures. Walking on a planar street in a city we can observe many planar surfaces. Entering a supermarket even more objects consisting of planar structures can be found, such as box-shaped packaging located on shelves surrounded by a planar floor and walls.

Because of their attractive properties, planes are used in various vision tasks. Examples include application for camera self-calibration [25, 28] and for feature matching and grouping [7, 15]. Planes are also used for 3D reconstruction and scene analysis [1, 3, 6, 10, 18] and in robotics applications for obstacle detection [14], camera localization [19] and for object recognition [23].

Our overall goal is to build a cognitive robotic experimentation framework. The rationale behind our system is to enable human tutor driven learning-by-showing as well as completely automatic on-line model acquisition by the robot (More details about the cognitive

BMVC 2010 doi:10.5244/C.24.87

robotic framework can be found in [20]). Schindler et al. [24] use a model selection framework for multibody Structure-and-Motion estimation of image sequences. In contrast we use model selection to detect piecewise planar surfaces necessary to attach affordances such as graspable and stability. Our model is simpler but enables the robot to interact in more realistic environments. Therefore we developed an incremental model selection method, where planes once detected are tracked and serve as priors in subsequent images. This approach formalizes model selection for plane detection with Minimal Description Length (MDL), which is well-known for segmentation of range images [12] and robust object detection [11]. We describe plane hypotheses using the 2D projective transformation (homography) computed from four interest point pairs in two uncalibrated images. In an iterative scheme tracked planes and new planes computed from randomly sampled interest points are evaluated, the hypotheses which best explain the scene are retained and their supporting points marked so that in the next iteration random sampling is guided to unexplained points. Hence, the remaining finer scene details can be detected.

After a review of the related work, we give an overview of the approach in Section 2 and its core parts, namely how to select good plane hypotheses (Section 2.1) and the model selection scheme (Section 2.2). Section 3 describes the post-processing and Section 4 shows how to incorporate plane tracking. Finally, results of the experiments are shown in Section 5.

1.1 Related work

Various approaches for plane detection in uncalibrated image pairs exist. Most of them use a hypothesize-and-test framework. A popular method for detecting multiple models is to use the robust estimation method RANSAC [4], to sequentially fit the model to a data set and then to remove inliers. To generate plane hypotheses Vincent et al. [29] use groups of four points which are likely to be coplanar to compute the homography. To increase the likelihood that the points belong to the same plane they select points lying on two different lines in an image. In contrast Kanazawa et al. [9] define a probability for feature points to belong to the same plane using the Euclidean distance between the points. Both approaches use a RANSAC scheme, iteratively detect the dominant plane, remove the inliers and precede with the remaining interest points.

The success of the plane computation depends on the coplanarity of four matched points. Nicolas et al. [13] propose to use the knowledge of the first homography to guide the computation of further homographies and thus reduce the set of points/lines to three pairs. After the first iteration they use three points to generate hypotheses. The selection scheme is also iterative and RANSAC based. Piazzi et al. [22] also need only three corresponding points. They propose to use two cameras aligned to the same orientation to compute the normal vector to a plane. The normal vector is then used to cluster triangles. This approach does not use RANSAC, instead sequentially similar neighbouring normal vectors of Delaunay triangles are clustered. Lourakis et al. [16] first estimate the fundamental matrix and the epipoles. Then the homography is computed for each set of point and line feature and a voting scheme is applied. The homography with the highest number of votes is selected and refined using Least Median of Squares. This approach also detects a plane and removes the inliers in an iterative scheme.

More recent approaches concentrate on robust estimation of multiple structures. Toldo et al. [27] propose j-linkage, an approach based on random sampling and conceptual data representation. Each point is represented with the characteristic function of the set of random models that fit the point. Then agglomerative clustering is used to group points belonging to the same model. In [5] this method is used for the robust detection and matching of multiple planes. For hypothesis generation random sampling is used. To select the best hypotheses this approach clusters homographies. In [2] Chin et al. propose a novel Mercer kernel for the robust estimation problem which elicits the potential of two points to have emerged from the same underlying structure and thus can cope with more than 90% outliers. While random sampling is used to generate hypotheses, principal component analysis and spectral clustering are applied for robust fitting. The methods of Toldo et al. [27], Fouhey et al. [5] as well as Chin et al. [2] use clustering schemes and avoid to remove inliers and iterative detection of planes.

Given a fixed threshold to detect inliers incremental methods favour planes detected first over subsequent planes by greedily consuming features. Recently developed approaches overcome this drawback by treating hypotheses equally, but plane hypotheses have to be created independently of each other and thus it is not possible to restrict the search space, which leads to higher computational complexity. We propose model selection based on the MDL principle: Instead of creating all hypotheses at once, pruning models and then using model selection, we propose to embed model selection in an incremental scheme and thus guide randomized selection of interest points to compute more likely plane hypotheses. This allows us to avoid an additional hypotheses pruning step without decrease of performance. Finally, this formulation allows us to explicitly introduce priors, hence we can detect and track planes in one scheme which is not possible in any of the approaches described above.

2 Approach

We developed a method to detect multiple planes in image pairs. The idea is to embed Minimal Description Length (MDL) based model selection in an iterative scheme. Thus existing planes compete with newly created hypotheses to ensure that interest points are assigned to the best current available hypothesis. Additionally hypothesis generation can be guided to unexplained regions. This method avoids the bias towards dominant planes typical for iterative methods, and it limits the search space which leads to a faster explanation of the entire image in terms of piecewise planar surfaces.

Algorithm 1 Plane detection

$P \leftarrow 0$, $T \leftarrow 0$
$k \leftarrow 0$, $\varepsilon \leftarrow M/N$, $S \leftarrow 0$
while $\eta = (1-\varepsilon^M)^k \leq \eta_0$ **do**
  $T \leftarrow P$
  Add Z random plane hypotheses to T
  Select plane hypotheses from T and store in P
  Count number of explained interest points (inliers) I for P
  if $I > S$ **then**
    $S \leftarrow I$
    $\varepsilon \leftarrow S/N$
  end if
  $k \leftarrow k+1$
end while

Algorithm 1 shows our proposed method for plane detection. In each iteration a small number Z of new plane hypotheses T is computed which have to compete with the selected hypotheses P of the last iteration. The termination criterion is based on the true inlier ratio ε and the number of samples M which are necessary to compute the homographies. As long as we do not know these values we use the best estimate available up to now. For ε that is the ratio of the number of explained interest points S of the current best plane hypotheses and the number of matched interest points N to explain. Accordingly M is the number of plane hypotheses currently selected multiplied with the minimal set of interest points $m = 4$ to compute one plane homography. Furthermore in Algorithm 1 k is the number of iterations, η stands for the probability that no correct set of hypotheses is found and η_0 is the desired failure rate. Due to the incremental scheme it is possible to guide the computation of new hypotheses to unexplained regions.

2.1 Incremental computation of good hypotheses

One of the key issues of approaches which use random samples is to select *good* features. Our method addresses this fact in two ways. In [21] Myatt et al. propose to bias random selection towards clusters in a multi-dimensional space. Following this approach the first interest point A is selected randomly. Then all other points are ordered in increasing Euclidean distance from A and further three nearby points are selected, depending on their position in the sorted list using a Gaussian distribution. The second assumption is that in the following iteration already selected homographies are good and thus the selection of the first interest point A is biased to unexplained interest points.

2.2 Minimal Description Length (MDL) based model selection

In each iteration selected homographies of the last iteration have to compete with newly sampled hypotheses. For the selection, the idea is that the same feature cannot belong to more than one plane and that the model cannot be fitted sequentially. Thus an over-complete set of homographies is generated and the best subset in terms of a Minimum Description Length criterion is chosen. The basic mathematical tool for this is introduced in [12] and adapted in [11]. We briefly describe the general formulation and explain our specific version for plane detection. To select the best model, the savings for each hypothesis H are expressed as

$$S_H = S_{data} - \kappa_1 S_{model} - \kappa_2 S_{error} \tag{1}$$

where in our case S_{data} is the number of interest points N explained by H and S_{model} stands for the cost of coding the model itself. We use one model (the homography of a plane) and thus $S_{model} = 1$. S_{error} describes the cost for the error added, which we express with the log-likelihood over all interest points f_i of the plane hypothesis H. Experiments have shown that the Gaussian error model

$$p(f_i|H) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right) \tag{2}$$

in conjunction with an approximation of the log-likelihood comply with our expectations. Thus the cost of the error results in

$$S_{error} = -\log \prod_{i=1}^{N} p(f_i|H) = -\sum_{i=1}^{N} \log(p(f_i|H)) \quad (3)$$

$$= -\sum_{i=1}^{N} \sum_{n=1}^{\infty} \frac{1}{n}(1-p(f_i|H))^n \approx N - \sum_{i=1}^{N} p(f_i|H) \quad (4)$$

where $\log(p(f_i|H))$ is the log-likelihood that an interest point belongs to the plane. For ε_i we use the Euclidean distance of inliers to the estimated homography. Substitution of Equation 4 into Equation 1 yields the merit of a model

$$S_H = -\kappa_1 + \sum_{k=1}^{N} ((1-\kappa_2) + \kappa_2 p(f_k|H)) \quad (5)$$

An interest point can only be assigned to one model. Hence, overlapping models compete for interest points which can be represented by interaction costs

$$s_{ij} = -\frac{1}{2} \sum_{f_k \in H_i \cap H_j} ((1-\kappa_2) + \kappa_2 \min\{p(f_k|H_i), p(f_k|H_j)\}). \quad (6)$$

In contrast $s_{ii} = S_{H,i}$ represents a merit term of a plane hypothesis. Finding the optimal possible set of homographies for the current iteration leads to a Quadratic Boolean Problem (QBP)[1]

$$\max_n \mathbf{n}^T S\mathbf{n} \; , \; S = \begin{bmatrix} s_{11} & \cdots & s_{1N} \\ \vdots & \ddots & \vdots \\ s_{N1} & \cdots & s_{NN} \end{bmatrix} \quad (7)$$

where $\mathbf{n}= [n_1, n_2, \cdots, n_N]$ stands for the indicator vector with $n_i = 1$ if a plane hypothesis is selected and $n_i = 0$ otherwise. We embed model selection in an iterative algorithm to keep the number of hypotheses tractable. Furthermore experiments have shown that a greedy approximation gives good results and thus the solution can be found very fast.

3 Splitting of clusters

Plane hypotheses often capture interest points that match the underlying homography by chance. To account for this we introduce a postprocessing step and split them into separate planes. For this, we build a neighbourhood graph of the interest points of a plane using the Delaunay triangulation. Then the mean and the standard deviation of the distance between connected interest points is computed and edges longer than s times the standard deviation are removed. For each split plane hypothesis, we verify that the support is high enough, i.e. that their merit still surpasses κ_1. We found that for our scenarios a factor of $s = 1$ works best and leave it fixed for all following experiments. Figure 1 shows the edges of the Delaunay neighbourhood graph in white and the plane hypothesis split into two groups of interest points. The dark cluster is accepted and the weaker white group is rejected, since it does not surpass the base cost κ_1.

[1] QBP assumes pairwise interaction, which in our case can be violated. But this is still a good approximation because interaction always increases cost, yielding a desirable bias against weak hypotheses.

Figure 1: A plane (top surface of packaging) accidentally picks up interest points in the background. Therefore we split the interest point clusters (coloured dots) using the mean and standard deviation of edges of a Delaunay graph (white edges).

4 Combined detection and tracking of planes

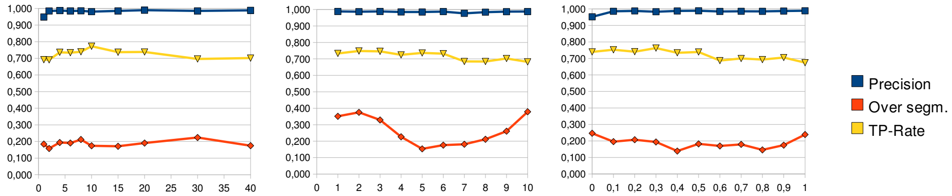
One of the key benefits of our algorithm is that prior knowledge can be introduced easily. We exploit this in image sequences where detected planes are propagated to subsequent frames. For this, the interest points of planes detected in the previous image pair are matched with interest points of the current frame, followed by a robust homography estimation using RANSAC. Thus Algorithm 1 is extended with tracked planes $P_{tracked}$, which are used to initialize P. Following $P_{tracked}$ the initialization value of the inlier ratio ε increases to the number of accumulated interest points of the tracked planes divided by the total number of matched interest points. Hence, plane detection already starts with an initial guess of planes which have to survive the following hypothesis selection stage.

5 Experiments

For all experiments we use SIFT, the well known interest point proposed by Lowe [17]. SIFT features are matched in image pairs using the Euclidean distance of the descriptors and matches are accepted if the NNDR (nearest/next ratio) d is below 0.8. To compute the homography we follow Hartley [8], that is points are normalized to zero mean and scaled to get an average distance of $\sqrt{2}$ from the origin. Then the homography is estimated using the Direct Linear Transform (DLT) algorithm.

To test our method we use two completely different sets of images. Motivated by our cognitive robotic scenarios the first set of images show packaging of arbitrary shapes typically found in a supermarket (see Figures 5). We placed each object in front of a low textured background as well as in a highly cluttered scene. For comparison, we additionally test the system with the houses data sets published by the Visual Geometry Group at the University of Oxford (see Figures 6). To get ground truth we manually marked all planes in the foreground and dominant ones of the background, resulting in 231 planes in 56 images. To test the tracking of our method the packaging data set consists of 8 sequences with 4 subsequent images, whereas we used 6 sets from Oxford also with 4 images, but these images are not ordered in a sequence.

Three numbers are computed to compare the methods, that is the feature based precision



(a) Random hypotheses Z (b) Parameter κ_1 (c) Parameter κ_2

Figure 2: Parameter optimisation

$$p_{pr} = \frac{n_{f,tp}}{n_{f,tp} + n_{f,fp}} \tag{8}$$

which is the ratio of the number of inliers $n_{f,tp}$ correctly located on a ground truth plane and the total number of features per detected plane $n_{f,tp} + n_{f,fp}$. The second number is the oversegmentation rate

$$p_{ov} = \frac{n_{p,fp}}{n_{p,tp} + n_{p,fp}} \tag{9}$$

per plane which indicates if an algorithm splits a plane in several parts. $n_{p,fp}$ the number of false positives is the number of detected planes minus the number of correct detected planes $n_{p,tp}$. Furthermore we computed the plane based true-positive rate (tp-rate)

$$p_{tp} = \frac{n_{p,tp}}{n_{p,tp} + n_{p,fn}} \tag{10}$$

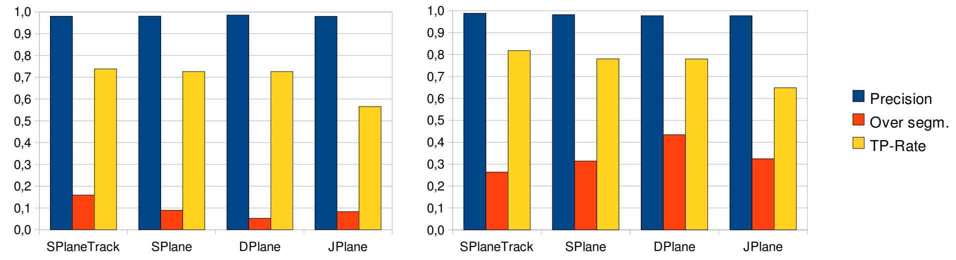
which describes the ratio of the correctly detected planes $n_{p,tp}$ and the total number of planes $n_{p,tp} + n_{p,fn}$.

5.1 Parameter optimisation

To get an idea about the behaviour of the parameters of our algorithm we tested it with the first half of the packaging data set. For this, we vary the parameters: number of random hypotheses $Z = [1 \ldots 40]$, $\kappa_1 = [1 \ldots 10]$ and $\kappa_2 = [0 \ldots 1.]$ and plot the performance measures. Figure 2 shows that our algorithm is quite robust against variation of the parameters. While the over-segmentation-rate in Figure 2(a) is almost constant the precision slightly increases at the beginning and the tp-rate has a peak at $Z = 10$. The Parameter κ_1 mostly influences the over-segmentation-rate, the tp-rate slightly decreases and thus we set $\kappa_1 = 5$ to the minimum of p_{ov}. In Figure 2(c) it can be seen, that the Parameter κ_2 is stable in a wide range. We set it to $\kappa_2 = 0.3$ where the tp-rate has a maximum and over-segmentation-rate is rather low.

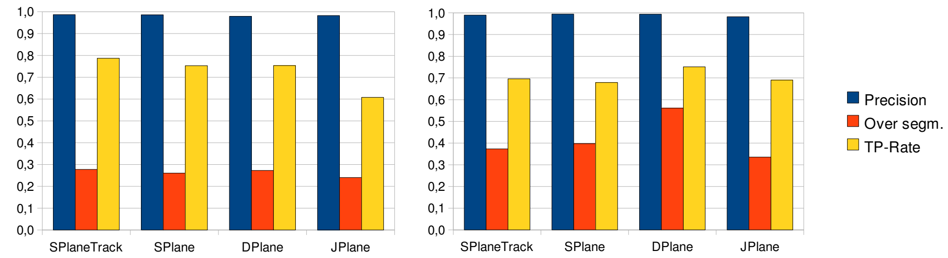
5.2 Plane detection

For this test all images of the packaging data set and the oxford houses data set are used. We compare the proposed method with/without tracking of planes *(SPlaneTrack/SPlane)* with iterative plane detection *(DPlane)* and j-linkage based plane detection *(JPlane, see [5])*. DPlane is a simple RANSAC based method, where in each iteration dominant planes are detected, supporting features marked and excluded in the following iteration. For all the following tests we use our own implementation of the algorithms.



(a) (b)

Figure 3: Comparison of plane detection methods. Left graph shows the plane detection result for images with no background texture. Tested images of the right graph have a highly textured background.



(a) (b)

Figure 4: Results for all our images (left) and for the oxford houses data set (right).

The experimental evaluation shows that our model based selection method outperforms the other methods in terms of tp-rate and lower over-segmentation especially for complex scenes. Although it is not optimized for outdoor environments of the Oxford houses it competes with JPlane. Only the incremental RANSAC approach DPlane has a higher tp-rate of the cost of over segmentation. If one compares Figure 3(a) with the Figures 3(b) and 4(b) an interesting detail can be seen. Although we use the same postprocessing step for the methods (see Section 3) for the case of highly cluttered images over-segmentation increases for DPlane and JPlane, while it remains low for SPlane. Comparing Figure 3(a) and 3(b) it seems that all methods have a higher tp-rate in case of a cluttered background. For foreground objects some of the marked ground truth planes are very small and thus easily missed, while the background planes of the cluttered scenes are generally rather large and thus more easily detected, which explains the higher overall tp-rate for this scenes.

Table 1 shows the results for the packaging data set depict in Figure 4(a). Comparing SPlane and DPlane it can be seen that although SPlane converges faster and the mean number of random sample per image is lower the tp-rate is higher. A reason therefore is that incrementally filtering out of interest points which support planes detected first in DPlane leads to a decreasing inlier/outlier ratio and thus to a increasing number of samples for planes detected later. In contrast SPlane treats all planes together and thus the number of samples has an appropriate lower value. In Figure 5 planes detected in our packaging data set are depicted in different colours. Whereas in Figure 5(a) and 5(b) flat objects located on the groundplane are “correctly” detected as one plane, they are separated in Figure 5(c). Because of the fore-

method	precision	ov. segm. rate	tp-rate	num. samples
SPlaneTrack	0.987	0.278	0.787	2455
SPlane	0.986	0.261	0.753	2407
DPlane	0.979	0.273	0.753	8278
JPlane	0.983	0.241	0.608	5000

Table 1: Comparison of different methods for plane detection (packaging data set)

(a) (b) (c)

Figure 5: Examples of our packaging data set.

ground object, which separates the magazine and the CD's, our postprocessing described in Section 3 also splits this plane in two separate groups. Figure 6(a) shows MertonCollege1 from the Oxford data set. Features assigned to detected planes are drawn in different colours, whereas false matches, inaccurate features and features which belong to not accepted weak planes are depicted in white. The white lines indicate the motion of the interest points in the image pair. Another example of the Oxford data set is shown in Figure 6(b).

6 Conclusion and Further Work

We formalize model selection with Minimal Description Length to detect multiple planes in images. Planes once detected are tracked and serve as priors in subsequent images. Instead of creating all hypotheses at once an incremental scheme is proposed where tracked planes as well as planes of former iterations serve as prior to guide randomized selection of interest

(a) MertonCollege1 (b) WadhamCollege

Figure 6: Examples of the Oxford Visual Geometry data set.

points to unexplained regions. Hypotheses which best explain the image are selected using an MDL criterion and retained to the next iteration. Although the method is incremental planes have to compete for features and thus features are treated equally. The result of our algorithm is planes specified by 2D homographies and sparse point clouds. For future work we want to extend the post-processing to get a dense piecewise planar object model. One possibility would be to introduce a multi-label segmentation using a Markov Random Field (MRF) optimization and graph-cuts, e.g., such as proposed by Sudipta et al. [26] and Micusik et al. [18].

Acknowledgement

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement No. 215181, CogX and No. 215821, GRASP. Bastian Leibe's research was funded by the DFG cluster of excellence UMIC (DFG EXC 89).

References

[1] C. Baillard and A. Zisserman. Automatic reconstruction of piecewise planar models from multiple views. In *Computer Vision and Pattern Recognition (CVPR 1999)*, pages II: 559–565, 1999.

[2] T.-J. Chin, H. Wang, and D. Suter. Robust fitting of multiple structures: The statistical learning approach. In *IEEE International Conference on Computer Vision (ICCV)*, 2009.

[3] Antonio Criminisi, Ian D. Reid, and Andrew Zisserman. Duality, rigidity and planar parallax. In *ECCV '98: Proceedings of the 5th European Conference on Computer Vision-Volume II*, pages 846–861, London, UK, 1998. Springer-Verlag.

[4] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, 1981.

[5] D. Fouhey, D. Scharstein, and A. Briggs. Multiple plane detection in image pairs using j-linkage. In *20th International Conference on Pattern Recognition (ICPR 2010)*, To appear 2010.

[6] Friedrich Fraundorfer, Konrad Schindler, and Horst Bischof. Piecewise planar scene reconstruction from sparse correspondences. *Image and Vision Computing*, 24(4):395 – 406, 2006.

[7] Luc Van Gool, Marc Proesmans, and Andrew Zisserman. Planar homologies as a basis for grouping and recognition. *Image and Vision Computing*, 16(1):21 – 26, 1998. Geometric Modeling and Invariants for Computing Vision.

[8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[9] Y. Kanazawa and H. Kawakami. Detection of planar regions with uncalibrated stereo using distributions of feature points. In *British Machine Vision Conference (BMVC)*, 2004.

[10] Robert Kaucic, Richard Hartley, and Nicolas Dano. Plane-based projective reconstruction. *Computer Vision, IEEE International Conference on*, 1:420, 2001.

[11] Bastian Leibe, Aleš Leonardis, and Bernt Schiele. Robust object detection with interleaved categorization and segmentation. *Int. J. Comput. Vision*, 77(1-3):259–289, 2008.

[12] Aleš Leonardis, Alok Gupta, and Ruzena Bajcsy. Segmentation of range images as the search for geometric parametric models. *Int. J. Comput. Vision*, 14(3):253–277, 1995.

[13] G. Lopez Nicolas, J.J. Guerrero, O.A. Pellejero, and C. Sagues. Computing homographies from three lines or points in an image pair. In *CIAP 2005*, pages 446–453, 2005.

[14] Manolis I. A. Lourakis and Stelios C. Orphanoudakis. Visual detection of obstacles assuming a locally planar ground. In *ACCV '98: Proceedings of the Third Asian Conference on Computer Vision-Volume II*, pages 527–534, London, UK, 1997. Springer-Verlag.

[15] M.I.A. Lourakis, S.V. Tzurbakis, A.A. Argyros, and S.C. Orphanoudakis. Using geometric constraints for matching disparate stereo views of 3d scenes containing planes. In *ICPR '00: Proceedings of the International Conference on Pattern Recognition*, page 1419, Washington, DC, USA, 2000. IEEE Computer Society.

[16] M.I.A. Lourakis, A.A. Argyros, and S.C. Orphanoudakis. Detecting planes in an uncalibrated image pair. In *British Machine Vision Conference (BMVC)*, 2002.

[17] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, 2004.

[18] B. Micusik and J. Kosecka. Piecewise planar city 3d modeling from street view panoramic sequences. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:2906–2912, 2009.

[19] Roger Mohr, Luce Morin, and Enrico Grosso. Relative positioning with uncalibrated cameras. In *Geometric invariance in computer vision*, pages 440–460, Cambridge, MA, USA, 1992. MIT Press.

[20] Thomas Mörwald, Johann Prankl, Andreas Richtsfeld, Michael Zillich, and Markus Vincze. Blort - the blocks world robotic vision toolbox. In *Best Practice in 3D Perception and Modeling for Mobile Manipulation at ICRA 2010*, Anchorage, Alaska, 2010.

[21] D.R. Myatt, P.H.S. Torr, S.J. Nasuto, J.M. Bishop, and R. Craddock. Napsac: High noise, high dimensional robust estimation - it's in the bag. In *British Machine Vision Conference (BMVC)*, 2002.

[22] J. Piazzi and D. Prattichizzo. Plane detection with stereo images. In *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference*, pages 922 –927, may 2006.

[23] C. A. Rothwell, A. Zisserman, D. A. Forsyth, and J. L. Mundy. Planar object recognition using projective shape representation. *Int. J. Comput. Vision*, 16(1):57–99, 1995.

[24] Konrad Schindler, David Suter, and Hanzi Wang. A model-selection framework for multibody structure-and-motion of image sequences. *Int. J. Comput. Vision*, 79(2): 159–177, 2008.

[25] Peter F. Sturm and Stephen J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 1:1432, 1999.

[26] N. Sinha Sudipta, Drew Steedly, and Richard Szeliski. Piecewise planar stereo for image-based rendering. In *International Conference on Computer Vision (ICCV 2009)*, September 2009.

[27] Roberto Toldo and Andrea Fusiello. Robust multiple structures estimation with j-linkage. In *ECCV '08: Proceedings of the 10th European Conference on Computer Vision*, pages 537–547, Berlin, Heidelberg, 2008. Springer-Verlag.

[28] Bill Triggs. Autocalibration from planar scenes. In *ECCV '98: Proceedings of the 5th European Conference on Computer Vision-Volume I*, pages 89–105, London, UK, 1998. Springer-Verlag.

[29] E. Vincent and R. Laganiere. Detecting planar homographies in an image pair. In *Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on*, pages 182 –187, 2001.