Incremental Model Selection for Detection and Tracking of Planar Surfaces

Johann Prankl¹
prankl@acin.tuwien.ac.at
Michael Zillich¹
zillich@acin.tuwien.ac.at
Bastian Leibe²
leibe@umic.rwth-aachen.de
Markus Vincze¹
vincze@acin.tuwien.ac.at

 Automation and Control Institute Vienna University of Technology Austria
 UMIC Research Centre RWTH Aachen Germany

Man-made environments are abundant with planar surfaces which have attractive properties and are a prerequisite for a variety of vision tasks. Examples include application for camera self-calibration, scene analysis and 3D reconstruction. Planes are also used in robotics applications for obstacle detection, camera localization and for object recognition.

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. Schindler et al. [3] 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. 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 \le \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 S to explain. Accordingly S is the number of plane hypotheses currently selected multiplied with the minimal set of interest points S is the number of iterations, S stands for the probability that no correct set of hypotheses is found and S is the desired failure rate. Due to the incremental scheme it is possible to guide the computation of new hypotheses to unexplained regions.

One of the key issues of approaches which use random samples is to select *good* features. Our method addresses this fact in two ways. The first idea is to select the first interest point *A* randomly and sort the other points in increasing Euclidean distance from *A*. Then 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





Figure 1: Examples showing detected planes in our packaging data set and in the Oxford Visual Geometry data set.

iteration already selected homographies are good and thus the selection of the first interest point *A* is biased to unexplained interest points.

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 [2] and adapted in [1]. 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 in conjunction with an approximation of the log-likelihood comply with our expectations. Thus the cost of the error results in

$$S_{error} = -\sum_{i=1}^{N} \log(p(f_i|H)) \approx N - \sum_{i=1}^{N} p(f_i|H)$$
 (2)

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. An interest point can only be assigned to one model. Hence, overlapping models compete for interest points and Equation 1 leads to interaction costs and merit terms of a plane hypotheses. Finding the optimal possible set of homographies for the current iteration leads to a Quadratic Boolean Problem (QBP). 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.

- 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.
- [2] 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.
- [3] 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.