Adapting Polynomial Mahalanobis Distance for Self-supervised Learning in an Outdoor Environment

Miloslav Richter, Petr Petyovsky and Ondrej Miksik

Laboratory of Telepresence and Robotics & Computer Vision Group, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Introduction

- Autonomous robot navigation is one of the most extensively studied problem in the field of mobile robotics
  - The goal is to detect the drivable surface ahead of the robot
  - Common sensors provide information about obstacles in a near field, however long-range sensing is needed for high speed vehicles
  - A combination of short-range sensors with a camera is commonly used to overcome such limitations
- Generally, state-of-the-art approaches use color based segmentation of road/non-road regions in particular
  - The question is, how the belongness to a model is measured
  - Many algorithms employ Mahalanobis distance, however it is assumed, that the data points have a normal distribution.
  - Recently proposed Polynomial Mahalanobis Distance (Grudic & Mulligen, RSS 2006) represents more discriminative metric
- Polynomial Mahalanobis Distance (PMD)
  - Outperforms others commonly used metrics in high dimensional and non-linear feature spaces
  - The main drawback: its precision is strongly dependent on training data
  - Only carefully selected image patches, which perfectly represent road surface, are used as training areas
  - Such assumption can not be fulfilled in a real application, because we are able to estimate a ground plain and determine a training area, however we are not able to determine only a very small image patch, which perfectly represents surface of a whole road
- Proposed Three Stage Algorithm (TSA)
  - Deals with picking of suitable data points from the training area
  - Provides self-supervised learning algorithm for long-term road representation

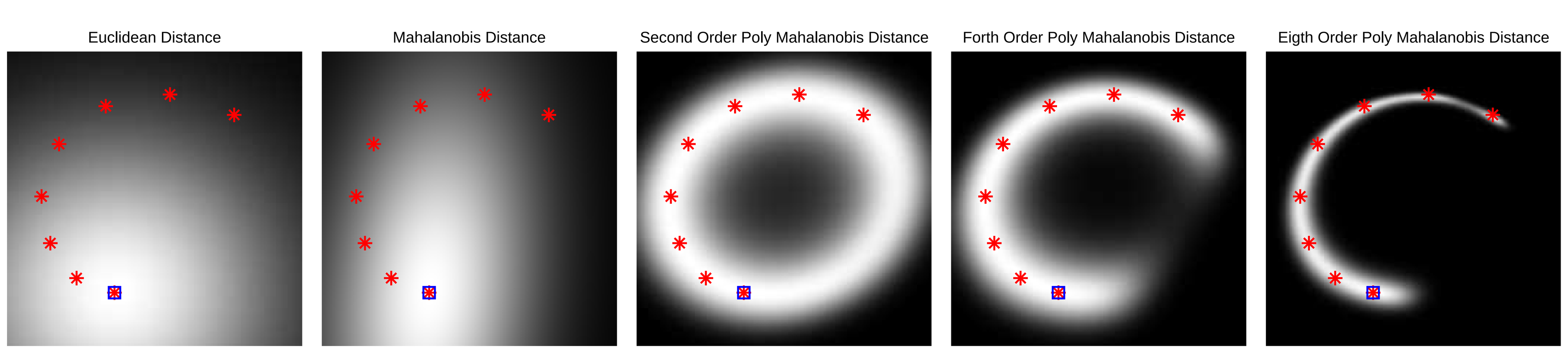
Figure 1: Output of our system

Polynomial Mahalanobis Distance

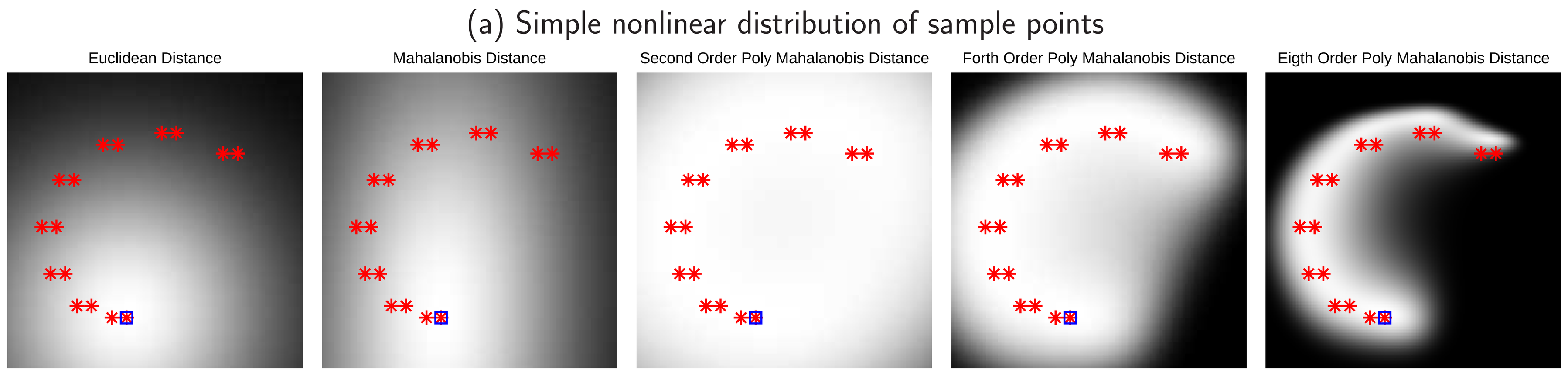
- Grudic and Mulligen proposed framework for efficient computation of a PMD by iterative forward mapping from lower to higher order polynomial terms:

$$d_{PM}(\mathbf{z}_i, \mathbf{z}_j) = d_{M_{\sigma^2}}(\mathbf{z}_i, \mathbf{z}_j) + \sum_{l=1}^{L} d_{M_{\sigma^2}}(\mathbf{g}_l^i, \mathbf{g}_l^j), \qquad (1)$$

  where $L > 1$ defines $q = 2^L$-order PMD and $\mathbf{g}_l^i, \mathbf{g}_l^j$ are projected polynomial terms into level l
- Neither Euclidean, nor standard Mahalanobis distance respect highly non-linear character of sample points. The superior results are obtained with PMD, because it maps its feature points into higher order polynomial terms
- The problem is, that PMD expects, that the sample points could be approximated by a higher order polynomial term – even with only a small number of similar points, PMD degenerates and does not follow the sample points precisely



(a) Simple nonlinear distribution of sample points



(b) Duplicated nonlinear distribution of sample points

Figure 2: Synthetic sample points distribution (the brighter the background is, the closer to the reference point denoted by the blue rectangle). Distance metrics from left: Euclidean, Mahalanobis, Second, Fourth and Eigth Order PMD

Figure 3: Various sizes of a training area - stability strongly decreases with larger training area (images from left, to right: purple, blue, green, red)

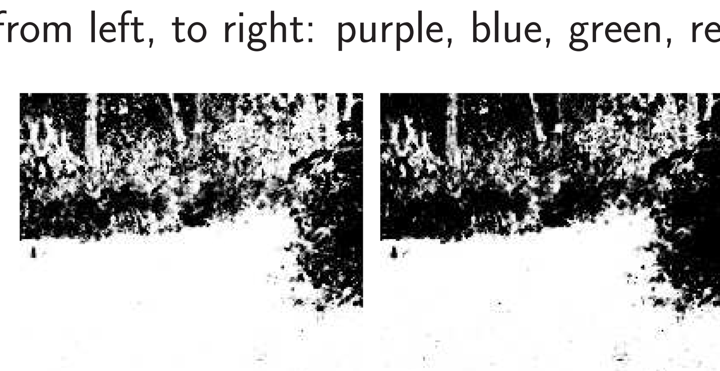


Figure 4: Various orders of PMD with $L = \{4, 8\}$, $\sigma^2 = 10^{-5}$

Figure 5: Influence of various σ^2: $L = 4$, $\sigma^2 = \{10^{-4}, 10^{-6}, 10^{-8}\}$

Three Stage Algorithm

- Continues re-learning of models stored in the pool
- Each model is represented by a codebook, which consists from parts:
  1. The first one represents "strong model" (memory) – cf. Fig. 6, blue part
  2. "Temporary model", cache, is used for weak hypothesis (orange part)
  - Both parts consist of codewords represented by their mean value, min/max brightness, mass and time of last update

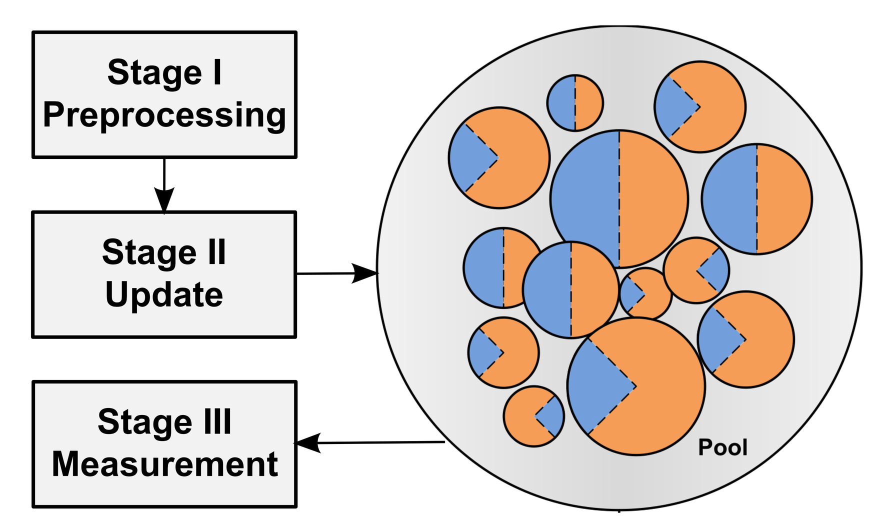
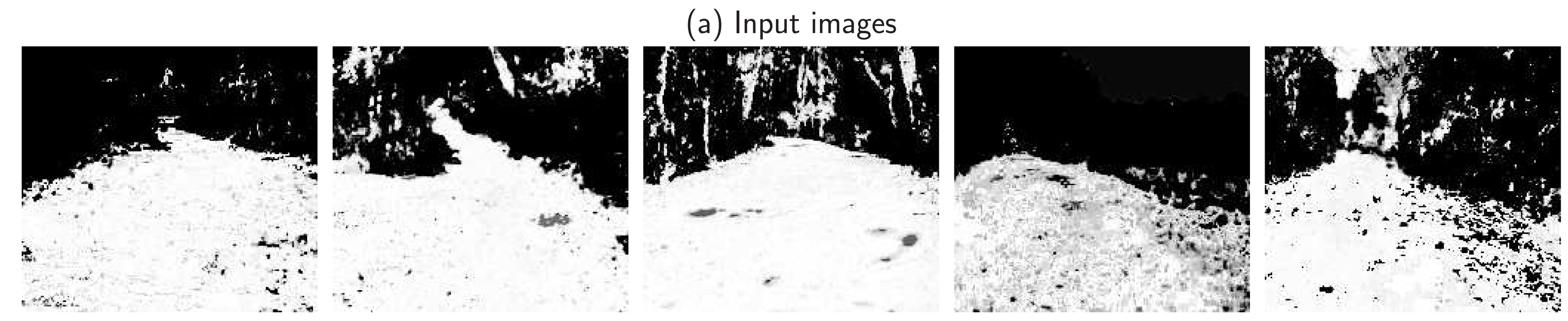


Figure 6: A Three Stage Algorithm overview.

1. Preprocessing (k-means clustering)
   - To split the training samples into different parts to create multiple models in the pool
   - Reduces computational complexity and deals with outliers
2. Update
   - Iteratively updates all road models (codebooks) in the pool, which overlap clusters from the preprocessing stage, otherwise the new codebook is created
3. Measurement
   - Scores the pixels outside of the training area if they belongs to the road/non-road region

Results

(a) Input images



(b) Polynomial Mahalanobis Distance with the three stage algorithm

(c) Output of proposed algorithm

Figure 7: Proposed algorithm – various road surfaces and light conditions.

Conclusions

- TSA successfully deals with picking of suitable training samples and provides self-supervised learning algorithm for long-term road representation
- TSA is not sensitive to a regularization constant σ^2 and a level L of the PMD
- "Strong" and "Weak" parts of models represent "road history" and sudden changes of road appearance
- Proposed algorithm is robust to outliers

This research has been supported by the Czech Ministry of Education in the frame of MSM 0021630529 Research Intention Inteligent Systems in Automation.
ondra.miksik@gmail.com http://www.miksik.co.uk