Learning to Associate with CRF-Matching

Fabio Ramos Dieter Fox

Data association remains a difficult and fundamental problem for many robotics tasks. It is a crucial component in problems such as tracking, image registration, reconstruction, and simultaneous localisation and mapping (SLAM).

Most existing data association algorithms consider only a limited set of features and require substantial manual tuning to work in practice. For instance, in robotics SLAM, data associations are typically determined solely based on the locations of landmarks, thereby ignoring important appearance information. However, incorporating more complex information makes manual tuning extremely cumbersome. Another limitation of existing data association algorithms is the fact that they only provide a deterministic result on the association. This makes them less robust and difficult to incorporate into probabilistic filtering approaches since no uncertainty on the association is returned.

This poster introduces CRF-Matching; a general multi-sensor data association framework that can be learnt from data [5]. CRF-Matching is a supervised probabilistic model able to jointly reason about the association of points. This is obtained by overcoming the independence assumption through the use of Conditional Random Fields (CRFs) [2]. CRFs are an extremely flexible technique for integrating different features in the same probabilistic framework. Features can be defined over different sensor modalities or designed to capture neighbourhood information. The power of CRFs is enhanced through the possibility to use statistical measurements (such as the likelihood of the data given the model) to learn a parametrisation of the model given some training data. This process estimates weights for the features, thus quantifying the importance of each feature for the particular task. Inference can be performed efficiently using Loopy Belief Propagation once the model has been fully specified.

We demonstrate the capabilities of CRF-Matching in two data association tasks: laser scan matching and image feature matching. In the first case data association between two laser scans is accomplished by converting the individual measurements of one laser scan into hidden nodes of a CRF. The states of each node range over all measurements in the other scan. The CRF models arbitrary information about local appearance and shape of the scans. Consistency of the association is achieved by connections between nodes in the CRF. CRF-Matching learns model parameters discriminatively from sets of aligned laser scans. When applied to a new pair of scans, maximum *a posteriori* estimation is used to determine the data associations, which in turn specify the spatial transformation between the scans. Extensive experiments show that CRF-Matching significantly outperforms ICP when matching laser range-scans with large spatial offset. Furthermore, they show that our approach is able to reliably match scans without a priori information about their spatial transformation, and to incorporate visual information so as to further improve matching performance.

An extension of the previous approach can be employed for association of image features. This is obtained through the use of the Delaunay triangulation [4] as the graph structure for CRF-matching. The graph defines neighbour points by respecting interesting geometric constraints such as the *empty circle property*. This creates a graph that is not over-connected while still encoding most of the geometric relationship between neighbour points. We demonstrate how pairwise potential functions can be defined over edges to jointly reason about the associations. In addition to pairwise potential functions, local potential functions can also be defined to directly incorporate sensor observations into the model. In our implementation, SIFT features and descriptors [3] were used, although any image feature descriptor or detector could also be employed. As opposed to the SIFT match procedure described in [3] where Euclidean distance is used to measure the compatibility of matches, we show how a boosting classifier can be learnt and integrated in CRF-Matching to capture non-linear relationships between image descriptors to best match the features. We perform extensive experiments in challenging indoor and outdoor datasets where images were obtained while the robot was in motion. Some of complexities in the datasets include occlusion, different illumination conditions, blurring, translation and rotation transformations. We show that CRF-Matching outperforms the commonly applied RANSAC procedure [1] when there is a small set of detected features.

References

[1] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Comms. Assoc. Comp. Mach.*, 24(6):381–395, 1981.

[2] J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proc. of the International Conference on Machine Learning (ICML)*, 2001.

[3] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[4] F. R. Preparata and M. I. Shamos. *Computational Geometry: An Introduction.* Springer-Verlag, New York, USA, 1985.

[5] F. Ramos, D. Fox, and H. Durrant-Whyte. Crf-matching: Conditional random fields for feature-based scan matching. In *Proc. of Robotics: Science and Systems*, 2007.