Visual Odometry: Literature Survey

Alex Kreimer

October 27, 2015

Abstract

0.1 Refining essential matrix estimates from RANSAC

The paper deals with the problem of estimating relative pose of two cameras from outlier contaminated feature correspondences. It is a common practice to use RANSAC with conjunction with a linear method to estimate a model (in this case the Essential matrix) and then to refine it using non-linear optimization method. [2] evaluate several refinement methods which minimize functions of Sampson's error. All perform well on range sets of correspondences or sets with low outlier rates; but many perform poorly otherwise. The most accurate solution is give by minimizing a robust function (Blake-Zisserman) of a Sampson's error. The rotations are parameterized as quaternions and the optimization is performed over the essential manifold, see [6]. The authors use IRLS combined with LM [5] method to optimize the objective.

0.2 Vision-based robot localization without explicit object models

[3] propose a solution to localize a robot in an unknown 2d environment using visual input. The authors train neural net as a regression model N(I)=(x,y). As an input for the net they use image statistics (first and second moments of the edge distributions, mean edge orientations, densities of lines).

0.3 Visual Homing: Surfing on the Epipoles

[1] propose a solution for visual homing. Using this method a robot can be sent to desired positions and orientations in 3D space specified by single images taken from these positions. The method is based on estimating epipolar geometry between a pair of views. A 3D model of the environment is not required. Using the epipolar geometry most of the parameters which specify the differences in position and orientation of the camera between the two images are recovered. From a pair of images translation may be recovered only up to a signed scale. In order to find out the real distance the robot make an extra step and takes an

additional image. The authors prove that when the camera moves on a straight line the features move along the epipolar lines while their coordinates and the coordinate of the epipole obey the cross-ratio relation.

0.4 A way to parameterize rotations

[6] proposes a method to use quaternions in an unconstrained nonlinear optimization. Quaternions representing rotation have four elements but only three degrees of freedom, since they have to be of norm one. This constraint has to be taken into account when applying e.g. Levenberg-Marquardt algorithm. One of the ways to address this issue is to use appropriate parameterization (others are a projection step and Lagrange multipliers). Well known parameterizations are Euler angles and axis-angle representation.

[4] call a parameterization fair if it does not introduce more numerical sensitivity than inherent to the problem itself. This is guaranteed, if any rigid transformation of the space to be parameterized results in an orthogonal transformation of the parameters. Both axis-angle and quaternion parameterizations are fair, while Euler angles is not.

Authors search for a parameteriation that:

1. is minimal, i.e. uses only three parameters
2. the three parameters may be changed arbitrarily by the optimization algorithm
3. the resulting quaternion has always norm 1.

This new approach is based on the observation that all quaternions of norm-1 lie on the unit sphere in $\mathbb{R}^4$. The authors use the shortest connection between two points on a sphere, i.e. a great circle. For describing a movement on the sphere starting at $\mathbf{h_0}$ they use a vector v_4 lying in the tangential hyperplane that touches the sphere at $\mathbf{h_0}$. This hyper-plane is a subspace of $\mathbb{R}^4$, thus vectors in this plane may be represented as 3-vectors with respect to a plane-local coordinate frame.

Experiments are made on a synthetic (small) data-set. The authors perform bundle adjustment and compare their approach with axis-angle representation. The conclusion is that this representation performs better for rotations, for transnational motion both method are approximately equal.

References

[1] Ronen Basri, Ehud Rivlin, and Ilan Shimshoni. Visual homing: Surfing on the epipoles. *International Journal of Computer Vision*, 33(2):117–137, 1999.

[2] Tom Botterill, Steven Mills, and Richard Green. Refining essential matrix estimates from ransac.

[3] Gregory Dudek and Chi Zhang. Vision-based robot localization without explicit object models. In *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, volume 1, pages 76–82. IEEE, 1996.

[4] Joachim Hornegger and Carlo Tomasi. Representation issues in the ml estimation of camera motion. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 640–647. IEEE, 1999.

[5] Donald W Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics*, 11(2):431–441, 1963.

[6] Jochen Schmidt and Heinrich Niemann. Using quaternions for parametrizing 3-d rotations in unconstrained nonlinear optimization. In *VMV*, volume 1, pages 399–406, 2001.