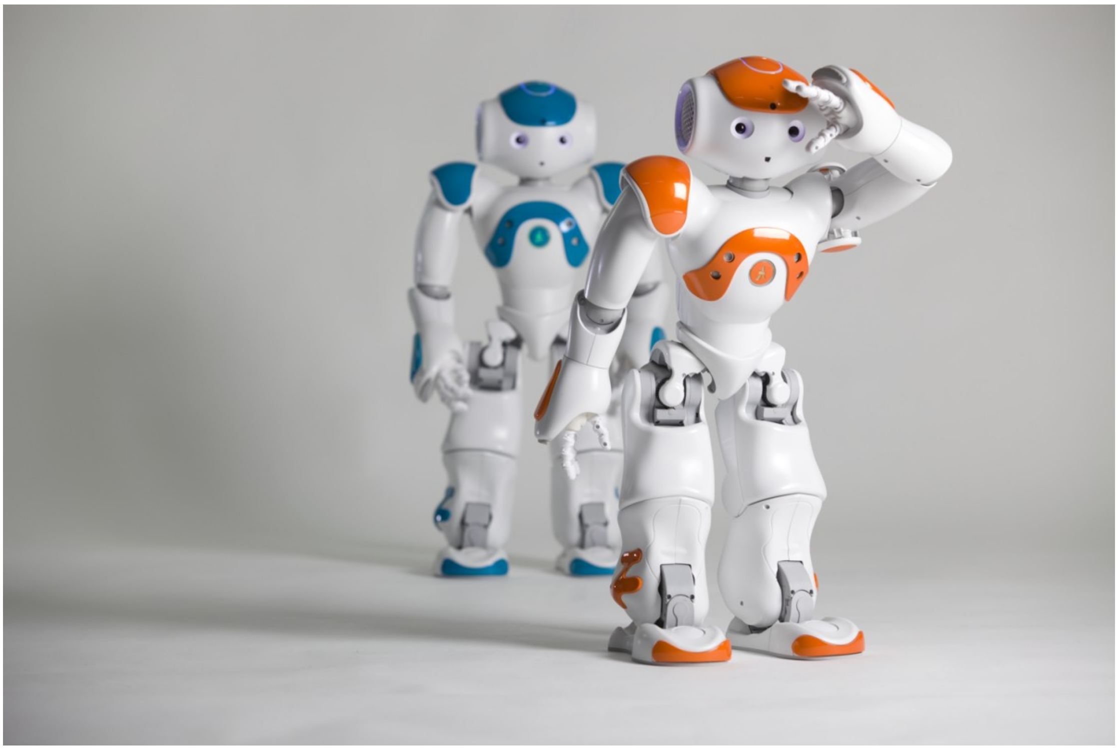
Collaborative Visual SLAM
Multi-Agent Visual Odometry and SLAM with humanoid robots.



An Artificial Intelligence project by Auke J. Wiggers, Camiel R. Verschoor, Chiel Kooijman and Steven Laan

Collaborative Visual SLAM

Multi-Agent Visual Odometry and SLAM with humanoid robots.
Project AI (6 EC)
Artificial Intelligence
Faculty of Science
University of Amsterdam

Chiel Kooijman
5743028
Chiel999@gmail.com

Steven Laan
6036031
S.Laan@uva.nl

Camiel Verschoor
10017321
Verschoor@uva.nl

Auke Wiggers
6036163
A.J.Wiggers@uva.nl

January 28, 2013

Contents

1 Introduction

2 Related Work

3 Theory

4 Pipeline

In this section, the pipeline of our proposed system is described stepwise.

$$\begin{tikzcd} Calibration \arrow[d] \\ Egomotion \arrow[d] \\ Reconstruction \arrow[d] \\ Modelling \arrow[d] \\ Combining Visual Odometry \end{tikzcd}$$

Figure 1: Schematic overview of the pipeline of the system

4.1 Calibration

Camera calibration is the process of estimating the intrinsic and extrinsic parameters of the camera. Images taken by the Nao robot's camera are subject to barrel distortion. Calibration provides us with camera parameters that allow us to transform images in a way that removes the distortion.

To find the camera parameters we use the method proposed by Zhang (1999) and implemented in the OpenCV library (Bradski, 2000). In this method, a chessboard-like panel of which the dimensions are known, is held up in front of the camera in different positions. The algorithm uses the knowledge that the panel is planar to infer the camera parameters f_x, f_y, c_x, c_y and distortion coefficients ?, forming the matrix

$$F = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

4.2 Egomotion

Egomotion is the process of estimating a camera's motion relative to a rigid scene. There are two main approaches for estimating the relative motion between two frames, namely, Image-to-Image and Image-to-World.

4.2.1 Feature extraction

In this section, we concisely describe the feature extraction methods that were used in the system to extract features. In total three feature extractions methods were applied, namely, Binary Robust Invariant Scalable Keypoints (BRISK) (Leutenegger et al., 2011), Oriented FAST and Rotated BRIEF (ORB) (Rublee et al., 2011) and Fast Retina Keypoints (FREAK) (Ortiz, 2012).

Robust Invariant Scalable Keypoints

BRISK relies on an easily configurable circular sampling pattern from which it computes brightness comparisons to form a binary descriptor string. The unique properties, rotation and scale invariance, of BRISK can be useful for a wide spectrum of applications, in particular for tasks with hard real-time constraints or limited computation power: BRISK ïňĄnally offers the quality of high-end features in such time-demanding applications.

Oriented FAST and Rotated BRIEF

[Insert]

Fast Retina Keypoint

[Insert]

4.2.2 Feature Matching

FLANN FEATUREMATCHER (Muja and Lowe, 2009)

4.2.3 Image-to-Image

- Match 2D with 2D descriptors. - Normalize matches. - Compute transformation matrices. - Scale points. - Find fundamental matrix and use RANSAC to reject outliers. - Compute essential matrix. - Estimation of projection matrix. - Decide on all candidates.

4.2.4 Image-to-World (PnP)

- Match 2D with 3D descriptors. - PNP Ransac. - Obtain rotation matrix from rotation vector. - Triangulate any unknown points.

4.3 Reconstruction

The previously described methods estimate the motion of the camera and this pose estimation process yields the position of the different cameras in space. The matching obtained matching feature points across frames and then esitmate the 3-Dimensional (3D) position of those points in space. To obtain these 3D points there are at least two camera poses needed to estimate the spatial position.

4.4 Modelling

4.5 Combining Visual Odometry

5 Experimental Setup

6 Results

7 Discussion

8 Conclusion

References

G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

Stefan Leutenegger, Margarita Chli, and Roland Y. Siegwart. Brisk: Binary robust invariant scalable keypoints. In *Proceedings of the 2011 International Conference on Computer Vision*, ICCV '11, pages 2548–2555, Washington, DC, USA, 2011. IEEE Computer Society. ISBN 978-1-4577-1101-5. doi: 10.1109/ICCV.2011.6126542. URL `http://dx.doi.org/10.1109/ICCV.2011.6126542`.

Marius Muja and David G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *International Conference on Computer Vision Theory and Application VISSAPP'09)*, pages 331–340. INSTICC Press, 2009.

Raphael Ortiz. Freak: Fast retina keypoint. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, CVPR '12, pages 510–517, Washington, DC, USA, 2012. IEEE Computer Society. ISBN 978-1-4673-1226-4.

Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary R. Bradski. Orb: An efficient alternative to sift or surf. In Dimitris N. Metaxas, Long Quan, Alberto Sanfeliu, and Luc J. Van Gool, editors, *ICCV*, pages 2564–2571. IEEE, 2011. ISBN 978-1-4577-1101-5.

Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 666–673. Ieee, 1999.