DCCAL - Discrete Cameras Calibration using Properties of Natural Scenes

Milestone 1
State of the Art

Funded by:

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

FCT PTDC/EEA-CRO/105413/2008
January 2010 - December 2012

July 2010

DCCAL - Discrete Cameras Calibration using Properties of Natural Scenes

Milestone 1 - State of the Art

Webpage

Electronic address of the webpage of the DCCAL project:
http://users.isr.ist.utl.pt/~jag/project_dccal

State of the Art

Natural image statistics [Hyvarinen08] is a recent theoretical framework based on a principle that the properties of the biological visual systems emerged from evolutionary adaptation processes and tend to be reflections of the statistical structure of natural images. Hence, ambitioning to replicate the amazing plasticity capabilities of the biological visual systems implies that one has to study directly the statistics of natural images. Despite recent, this kind of research already proved fruitful by evidencing fundamental properties, in terms of local, global and spectral statistics, of real-world images, that has been also exploited for computer vision tasks, such as classification [Torralba03], image restoration [Freeman00] and 3D inference [Potetz06].

Considering an even more general setup, Pierce and Kuipers [Pierce97], measure the dissimilarity, or distance, between sensor elements that are not necessarily light sensors. The elements are then embedded in a metric space using metric scaling [Krzanowski88], which also determines the dimension of the space. A relaxation method then improves this embedding, so that the Euclidean distance between sensor elements better matches the dissimilarity between the sensor inputs. In an experiment, the authors use this method to reconstitute the geometry of an array of visual sensors that scans a fronto-parallel image (under the assumption that the array has a rectangular shape). Our approach is closely related to this work.

Going further, Olsson et al. [Olsson04] use the information metric of [Crutchfield90] as a more appropriate method to measure the distance between visual or other sensor elements. They also show how visual sensors - the pixels of the camera of a mobile robot- can be mapped to a plane, either using the method of [Pierce97], or their own, that embeds sensor elements in a square grid.

The works of Olsson et al. and of Pierce and Kuipers are very interesting to computer vision researchers but, since the geometry of the embedding space is either abstract or fixed to a grid, in either case, it lacks an explicit connection to the geometry of the sensor. Thus, although these results are of great interest, they are not directly applicable in our case, mainly because we lack images (have just pixel streams). Moreover, these statistics are about images formed on a planar image plane, which is a hindrance in our case: first, we do not want to exclude the case of visual sensor elements that are separated by more than 180 degrees, such as the increasingly popular omnidirectional cameras. Second, in the statistical analysis of Cartesian images, some assumptions are implicitly made: a viewing direction is privileged and an image scale (equivalent to a focal length) is implicitly chosen. Third, the local statistical properties of perspective images depend of the orientation of the image plane with respect to the scene, except in special constrained cases such as the fronto-parallel “leaf world” of Wu et al. [Wu04]. In other words, planar image statistics lack generality.

Defining images on the unit sphere thus appears as a natural way to render image statistics independent of the sensor orientation, at least with proper assumptions on the surrounding world and/or the motion of the sensor. Following these representation and assumptions, our previous work [selfRef07, selfRef08, selfRef10] showed that a definite relation can be found between streams of photocell data and the arbitrary geometry of the sensor. The number of cells is however still small, about one hundred.

Recently, Censi and Scaramuzza [Censi13] showed that the reconstruction of a sensor can include also a global scale factor. The authors formalize calibration as a generalization of multidimensional scaling (MDS) and iterate on the distances matrix to minimize a ratio of singular values.

In [selfRef13] our goal is to do auto-calibration of central sensors with a number of pixels orders of magnitude larger than [Olsson06, Grossmann10] and 50% larger than [Censi13]. We also approach the computational complexity with MDS like algorithms. A relatively old but very effective in the presence of noise free data is the Classical MDS [Cox01, Datorro10], based on Euclidean distances. Its goal is to find a representation of a data set on a given dimensionality from the knowledge of all interpoint distances. Several new algorithms evolved from MDS, such as ISOMAP [Tenenbaum00], where geodesic distances induced by a neighbourhood graph are used instead of Euclidean distances.

Regarding the computational complexity, ISOMAP, and even Landmark ISOMAP [Silva08], are still considered expensive. Other large scale approximate solutions are currently known, e.g. the Nystrom method and column sampling [Talwalkar08]. These methodologies are definitely worth exploring.

References

[Agapito98] L. Agapito, E. Hayman, I. Reid. Self-calibration of a rotating camera with varying intrinsic parameters. In: Proc 9th British Machine Vision Conf, 1998.

[Agapito01] L. Agapito, E. Hayman, I. Reid. Self-calibration of rotating and zooming cameras. Int. J. Comput. Vision 52 (2001) 107–127

[Bouguet08] Jean-Yves Bouguet, Camera Calibration Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc/, Last updated June 2nd, 2008

[Censi13] A. Censi, D. Scaramuzza. Calibration by correlation using metric embedding from non-metric similarities. In IEEE T-PAMI (to appear 2013)

[Cox01] T. Cox, M. Cox. Multidimensional scaling. Chapman & Hall/CRC, 2001

[Crutchfield90] J. P. Crutchfield. Information and its metric. In L. Lam and H. C. Morris, editors, Nonlinear Structures in Physical Systems–Pattern Formation, Chaos and Waves, pages 119–130. Springer-Verlag, 1990.

[Dattorro10] J. Dattorro. Convex Optimization and Euclidean Distance Geometry. Meboo Publishing USA (2010) [https://ccrma.stanford.edu/~dattorro/mybook.html]

[DCCAL10] DCCAL - Discrete Cameras Calibration using Properties of Natural Scenes, 2010-2013, FCT project PTDC/EEA-CRO/105413/2008, http://www.isr.ist.utl.pt/~jag/project_dccal/

[Fitzgibbon05] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. IEEE CVPR 2001

[Freeman00] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. International Journal of Computer Vision, 40(1):25–47, 2000.

[Grossmann10] E. Grossmann, J. António Gaspar and F. Orabona. Discrete camera calibration from pixel streams. In Computer Vision and Image Understanding (Special issue on Omnidirectional Vision, Camera Networks and Non-conventional Cameras), Volume 114, Issue 2, Pages 198-209, February 2010.

[Hartley00] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2000 (or 2nd ed 2004).

[Hyvarinen08] A. Hyvarinen, J. Hurri, and P. O. Hoyer. Natural Image Statistics — A probabilistic approach to early computational vision. to be published by Springer-Verlag, 2008. Preprint available at http://www.naturalimagestatistics.net/.

[Karlsson05] Karlsson, N.; Bernardo, E.D.; Ostrowski, J.; Goncalves, L.; Pirjanian, P. & Munich, M. The vSLAM Algorithm for Robust Localization and Mapping Proc. IEEE Int. Conf. on Robotics and Automation, 2005, 24 - 29

[Kohler62] I. Kohler. Experiments with goggles. Scientific American, 206:62–72, 1962.

[Krzanowski88] W. J. Krzanowski. Principles of Multivariate Analysis: A User's Perspective. Clarendon Press, Statistical Science Series, 1988.

[Lowe04] "Distinctive image features from scale-invariant keypoints", D. Lowe, Int. Journal of Computer Vision, 2004, 60, 91-110

[Olsson04] L. Olsson, C. L. Nehaniv, and D. Polani. Sensory channel grouping and structure from uninterpreted sensor data. In NASA/NoD Conference on Evolvable Hardware, 2004.

[Olsson06] L. Olsson, C.L. Nehaniv, D. Polani. Measuring informational distances between sensors and sensor integration. In: Artificial Life X, MIT Press, (2006)

[Paninski03] L. Paninski. Estimation of entropy and mutual information. Neural Computation, 15:1191– 1254, 2003.

[Pierce97] D. Pierce and B. Kuipers. Map learning with uninterpreted sensors and effectors. Artificial Intelligence Journal, 92(169–229), 1997.

[Potetz06] B. Potetz and T. S. Lee. Scaling laws in natural scenes and the inference of 3d shape. In NIPS – Advances in Neural Information Processing Systems, pages 1089–1096.MIT Press, 2006.

[Ramalingam05] S. Ramalingam, P. Sturm, and S. Lodha. Towards complete generic camera calibration. In Proc. CVPR, volume 1, pages 1093–1098, 2005.

[Sammon69] J. W. Jr. Sammon. A nonlinear mapping for data structure analysis. IEEE Transactions on Computers, C-18:401–409, 1969.

[selfRef05] D. Nistér, H. Stewenius, and E. Grossmann. Non-parametric self-calibration. In proc. ICCV, 2005.

[selfRef06] E. Grossmann, E-J Lee, P. Hislop, D. Nistér, and H. Stewénius. Are two rotational flows sufficient to calibrate a smooth non-parametric sensor? In proc. IEEE CVPR, 2006.

[selfRef07] E. Grossmann, F. Orabona, and J. A. Gaspar. Discrete camera calibration from the information distance between pixel streams. In Proc. Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras, OMNIVIS, 2007.

[selfRef08] E. Grossmann, J. A. Gaspar, and F. Orabona. Calibration from statistical properties of the visual world. In European Conf. on Computer Vision, 2008, 2008.

[selfRef10] E. Grossmann, J. António Gaspar and F. Orabona. Discrete camera calibration from pixel streams. In Computer Vision and Image Understanding (Special issue on Omnidirectional Vision, Camera Networks and Non-conventional Cameras), Volume 114, Issue 2, Pages 198-209, February 2010.

[selfRef12] R. Galego, A. Bernardino, J. Gaspar. Auto-calibration of Pan-Tilt Cameras including Radial Distortion and Zoom. In International Symposium on Visual Computing (ISVC), July 16-18, Crete, Greece, 2012.

[selfRef12-Moutinho] Online Calibration of a Humanoid Robot Head from Relative Encoders, IMU Readings and Visual Data, Nuno Moutinho, Martim Brandão, Ricardo Ferreira, José Gaspar, Alexandre Bernardino, Atsuo Takanishi, José Santos-Victor, IROS 2012.

[selfRef12-Ruesch] J. Ruesch, R. Ferreira, A. Bernardino. Predicting visual stimuli from self-induced actions: an adaptive model of a corollary discharge circuit. IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, accepted 2012.

[selfRef13] R. Galego, R. Ferreira, A. Bernardino, E. Grossmann and J. Gaspar. Topological Auto-Calibration of Central Imaging Sensors. IbPRIA 2013.

[selfRef13-Ruesch] J. Ruesch, R. Ferreira and A. Bernardino. A computational approach on the co-development of artificial visual sensorimotor. In SAGE journal on Adaptive Behavior, accepted 2013

[Silva08] V. de Silva, J.B. Tenenbaum. "Sparse multidimensional scaling using landmark points". Technical Report, Stanford University.

[Snyder52] F. W. Snyder and N. H. Pronko. Vision with spatial inversion. University of Wichita Press, 1952.

[Stratton96] G. M. Stratton. Some preliminary experiments on vision without inversion of the retinal image. Psychological Review, 3(6):611–617, Nov 1896.

[Sturm10] Peter Sturm, Srikumar Ramalingam, Jean-Philippe Tardif, Simone Gasparini and Joao Barreto. Camera Models and Fundamental Concepts Used in Geometric Computer Vision. Foundations and Trends in Computer Graphics and Vision, Vol. 6, Nos. 1–2 (2010) 1–183, Now Publishers

[Talwalkar08] A. Talwalkar, S. Kumar, H. Rowley. "Large-Scale Manifold Learning". IEEE CVPR 2008

[Tenenbaum00] J. Tenenbaum, V. de Silva, J. Langford. A global geometric framework for nonlinear dimensionality reduction. Science 290 (2000)

[Torralba03] A. Torralba and A. Oliva. Statistics of natural image categories. Network: Computation in Neural Systems, 14:391–412, 2003.

[Torralba06] R. Fergus, A. Torralba, and W. T. Freeman. Random lens imaging. Technical Report MIT CSAIL TR 2006-058, Massachusetts Institute of Technology, 2006.

[Tsai86] R. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. In IEEE Conf. on Computer Vision and Pattern Recognition, 1986.

[Wexler03] Learning epipolar geometry from image sequences, Yonatan Wexler, Andrew W. Fitzgibbon and Andrew Zisserman, CVPR'03, pp:II- 209-16 vol.2

[Wu04] Y. N. Wu, S.-C. Zhu, and C.-E. Guo. From information scaling of natural images to regimes of statistical models. Technical Report 2004010111, Department of Statistics, UCLA, 2004.