

DCCAL - Discrete Cameras Calibration using Properties of Natural Scenes

Milestone 6 Final Report - Scientific Component

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 - June 2013

July 2013

DCCAL - Discrete Cameras Calibration using Properties of Natural Scenes

Milestone 6 - Final Report - Scientific Component

Webpage

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

Objectives

The fundamental objective of the DCCAL project is the research and development of new calibration tools for video cameras. Conventional and discrete cameras are both considered in the project. The term discrete cameras refer to cameras composed by a collection of photocells characterized geometrically by a pencil of 3D optical rays.

Most single camera calibration methodologies rely on geometrical features (e.g. points, lines) registered between known patterns and the image plane (Roger Tsai, CVPR 1986; J. Y. Bouguet, http://www.vision.caltech.edu/bouguetj/calib_doc). However, the need to use known calibration patterns and specialized feature detection and matching algorithms, makes calibration impractical in many situations.

In the DCCAL project we take an alternative approach and use a different kind of data, namely the brightness (or colour) values captured by a moving camera. Each photocell acquires a time series of brightness values, i.e. the so called pixel-stream (Grossmann et al., CVIU 2010). The correlations of the pixel-streams allow estimating inter-pixel angles (see figure 1). The advantages of pixel-streams based calibration are clear: works for a large variety of cameras and requires no specific calibration patterns or procedures.



Fig.1: Pixel-streams based calibration. The angles among the photocells (left) are reconstructed (right).

In summary, the main objectives of the project are:

- Characterization of the effect of relaxing the motion isotropy into the precision and accuracy of the calibration;

Proposal of novel methods for embedding the pixels over a 2-sphere and estimating the topology of an unknown discrete camera;
Proposal of methodology for calibrating cameras forming a network.

Developed Tasks

Project DCCAL encompasses four principal tasks: (i) Calibration of central cameras, bibliography revision and proposal of methodology improvements; (ii) Comparison of the proposed calibration methodologies with other methodologies; (iii) Applying the calibration methodologies to optics classification; and (iv) Calibrating Multiple Central Cameras. In the following we detail the work developed within each of the tasks.

Task 1. Calibration of central cameras (1st year + 1st semester of 2nd year)

Central cameras comprise both conventional and discrete cameras. Conventional cameras are characterized by pin-hole geometric models and are based in CCD/CMOS standard sensors. Discrete cameras are also described by the pin-hole geometric projection model, but are composed by collections of pixels organized as pencils of lines with unknown topologies. Note that distinctly from conventional cameras, discrete cameras can be formed just by some sparse, non regular, sets of pixels (e.g. log-polar CMOS

sensors). Central cameras can be fixed or mobile, and image static scenes or mobile objects. Mobile cameras can move both in translation or rotation, or just rotation (usually in the pan and tilt degrees of freedom). Central cameras can have also the zoom degree of freedom (e.g. pan-tilt-zoom - PTZ - cameras).

Mobile cameras can be auto-calibrated using for example methodologies based in the Kruppa equations. In the work [Greggio11] are estimated the parameters of a conic as a way to identify the location of an object in 3D, i.e. the dual problem of estimating the localization of the camera with respect to the scenario.

Pan-tilt-zoom (PTZ) cameras can be auto-calibrated. Research has been conduced in order to obtain efficient (computational time and memory) calibration methodologies for PTZ cameras including the case of having significant radial distortion [Leite11] (also MSC thesis [MSc-Leite-11]). In particular has been documented the case of calibrating a PTZ camera for various levels of zoom, while having significant radial distortion at minimum zoom level (maximum field of view). Using the odometry of the motors proved to be an important accelerator for the convergence of the iterative calibration process. Figure 2 shows the mosaic obtained by overlapping more than a thousand images acquired by a pan-tilt camera, and shows that a clear overlapping involves estimating and correcting the radial distortion.



Fig.2: Geometric and radiometric calibration of PTZ cameras. Mosaics formed from 1220 images assuming no radial distortion (top) vs correcting radial and radiometric distortions (bottom).

The application of mosaicing images acquired by a PTZ camera shows that one needs precise geometric calibration (good estimation of radial distortion) and, in addition, radiometric calibration in order to mitigate variable image gain and vignetting. Work carried out under the project allowed to develop methodologies for gain equalization over multiple overlapping images acquired at distinct poses [Galego11b] and to correct vignetting [Galego10] [Galego11] (see also MSc thesis [MSc-Galego-11]).

Discrete central cameras, being composed just by a collection of pixels (photocells), are described geometrically by the location of the pixels. More in detail, pixel localization means embedding pixels over a 2-sphere given inter-pixel (photocell) angles. Inter-pixel angles are estimated based on the correlation of the time series (pixel streams) acquired by the pixels. The pixel embedding is usually obtained using Multi-dimensional Scaling (MDS). The need to maintain all-to-all pixel distances implies using very large matrices (number of entries equal to the square of the number of pixels). In order to avoid keeping all-to-all distances, some Multiple Hypothesis Tracking (MHT) techniques have been explored. Preliminary tests of MHT applied to tracking objects in video were published in a national conference [Antunes10], and a supporting software library has been developed and made public to help developing MHT based applications (http://www.multiplehypothesis.com). From this work resulted also an international publication [Antunes11] (see also MSc thesis [MSc-Antunes-11]).

Accepting that the topology of an imaging sensor does not need to be a regular grid opens many novel research and development directions. Research with animals shows that there is a large variety of visual sensor topologies. The topologies of the visual sensors are in many cases associated to the tasks that the animals do most often. Therefore, one question that can be formulated is: given a certain task can one design (automatically) a discrete camera that helps the fulfilment of the task? Research realized under this subject resulted in two journal publications [Ruesch13] and [Ruesch13b], which shown that is possible to obtain sensors whose topologies fit well some specific tasks by using numeric optimization based on some simple biologically plausible principles.

Task 2. Comparing calibration methodologies (2nd semester of the 2nd year)

The task of comparing calibration methodologies is concerned both with conventional central cameras (fixed or PTZ) and discrete cameras.

Calibrating PTZ cameras having radial distortion, traditionally involves the non-linear minimization of the reprojection error. Within DCCAL has been developed a calibration methodology based in the Direct Linear Transformation (DLT) for PTZ cameras that can have significant radial distortion [Galego12]. More precisely, points matched between consecutive images are used to form a least squares problem where radial distortion just adds polynomial terms to the DLT. In this way the DLT is converted into a Polynomial Eigenvector Problem (PEP). Comparing with previous approaches, zoom and radial distortion are considered simultaneously, while keeping an efficient calibration methodology in terms of computational time.

The comparison of calibration methodologies is also present in the MSc dissertation [MSc-Silva-12]. More precisely, one compares calibration using DLT based in normalized points, proposed by R. Hartley, with a lines-based-DLT calibration methodology proposed within the project to handle cameras having, or not having, significant radial distortion. By using lines one has more effective noise filters which make therefore possible surpassing the de facto standard of normalized points DLT. In the MSc thesis [MSc-Silva-12], besides listing various calibration methodologies for conventional fixed cameras, it is also explored the calibration of all cameras considering one universal coordinate system. This aspect is important for the task of calibrating multiple cameras.

Recent results, [Galego13] show that one discrete camera can be auto-calibrated without the need of information taken from a (different) calibrated camera as studied by E. Grossmann, J. Gaspar and F. Orabona (Elsevier CVIU, 2010). Of particular interest is the finding that by considering scenarios where the contour-directions are uniformly distributed and the camera has uniformly distributed motion just in the pan and tilt degrees of freedom, one has a linear relationship between inter-pixel angles and the correlation of the respective pixel-streams.



Fig.3: Discrete camera based in an optic fibers bundle (\sim 2k fibers). Calibration data, image of a laser pointer, i.e. a red circular footprint distorted due to the twisting of the fibers (left). Topology correction and mosaicing of 200 images given the camera odometry.

Figure 3 shows a prototype of a discrete camera. The fibers bundle has about two thousand fibers, which are not organized, i.e. do not form a regular topology. The bundle has about 6mm diameter. The imaging is taken by a conventional camera equipped with a microscope lens attached to one end of the bundle, while the other end of the bundle simply collects light. The calibration methodology is detailed in [Galego13]. By correcting the topology and superposing overlapping images one obtains mosaics as documented in the figure.

Task 3. Application to optics classification (2nd semester of the 2nd year)

Task 3 had a preparatory step in the end of the first year with the collaboration of Escuela de Ingeniería de Antioquia (EIA), Colombia. More precisely, IST received the visit of Mauricio Areas Correa, instructor-researcher at EIA, during the period of

approximately two months. During that period the work at DCCAL comprised the calibration of a constant-angular-resolution lens (fisheye, more than 180deg field of view), and was assembled a setup where the camera moves inside a tubular structure (figure 4).



Fig.4: Wide angle lens calibration helps tubular-Shape Inspection. Camera with lens (FOV $> 180^{\circ}$) and illumination LEDs (left), Mockup tubular shape to inspect (middle-left), Captured image (middle-right) and mosaic of the inside of the tubular shape (right).

After the preparatory step, a MSc dissertation work continued the work over the setup composed by the camera equipped with a wide field of view lens and moving inside the tubular structure, with the purpose of reconstructing the tubular structure [MSc-Tomaz-13]. The work developed in the dissertation consisted of tracking characteristic points and estimating their pose considering the shape of the tubular structure. Robust reconstruction of 3D information requires camera calibration.

In parallel with the work of reconstructing the tubular structure, an auto-calibration process has been developed for a mobile humanoid robot equipped with two cameras [Moutinho12]. In this work, developed in the framework of a PhD, it is explored whether a robot can adapt itself to changes in its physiognomy. In particular, calibration encompasses estimating online backlash in the mechanical joints that link the cameras to the humanoid body. The PTZ calibration work detailed in [Galego12] is expected to provide an online calibration methodology with low complexity (time and memory) as it is based in a least squares problem formulated as a polynomial eigenvector problem.

The classification of a lens or, more precisely, estimating the respective field of view, is an aspect still under research and development. The base assumption is that the camera undergoes pure rotational motion (e.g. pan and tilt) and images a textured scene. Estimating the field of view can therefore be simply demonstrated with conventional cameras (CCD or CMOS sensors). It is however in the case of discrete cameras where the problem of estimating the field of view is more acute and interesting. The discrete camera can be formed, for instance, by a set of optic fibers as in the prototype shown in [Galego13]. However in [Galego13] the fibers bundle has a diameter too thick for allowing pan-tilt motion of the head of the camera. There is currently under construction a new setup where the fibers bundle is much thinner (diameter about 1mm) - see figure 5. By having a flexible head mounted on a pan-tilt unit it is expected that estimating the field of view of the optic system allows the successful demonstration of optics classification.



Fig.5: Outside view of a fiberscope, see cable cut near the pencil (left). Cable of optic fibers, precisely cut end, organized pattern, ~3k fibers (middle). Roughly cut end of the cable implies an unorganized pattern of fibers (right).

Task 4. Calibrating multiple central Cameras (3rd/last year)

The task of calibrating multiple central cameras involves two fundamental aspects: (i) providing for each camera the required calibration data, (ii) creating a global coordinate system. More in detail, (i) provides enough information to compute the pose of a camera relative to a specific coordinate system, and (ii) links into a global coordinate system all specific coordinate systems used for calibrating the cameras individually.

Considering that usually one has no intersection of the fields of view of the various cameras composing a network of cameras, it is normally not possible to use data from one fixed camera to calibrate a neighbour one. In the works [Silva12] and [MSc-Silva-12] is approached problem (i) by considering an auxiliary mobile colour-depth (RGBD) camera which is calibrated. The mobile RGBD

camera images the scenario, allows extracting 3D lines, and corresponding those lines with the ones (2D) imaged by the fixed cameras. These corresponded 3D-to-2D lines allow forming a Direct Linear Transformation (DLT) problem, whose solution is directly the calibration of the fixed camera. Figure 6 shows an example of the methodology, namely the input data, 2D and 3D lines, and the output, i.e. the location of the fixed camera with respect to the mobile camera.



Fig.6: Calibration of a surveillance camera, Axis P1347 (RGB), using a mobile robot equipped with a color-depth camera, Asus X-Tion (RGBD) (a). Lines in the RGBD image (b,c) define 3D lines (d,e). Each line formed directly from the depth map (cyan dots) is filtered using RANSAC (blue and black dots), as shown in (d), where the left/right plot has different/equal scales in the axis. RGBD (e) and RGB lines (f), form the input dataset for DLT-Lines. Decomposing the estimated projection matrix as K[R t], provides the camera position and orientation on the world coordinate system (g).

Problem (ii), creating a global coordinate system, is solved if one has the localization of the auxiliary mobile RGBD camera used to solve problem (i). One way to obtain this information is to perform Simultaneous Localization and Map Building (SLAM) with the mobile camera. The MSc work [MSc-Lucas-13] creates 3D indoor models by fusing RGBD images acquired by a MS-Kinect camera. The images are acquired with steps in the order of half a meter (much sparser than typical SLAM) by considering a Manhattan World assumption. The Manhattan world assumption in essence allows simplifying and improving the precision of an Iterated Closest Point (ICP) step used to register point clouds.

During the project collaboration has been established with Universitat Politècnica de Catalunya (UPC) for calibrating a local network of surveillance cameras. The network camera calibration methodologies developed in DCCAL [Silva12] have been applied to the UPC network of surveillance cameras based on a 3D global map acquired by a mobile robot equipped with a Laser Range Finder (LRF). A study of the uncertainty associated to calibration due to uncertainty in the calibration data was published in [Ortega13]. From the collaboration resulted also a submission of a journal paper (currently under review).

Fulfilled Objectives

-- General Objectives

Experiments and a formal demonstration have shown that the requirement of rotation and translation isotropic movement, found in seminal methodologies (Grossmann et all., CVIU 2010), can be relaxed. In particular has been shown that exists a simple scenario where rotation is enough to estimate inter-pixel angles [Galego13].

Given the inter-pixel angular distances, the embedding of the pixels over a spherical surface can be achieved with Multidimensional Scaling (MDS). There is however the memory complexity problem as the all-to-all pixels distance matrix has a number of entries equal to the square of the number of pixels. Along the project has been verified that some approximated embedding methodologies, such as Landmark ISOMAP, still provide accurate results while being much less complex. Hence it is expected that is possible to calibrate discrete cameras having large numbers of pixels. Figure 7 shows the calibration of a 10⁴ pixels sensor using Landmark ISOMAP. Qualitatively one observes that the reconstructions are topologically correct, despite the fact that Landmark ISOMAP is just an approximation to MDS.



Fig.7: Images used to estimate the topology of a 100×100 sensor (a). Estimated topology after random permutation of the pixel-streams (b,c). Test image before permutation (d). Permutations of pixel-streams in 2×4 blocks, 10×10 blocks, or 100×100 (all) pixels, illustrated on the test image, (e,f,g,h). Estimated topology applied to reconstruct the test image after the three permutations (i,j,k).

Along the project have been tested successfully methods to provide scene based calibration data to fixed cameras. This allows calibrating the cameras of a network, provided that the calibration data is referred to a global coordinate system. In particular has been shown that straight lines (3D and 2D) are effective features for calibrating cameras. The possibly long support regions of the lines allow doing robust calibrations. The fact that lines can form calibration problems as Direct Linear Transformations (DLT) shows that calibration can also be solved efficiently in terms of memory and computations.

-- Specific Objectives

One result of DCCAL is the prototype of a discrete camera based in a twisted bundle of optic fibers (~ 2500 optic fibers; cable diameter of about 6mm). This prototype has the effect of immediately illustrating the need for topological calibration as the image observed is not directly human readable.

The numbers of publications are the following: 2 journal articles (+1 submitted), 9 publications in international meetings and 5 publications in national meetings. The aspects covered in the publications range from pure calibration of discrete cameras, to the calibration of conventional active cameras. Particular attention was given to pan-tilt-zoom cameras as they constitute excellent omnidireccional camera setups which allow simulating a broad range of discrete cameras. Networks of cameras have also been considered. In the following is a summary of the main results obtained.

In the calibration of discrete cameras, a closed form solution has been obtained for estimating inter-pixel angle from pixel-streams correlation in a specific environment [Galego13].

Radiometric aspects, such as vignetting, have been studied and methodologies have been derived to correct it [Galego11]. Radiometric issues may bias significantly auto-calibration based on the similarities of pixel-streams, and is therefore a subject for future studies.

Pan-tilt-zoom cameras, have been proven to auto-calibrate even in the case of having significant radial distortion [Galego12]. These methodologies may be useful for fine calibrating discrete cameras whose topology is already estimated.

Mounting cameras on humanoid robots and using them in feedback loops, require internal knowledge of the calibration. In [Moutinho12] has been proved that zero offsets can be determined to auto-calibrate joints and cameras by using inertial sensors and calibrated cameras. Considering discrete cameras in this setup will involve further work to create compatible coordinate systems.

Calibrating the extrinsic parameters of cameras has been found to be a subject yet under development in the framework of surveillance camera networks. Transporting coordinate systems among cameras is a major concern since GPS is limited in some environments. In [Silva12a] and [Ortega13], "direct linear transformations" have been derived to work with large scene features (lines), and therefore form promise good estimations of the extrinsic parameters of the cameras considering one single, global, coordinate system. As soon as discrete cameras have their intrinsic calibration, extrinsic calibration may in many cases proceed as with conventional cameras.

The work [Ruesch13], summarizing the PhD defence of Jonas Ruesch (EEC / IST / UTL), paves the way for many future developments. Considering that the calibration of discrete cameras in essence enables their use, the next big question is: given a certain task can I design a discrete camera that helps the fulfilment of the task? The published work already answer positively to some simple scenarios.

Besides publications, other relevant results resulted. In particular was obtained one interesting software result namely an implementation of Multiple Target Tracking based in multiple hypothesis (D. Reid, IEEE Trans. on A.C. 1979). A library has been made public to help other applications:

[Antunes11-www] David Antunes, "Multiple Hypothesis Library", http://www.multiplehypothesis.com/

In particular is expected that a sequential embedding of pixels over a spherical surface can be implemented over this library.

Publications

(see PDF files in the webpage of the project)

International journals

[Ruesch13] Predicting visual stimuli from self-induced actions: an adaptive model of a corollary discharge circuit, Jonas Ruesch, Ricardo Ferreira, Alexandre Bernardino, IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, accepted 2012.

[Ruesch13b] A computational approach on the co-development of artificial visual sensorimotor, Jonas Ruesch, Ricardo Ferreira and Alexandre Bernardino, SAGE journal on Adaptive Behavior, accepted 2013

[Ortega13-submitted] Calibration of an Outdoor Distributed Camera Network with a Laser Range Map, Agustin Ortega, Manuel Silva, Ernesto H. Teniente, Ricardo Ferreira, Alexandre Bernardino, José Gaspar, Juan Andrade-Cetto, journal of Machine Vision and Applications / Springer (submitted)

International conferences

[Antunes11] Multiple Hypothesis Tracking in Camera Networks, David M. Antunes, Dario Figueira, David M. Matos, Alexandre Bernardino, José Gaspar, Int. WS on Omnidirectional Vision, Camera Networks and Non-classical Cameras (OMNIVIS) 2011, in conjunction with ICCV 2011, 6-13 November 2011, Barcelona, Spain, pp367-374.

[Galego11] Vignetting Correction for Pan-Tilt Surveillance Cameras, Ricardo Galego, Alexandre Bernardino, José Gaspar, Int. Conf. on Computer Vision Theory and Applications (Visapp), Vilamoura, Portugal, March 5-7, 2011, pp638-644.

[Greggio11] Monocular vs Binocular 3D Real-Time Ball Tracking From 2d Ellipses, Nicola Greggio, José Gaspar, Alexandre Bernardino, José Santos-Victor, Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO), Noordwijkerhout, Netherlands, July, 2011, vol2 pp67-73.

[Galego12] Auto-calibration of Pan-Tilt Cameras including Radial Distortion and Zoom, Ricardo Galego, Alexandre Bernardino, José Gaspar, International Symposium on Visual Computing (ISVC), July 16-18, Crete, Greece, 2012.

[Moutinho12] 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, IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS) 2012.

[Ruesch12] An Approach Toward Self-Organization of Artificial Visual Sensorimotor Structures, Jonas Ruesch, Ricardo Ferreira, Alexandre Bernardino, Proc. of BICA 2012 - International Conference on Biologically Inspired Cognitive Architectures, Palermo, Italy, 2012

[Silva12] Camera Calibration using a Color-Depth Camera: Points and Lines Based DLT including Radial Distortion, M. Silva, R. Ferreira, J. Gaspar, In WS in Color-Depth Camera Fusion in Robotics, held with IROS 2012, October 7, 2012, Vilamoura, Portugal.

[Galego13] Topological Auto-Calibration of Central Imaging Sensors, R. Galego, R. Ferreira, A. Bernardino, E. Grossmann and J. Gaspar, Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA) 2013.

[Ortega13] Uncertainty in Network Camera Calibration based in LIDAR, A. Ortega, R. Galego, R. Ferreira, A. Bernardino, J. Gaspar and J. Andrade-Cetto, European Conference on Mobile Robots (ECMR) 2013 (accepted).

National conferences

[Antunes10] Multiple Hypothesis Group Tracking in Video Sequences, David Antunes, David Martins de Matos, José Gaspar, In 16th Portuguese Conference on Pattern Recognition (RecPad 2010), Vila Real, Portugal, October 2010.

[Galego10] Surveillance with Pan-Tilt Cameras: Background Modeling, Ricardo Galego, Alexandre Bernardino, José Gaspar, In 16th Portuguese Conference on Pattern Recognition (RecPad 2010), Vila Real, Portugal, October 2010.

[Galego11b] Enforcing Consistency of Image Gains in Panoramic Mosaics, Ricardo Galego, Ricardo Ferreira, Alexandre Bernardino, José Gaspar, In 17th Portuguese Conference on Pattern Recognition (RecPad 2011), Porto, Portugal, October 2011.

[Leite11] Auto-Calibration of Pan-Tilt-Zoom Cameras: Estimating Intrinsic and Radial Distortion Parameters, Diogo Leite, Alexandre Bernardino, José Gaspar, In 17th Portuguese Conference on Pattern Recognition (RecPad 2011), Porto, Portugal, October 2011.

[Silva12b] Effect of Image Resolution on DLT-Lines Camera Calibration Including Radial Distortion, M. Silva, R. Ferreira, J. Gaspar, In 18th Portuguese Conference on Pattern Recognition (RecPad 2012), Coimbra, Portugal, October 2012.

MSc dissertations

[MSc-Galego-11] Geometric and Radiometric Calibration for Pan-Tilt Surveillance Cameras Ricardo Galego nmec=52118 (superv: J. Gaspar, Alexandre Bernardino) IST/MEEC 2010/2011 2°Sem. (concl. 12-07-2011)

[MSc-Antunes-11] Multi-sensor based Localization and Tracking for Intelligent Environments David Antunes nmec=58634 (superv: J. Gaspar, David Matos (DEI)) IST/MEIC 2010/2011 2°Sem. (concl. 02-11-2011)

[MSc-Leite-11] Target Tracking with Pan-Tilt-Zoom Cameras Diogo Leite nmec=58094 (superv: J. Gaspar, Alexandre Bernardino) IST/MEEC 2010/2011 2°Sem. (concl. 22-11-2011)

[MSc-Silva-12] NetCam - Network Cameras Calibration Manuel Silva nmec=48026 (superv: J. Gaspar) IST/MEEC 2011/2012 1°Sem. (concl. 14-11-2012)

[MSc-Lucas-13] Building World Representations using Color-Depth Cameras Ricardo Lucas nmec=56719 (superv: J. Gaspar, R. Ferreira) IST/MEEC 2011/2012 1°Sem. (concl. 14-05-2013)

[MSc-Tomaz-13] Mosaicing the interior of tubular structures João Tomaz nmec=58134 (superv: J. Gaspar, R. Ferreira) IST/ MEEC 2011/2012 1°Sem. (concl. 21-06-2013)