Contributions to Real-time Metric Localisation with Wearable Vision Systems

Daniel Gutiérrez-Gómez

Dept. de Informática e Ingeniería de Sistemas , I3A – Universidad de Zaragoza, María de Luna 1 50018, Zaragoza, Spain
Advisor: José Jesús Guerrero
Date and location of PhD Thesis defense: April 26th, 2016, Universidad de Zaragoza

Abstract

Under the rapid development of electronics and computer science in the last years, cameras have become omnipresent nowadays, to such extent that almost everybody is able to carry one at all times embedded into their cellular phone. What makes cameras specially appealing for us is their ability to quickly capture a lot of information of the environment encoded in one image or video, allowing us to immortalize special moments in our life or share reliable visual information of the environment with other persons. However, while the task of extracting the information from an image may by trivial for us, in the case of computers complex algorithms with a high computational burden are required to transform a raw image into useful information. In this sense, the same rapid development in computer science that allowed the widespread of cameras has enabled also the possibility of real-time application of previously practically infeasible algorithms.

Among the current fields of research in the computer vision community, this thesis is specially concerned in metric localisation and mapping algorithms. These algorithms are a key component in many practical applications such as robot navigation, augmented reality or reconstructing 3D models of the environment.

The goal of this thesis [*] is to delve into visual localisation and mapping from vision, paying special attention to conventional and unconventional cameras which can be easily worn or handled by a human (Fig. 1). In this thesis I contribute in the following aspects of the visual odometry and SLAM (Simultaneous Localisation and Mapping) pipeline:

- **Generalised Monocular SLAM:** State of the art visual SLAM algorithms are usually designed for conventional cameras, which can be modelled by a simple pin-hole camera model but have a narrow field of view. Vision systems consisting in a combination of a camera and a conic-shaped mirror, catadioptric cameras, offer a larger field of view, but in turn require of the use of a more complex projection model, which generalises the projection of central projection systems. Our proposal in this field is the adaptation to catadioptric cameras of a real time monocular SLAM system initially designed for conventional cameras. *Associated publications:* [1]

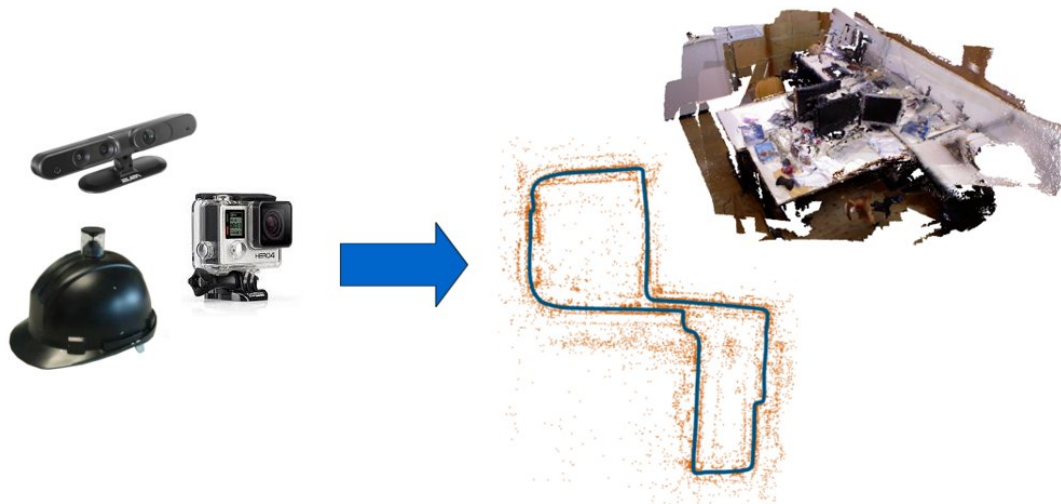[*]`http://webdiis.unizar.es/~danielgg/PhDThesis_DGG.pdf`

Figure 1: Small wearable and portable cameras can be used for precise localisation and building accurate 3D models of the environment.

- **Scale problem in monocular vision:** One problem of localisation and mapping algorithms for monocular systems is their inability to provide the real scale of the camera motion and the observed map. To solve this problem, extra information has to be obtained either from additional sensors or from known or assumed priors about the environment. Our proposal in this field is an algorithm for scale estimation and scale drift avoidance when monocular SLAM is performed with a wearable camera by capturing the walking speed of the user, from the step frequency. *Associated publications:* [10], [2], [6]

- **Dense RGB-D SLAM:** Recent algorithms for odometry estimation with RGB-D sensors, estimate the camera motion by performing pixelwise minimisation of the photometric and/or the geometric error. However, in many cases important properties of the error model of the depth sensor are ignored, which could affect the performance of odometry computation. In this thesis we propose a direct RGB-D odometry method which uses the inverse depth for the parametrisation of the geometric error, whose probability distribution is closer to the error model of a depth sensor, improving state of the art results. Building upon this odometry method we further developed a RGB-D SLAM system for localisation and dense mapping for 3D cameras which is available for download [†]. *Associated publications:* [7], [9], [5]

- **Robust place recognition:** In visual SLAM, a place recognition module is a key component for camera relocalisation when lost, or to close loops at revisited areas. However during place recognition, it is frequently assumed that a scene does not suffer changes between visits, which can be a source of fails, affecting robustness of place recognition algorithms. Our work in this field and in the context of RGB-D sensors proposes the pruning of scene parts with a high entropy in the surface normals in order to increase the robustness of place recognition under long term changes in the scene. *Associated publications:* [8]

- **Pose-graph optimisation:** In SLAM approaches, pose-graph optimisation is a frequently used approach to enforce loop constraints between pairs of camera poses. Pose-graph problems usually mix the camera rotation and translation in the same optimisation problem. The different units in which rotation and translation are measured make necessary a normalisation of these variables, but it is not always possible to do this in a rigorous way. In addition to this, it seems also unnecesary having to optimise the orientation together with the translation, when the main objective might be bending the curve which represents the

---

[†]`https://github.com/dangut/RGBiD-SLAM`

trajectory of the camera. Stemming from these observations we propose a reparametrisation of the pose-graph problem removing the orientations from the optimised vector and making possible the optimisation only on the position part of the poses. *Associated publications:* [3], [4]

References

[1] D. Gutierrez, A. Rituerto, J. M. M. Montiel, and J. J. Guerrero. Adapting a real-time monocular visual slam from conventional to omnidirectional cameras. In *11th OMNIVIS, held with Int. Conf. on Computer Vision (ICCV)*, pages 343–350, 2011.

[2] D. Gutierrez-Gomez and J. J. Guerrero. Scaled monocular slam for walking people. In *Int. Symp. on Wearable Computing (ISWC)*, pages 9–12, 2013.

[3] D. Gutierrez-Gomez and J. J. Guerrero. Curve-graph odometry: Removing the orientation in loop closure optimisation problems. In *Int. Conf. on Intelligent Autonomous Systems (IAS)*, pages 263–275, 2014.

[4] D. Gutierrez-Gomez and J. J. Guerrero. Curve-graph odometry: Orientation-free error parameterisations for loop closure problems. *Robotics and Autonomous Systems (RAS)*, 74(Part B):299–308, 2015.

[5] D. Gutierrez-Gomez and J. J. Guerrero. Rgbid-slam for accurate real-time localisation and 3d mapping. *Technical Report*, 2016.

[6] D. Gutierrez-Gomez and J. J. Guerrero. True scaled 6 dof egocentric localisation with monocular wearable systems. *Image and Vision Computing*, 52:178–194, 2016.

[7] D. Gutierrez-Gomez, W. Mayol-Cuevas, and J. J. Guerrero. Inverse depth for accurate photometric and geometric error minimisation in rgb-d dense visual odometry. In *Proc. IEEE/RSJ Int. Conf. on Robotics and Automation (ICRA)*, pages 83–89, 2015.

[8] D. Gutierrez-Gomez, W. Mayol-Cuevas, and J. J. Guerrero. What should i landmark? entropy of normals in depth juts for place recognition in changing environments using rgb-d data. In *Proc. IEEE/RSJ Int. Conf. on Robotics and Automation (ICRA)*, pages 5468–5474, 2015.

[9] D. Gutierrez-Gomez, W. Mayol-Cuevas, and J. J. Guerrero. Dense rgb-d visual odometry using inverse depth. *Robotics and Autonomous Systems (RAS)*, 75(Part B):571–583, 2016. Special Section on 3D Perception with PCL.

[10] D. Gutierrez-Gomez, L. Puig, and J. J. Guerrero. Full scaled 3d visual odometry from a single wearable omnidirectional camera. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robot Systems (IROS)*, pages 4276–4281, 2012.