Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 50-1 (2017) 10274–10279

Autonomous exploration with prediction of the quality of vision-based localization

Hélène Roggeman * Julien Marzat *
Anthelme Bernard-Brunel * Guy Le Besnerais *

* ONERA – The French Aerospace Lab, F-91123, Palaiseau, France.
e-mail: firstname.lastname@onera.fr.

Abstract: This paper presents an algorithm to perform autonomous exploration with robotic platforms equipped with a stereo-vision system in indoor, unknown and cluttered environments. The accuracy of the vision-based localization depends on the quantity of visible features in the scene captured by the cameras. A Model Predictive Control approach permits to perform the exploration task with obstacle avoidance and taking into account the quality of the scene in order to avoid areas where the visual odometry is likely to fail. Experiments were carried out with a mobile robot to assess the improvement in localization accuracy and coverage for exploration.

Keywords: Autonomous vehicles, Autonomous exploration, Stereo-vision, Model Predictive Control

1. INTRODUCTION

Autonomous exploration allows to build maps of unknown environments without human intervention. This can be interesting, for instance, for search-and-rescue missions in dangerous areas. UAVs and mobile robots can be complementary to perform this type of missions. Their success depends on the accurate localization of the robot. The localization algorithms are designed to be embedded on UAVs and mobile robots. The UAVs have a low payload that implies a limitation of the number of embedded sensors. Moreover, these exploration missions are most of the time conducted in indoor environments, where no global localization systems, such as GPS, are available. Thus, a good solution to address these constraints is to install camera sensors, because they are lightweight, inexpensive and the images provide a high amount of information. A stereo rig composed by two cameras with a visual odometry algorithm allows to compute the localization of the platform from features extracted in the images (Sanfourche et al., 2013).

All passive vision-based navigation systems are likely to fail in low-textured environments, where the lack of interest points in the images prevent the computation of a good localization. The aim of this work is to develop an autonomous exploration system with a command strategy that seeks to avoid the situations where the robot is likely to loose its localization due to lack of texture in the environment. We experiment with a mobile robot in order to validate the developed strategy, but the overall system can easily be adapted for a UAV.

1.1 Related Work

The issue of evaluating the scene quality for vision-based localization was studied in the context of navigation between waypoints. Sadat et al. (2014) take into account the richness of the environment for the path planning, they used a RRT* algorithm and add a criterion which computes a viewpoint score based on the density of the triangle in the 3D mesh of the environment. Mostegel et al. (2014) evaluate the quality of the camera motions for the localization quality and the possibility of seeing new features. A combination of criteria on the features is computed and used to define if a future position will be well suited for the localization. In these two references, the robotic platforms are equipped with a monocular camera, which involves specific depth estimation issues to compute the localization. In previous work (Roggeman et al., 2016), we focused on the same problematic but with stereo-vision system which makes it possible to obtain directly information about the depth of the points. The present paper elaborates on this work by addressing autonomous exploration missions.

Some authors were interested in the active reduction of the uncertainty during an autonomous exploration mission: Bourgaul et al. (2002) aimed at exploring and building an accurate map of the environment with a mobile robot equipped with a laser range finder selecting the control actions that maximize the accuracy of the localization. Bryson and Sukkarieh (2008) developed an information-based path planning method for a UAV. It plans a trajectory which maximizes the accuracy of the map and the vehicle location during the exploration of unknown areas. It is based on the computation of the entropic information gain before and after taking an action. This system can be used with vision sensors.

1.2 Problem statement

This paper describes a complete architecture for autonomous exploration on robotic platforms. The mission considered in this paper is the exploration by a mobile robot of an unknown and cluttered environment which presents some low textured areas. The robot has to complete its exploration mission avoiding the obstacles in the

10.1016/j.ifacol.2017.08.1479

room and keeping an accurate localization during all the mission.

To achieve this mission, it is necessary to localize the robot and map the environment. The localization is estimated from the images of the stereo-vision system by the visual odometry algorithm (Sanfourche et al., 2013), described in Section 2.1. The mapping task is possible thanks to a RGB-D sensor, explanations can be found in Section 2.2. For the detection of the low-textured areas, a criterion based on the prediction of the amount of information available in the future images permits to define if a position is appropriate for the localization. This criterion is explained in Section 3. The Model Predictive Control (MPC) strategy presented in Section 4, uses the information of mapping, localization and quality of texture of the scene in order to find the optimal control to send to the motors. Experiments were made in real situations and the results are reported in Section 5.

2. PERCEPTION

2.1 Visual odometry algorithm: eVO

The localization of the robot is ensured by a visual odometry algorithm, using the images providing by the stereo rig. The aim is to estimate the localization of the robot from its starting point. In our experiments, we use eVO (Sanfourche et al., 2013) but other algorithms could be used as well (Klein and Murray, 2007).

The following is a brief description of the algorithm. Two tasks are working in parallel:

- Mapping: this task consists in providing a map with a limited number of points localized in space. Interest points (Harris and Stephens (1988) or FAST Rosten and Drummond (2006)) are extracted from both images and matched. The 3D position of the points in space is then computed by a triangulation (see Eq. 1).
- Localization: The matching between the 2D points in the left image and the 3D points in the map derives from the temporal tracking of 2D points with KLT (Shi and Tomasi, 1994). The position and orientation of the left camera are computed by minimization of the reprojection error, within a RANSAC procedure (Fischler and Bolles, 1981).

The 3D points computed during the mapping task will serve to evaluate the ability of the robot to localize itself from a given position, see Section 3.

2.2 Environment reconstruction

For the obstacle avoidance and the exploration tasks, it is necessary to have an occupancy map of the environment. A Kinect sensor is installed on the robot and gives a 3D representation of the environment in a 3D point cloud format. We first remove the ground plan, using a RANSAC method. Then, the obtained point cloud is transformed into an Octomap model (Hornung et al. (2013)), which is a representation of the volumetric occupancy. Finally, the Octomap is projected onto the ground plane and two 2D maps are created: an exploration map with the explored and unexplored areas and an obstacle map with the occupied and free areas. At time t_n, the exploration map is represented as a matrice denoted $G(n)$ whose elements $g_{i,j}$ are 0 or 1: 0 if the location is not explored, 1 if it is explored.

3. VISUAL QUALITY FOR THE LOCALIZATION

From the 3D points extracted by the odometry algorithm and with a known future position, the proposed visual quality criterion is derived from the prediction of the number of visible points in the future images, considering the uncertainty (Roggeman et al., 2016). The whole process is illustrated in Figure 1 and described in the following.

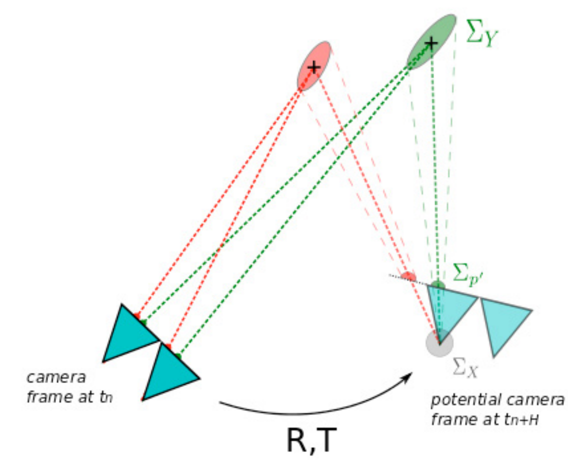

Fig. 1. Two 3D points are triangulated with their covariances at time t_n. The two points are projected onto the camera plane after a displacement (R, T): the green one is predicted to lie in the image whereas the red one is outside.

3.1 Future point position

From the 2D points extracted in the stereo images, the position of a 3D point in the current camera frame, defined by the position of the left camera, is given by

$$Y = (x, y, z)^T = \Pi^{-1}(u, v, d) = \frac{-b}{d} \cdot \begin{pmatrix} u - u_0 \\ v - v_0 \\ \alpha \end{pmatrix} \quad (1)$$

(u, v) is the 2D position of the point in the image and d is the disparity. b denotes the baseline between the left and right camera, α, the focal length and (u_0, v_0) are the coordinates of the principal point.

A change of basis is necessary in order to express the position of the 3D point in the future camera frame.

$$\tilde{Y'} = P \cdot \tilde{Y} = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \cdot (x, y, z, 1)^T \quad (2)$$

R and T are respectively the rotation and translation between the two frames.

The 3D point is then projected on the future image plane, corresponding to the camera frame.

$$\begin{aligned} p &= \Pi(Y) \\ \tilde{p} &= (u, v, 1)^T = z^{-1} K \cdot Y \end{aligned} \quad (3)$$