Features detection for Stereo Visual Odometry

Titus Iulian Ciocoiu[1] Florin Moldoveanu[1]

Abstract - **Estimating its ego-motion is one of the most important capabilities for an autonomous mobile platform. Without reliable ego-motion estimation no long-term navigation is possible. Besides odometry, inertial sensors, DGPS, laser range finders and so on, vision based algorithms can contribute a lot of information. Stereo odometry is a vision based motion estimation algorithm that estimates the ego-motion of a stereo camera through its environment by evaluatingthe captured images. In this paper, we want to give an integrated overview of stereo odometry and the accompanying literature. We want to emphasize the fact that stereo odometry is a chain of several single subprocesses where each relies on its predecessor's results. A variety of exchangeable methods for each of these subprocesses is available.The key to a more accurate and efficient stereo odometry lies in an integrated analysis of its single subprocesses and the many algorithms available.**

Keywords - **Stereo, Visual Odometry, Keypoints features depth triangulation, RANSAC algorithm**

I. Introduction

Motion estimation is one of the most important capabilities for autonomous mobile platform. Without motion estimation, long term navigation could not be performed. Besides visual odometry, inertial sensors, DGPS, vision based algorithms can provide a lot of information. Stereo odometry is an algorithm based on estimating the motion of a stereo camera system evaluating the images caught by the two cameras. This study is meant to provide an overview of odometric systems and to find the best feature detection algorithm from the related literature. The stereo odometry is a chain of multiple unique subprocesses where every process is based on the predecessor subprocess, therefore for each subprocess are available multiple solutions. The key to create a very precise visual stereo odometric system is completely analyzing these unique subprocesses and other available algorithms for solving all the steps[3]. First of all, why do we need stereo odometry in robotics as well in automotive industry? Classic odometry which calculates robots movements depending on robot's wheel number of rotations it has errors due to wheel slip.

Titus Iulian Ciocoiu
"Transilvania" University of Brasov
Romania

Florin Moldoveanu
"Transilvania" University of Brasov
Romania

GPS information might not be available at the desired quality. The robot may work in out of GPS signal areas like forests or mines or the robots operates on a planet that doesn't have its set of GPS satellites, like Mars.[4] Cheng gives an interesting perspective very important for stereo odometry during NASA MER missions with Spirit and Opportunity robots.

As a conclusion, visual stereo odometry systems are topical systems that in future could successfully replace the satellites that provides data for GPS systems, in industrial robotic field, as well in automotive[6].

II. Contribution to Sustainability

Stereo Visual Odometry Systems, as part of Active Vision Systems, is an important concept still to be integrated and supported by novel technologies, promising to expand into most of today's domains, including automotive, autonomous driving, robotics and directly aiming for a new concept, driving and navigation systems with no need of satellites connectivity and GPS position data.
A sustainable key feature that our system provides is given by the very low energy consumption compared with the actual GPS systems, which is caused by eliminating the permanent connectivity between the navigation system and satellites. To expand functionality, we propose two features algorithms that properly manage to help the system to have a much better accuracy. The small dimensions of our proposed stereo odometry system along with its developed control system make the device ideal for future autonomous mobiles needs. Our proposed system manages to reduce the main disadvantage of most of the present day navigation systems – the positioning error of the stereo visual odometry system which in our case is significant lower compared with same parameter of the modern navigation systems in the market.
The goal of the presented system, algorithms and tests is to obtain preliminary data for a complete independent navigation and mapping system that we intend to develop.
The proposed system will help future development of already existing domains by opening a path to new ideas and trends for self-sustained and advanced supportive environments. We anticipate that the results presented in this paper will also contribute to autonomous driving evolution, which represent in our view a small step for sustaining future technologies' development.

III. Experimental Results

A. Stereo Visual Odometry System

In automotive and computer vision, visual odometry is the process of determining the position and orientation of a robot by analyzing the associated camera images.

The project was developed in multiple steps, as it follows. First step was to develop the mechanical model necessary for sustaining the servo motors and web cameras and afterwards it was developed the board data acquisition and control servo motors.

The experimental prototype is made of (Fig.1):

a) acquisition and control board consists of: Supply via a laptop charger, one LM7805 voltage regulators, MAX 232 Integrated Circuit ,ATmega8 microcontroller, DB-9 connector;

b) 2 Futaba S3305 servo motors

c) 2 Logitech Quick Cam Pro webcams 9000;

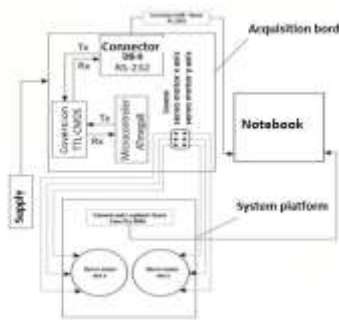d) The functional block diagram is the following:

Fig.1. General block diagram of the entire system active vision;

First step is to detect key points for each captured frame from the cameras system with two different features detection algorithm (ORB, FAST). After that we eliminate outliers points using RANSAC algorithm and performs stereo correlation.

With the remaining points, coordinate transformation is performed from 2D (x,y) to 3D(x,y,z) with the triangulation method (Fig.2).

Having the camera parameters and using the triangulation method, the 3D coordinates of the points are obtained (x_l-coordinate of the feature from the left image, x_r- coordinate of the feature from the left image)[2].

$$Z=bf/(x_l-x_r)$$

The next step is to use the decomposition (SVD method) in order to obtain the values of the rotation matrix (R) and translation (T) of the cameras. To see more easily if the system calculates the correct system trajectory, we introduced additional calculation of the 3 angles of the camera system (roll, pitch, yaw).
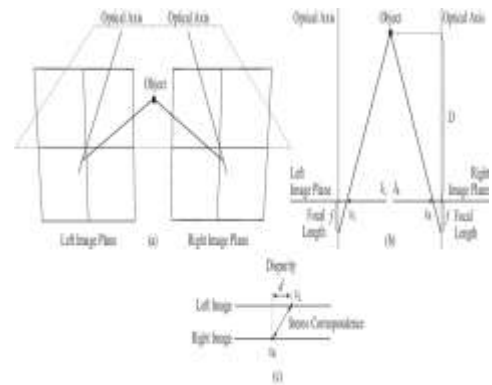

Fig.2. Triangulation method and stereo geometry f- focal length, x_l- coordinate of the feature from the left image, x_r-coordinate of the feature from the right image.

To be able to compare the odometry system result, meanwhile the latitude and the longitude of the stereo cameras system are saved for each frame. With those data the system trajectory is created. These trajectories are compared in order to determine if the odometry system error is approximately equal to 5% [7].

B. Features detection Algorithms

Feature matching is at the base of many computer vision problems, such as object recognition or structure from motion. Current methods rely on costly descriptors for detection and matching. In this paper, we propose to use a very fast binary descriptor based on BRIEF, called ORB, which is rotation invariant and resistant to noise. We demonstrate through experiments how ORB is at two orders of magnitude faster than SIFT or SURF, while performing as well in many situations [1].

In the first algorithm we evaluate the combination of oFAST and rBRIEF, which is called ORB, using two datasets: images with synthetic in-plane rotation and added Gaussian noise, and a real-world dataset of textured planar images captured from different viewpoints. For each reference image, we compute the oFAST keypoints and rBRIEF features, targeting 500 keypoints per image. [5][9]

In the second algorithm we start by detecting FAST points in the image. FAST takes one parameter, the intensity threshold between the center pixel and those in a circular ring about the center. We use circular radius of 9, which has good performance. FAST does not produce a measure of cornerness, and we have found that it has large responses along edges. We employ a Harris corner measure to order the FAST keypoints. For a target number of keypoints, we first set the threshold low enough to get more than N keypoints, then order them according to the Harris measure, and pick the top N points. We employ a scale pyramid of the image, and produce FAST features at each level in the pyramid [9].

C. Testing Method

In order to validate the stereo odometry system's functionality and to determine which of the features detection methods is the adequate solution for further development we conducted a testing session that included recording and comparison between some parameters like (number of detected points, inlier %, time per frame (ms)) and a questionnaire based evaluation. The inliens are the good features that was detected in the images that was caputerd from the stereo odometry system.

The method is as follows. We first set up a training set of some 300k keypoints, drawn from images in the PASCAL 2006 set [8]. We also enumerate all possible binary tests drawn from a 31×31 pixel patch. Each test is a pair of 5×5 sub-windows of the patch. If we note the width of our patch as wp = 31 and the width of the test sub-window as wt = 5, then we have $N = (w_p − w_t)^2$ possible sub-windows. We would like to select pairs of two from these, so we have N/2 binary tests. We eliminate tests that overlap, so we end up with M = 205590 possible tests.

The main algorithm steps are:

1. Run each test against all training patches.

2. Order the tests by their distance from a mean of 0.5, forming the vector T.

3. Greedy search:

(a) Put the first test into the result vector R and remove it from T.

(b) Take the next test from T and compare it against all tests in R. If its absolute correlation is greater than a threshold, discard it; else add it to R.

(c) Repeat the previous step until there are 256 tests in R. If there are fewer than 256, raise the threshold and try again.

At the end of the testing session the learned tests show better diversity and lower correlation[8].

D. Test Results

The average values of the features detection parameters number of detected points, inlier %, time per frame (ms) for ORB and FAST algorithms are represented in Table 1. Standard errors and standard deviations are represented as well for each case.

		Inlier %	N Number of detected points	Time per frame (ms)
1st Algorithm ORB	Value	45.8	789	15.3
	Std. Error	1.212	-	-
	Std. Deviation	1.427	-	-
2nd Algorithm FAST	Value	28.6	795	217.3
	Std. Error	1.514	-	-
	Std. Deviation	2.528	-	-
3rd and fourth Algorithm SURF and SIFT	Value	30.2	714	5228.7
	Std. Error	2.114	-	-
	Std. Deviation	2.142	-	-

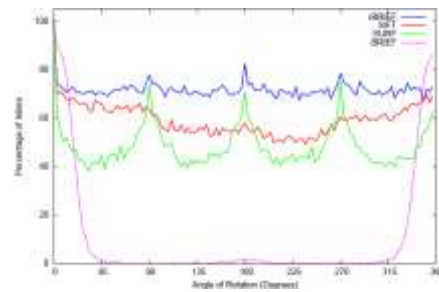**Table 1.** Comparative tests for feature detections algorithms.

Fig. 3. Matching performance of SIFT, SURF, BRIEF with FAST, and ORB (*oFAST +rBRIEF*) under synthetic rotations with Gaussian noise of 10.

Fig. 3 shows the results for the synthetic test set with added Gaussian noise of 10. Note that the standard BRIEF operator falls off dramatically after about 10 degrees. SIFT outperforms SURF, which shows quantization effects at 45-degree angles due to its Haar-wavelet composition. ORB has the best performance, with over 70% inliers. ORB is relatively immune to Gaussian image noise, unlike SIFT. If we plot the inlier performance vs. noise, SIFT exhibits a steady drop of 10% with each additional noise increment of 5. ORB also drops, but at a much lower rate (Fig.4). [9]
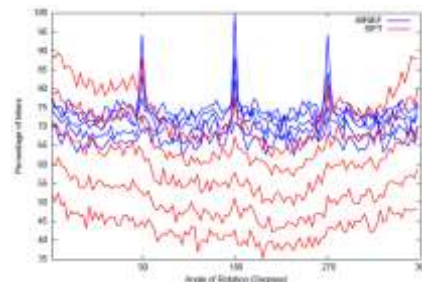

Fig. 4. Matching behavior under noise for SIFT and rBRIEF. The noise levels are 0, 5, 10, 15, 20, and 25. SIFT performance degrades rapidly, while rBRIEF is relatively unaffected.

E. Test Discussion

The results noted in Table 1 show that all the parameters, (number of features, inlines), are higher for the first algorithm case and also the time per frame (ms) is better in the first algorithm (ORB).

On the second algorithm the number of the features that was detected it is approximately the same like in the first algorithm, but the time of processing it is much bigger than in the ORB algorithm.

Even though we have implemented two different features detection algorithms with fair results, we can sustain that a proper detection is also dependent of system hardware

(cameras) and especially of data transmission between the odometry system and acquisition board. A slower data transmission rate is a major drawback that can be overcome only by a fast and precise algorithm.

We can conclude that an ideal algorithm has to balance inlines %, time per frame (s) parameter and also N- number of detected features. This means that the system must react very fast to any modification in the image. If we have a very good inlines percentage the time per is not the proper one. Thereby it should be a compromise between inlines percentage and the speed of the stereo odometry system.

IV. Conclusion and Future Work

Stereo visual odometry systems remain one of the major research issues for the Autonomus driving and navigation. Creating a totally independent system for the user like an alternative for the standard navigation systems is an important goal to be researched still.

The question, if the time of processing of the feature detection algorithm it is enough for real time results. In this paper we focus on providing some answers to this question.

The results of the comparison between those algorithms (ORB, SURF, SIFT) show us that the best results we have with the ORB algorithm which is resistant to noise.

A classic stereo odometry system purposes and the resulting system hardware and software components are presented. Two features detection algorithms needed for odometry systems are also described.

As a future development direction we intend to apply a fuzzy logic control algorithm in order to improve the system accuracy. Also we intend to implement a new algorithm with a better accuracy to convert the points from 2D to 3D. Our intention is to research and develop a complete stereo odometry system with application in actual and future areas of interest.

Acknowledgment

References

[1] 1. Cojocariu, D. „Acquisition, processing and image recognition", Ed. University, Craiova Romania (2002).

[2] 2. Corke, P.I. „Visual Control of Robots: A High-performance Visual Servoing", Research Studies Press Ltd., John Wiley & Sons Inc, Great Britain (1996).

[3] 3. Gary, B. , Adrian, K. „Learning Computer Vision with the OpenCV library", O'Reilly,

[4] United States of America (2008).

[5] 4. Constantin, V. „Image analyze and processing" , Ed. Politehnica (1999).

[6] 5. Cyganek, B., Siebert, J. P. „An Introduction to 3D Computer Vision Techniques and Algorithms", John Wiley & Sons, Great Britain (2009).

[7] 6. Jimenez, E.V.C., Navarro, D.Z., Rojas, R. „Intelligent Active Vision Systems for Robots", Cuvillier Verlag, Gottingen, Germany (2007).

[8] 7. Kisacanin, B., Bhattacharyya, S.S., Chai, S. „Embedded Computer Vision", Springer-Verlag, London, United Kingdom (2009).

[9] 8. Szeliski, R. „Computer Vision: Algorithms and Applications", Springer-Verlag New York Inc, New York (2008).

9. Rublee E., Rabaud V. , Konolige K., Bradski G. "ORB: an efficient alternative to SIFT or SURF", Willow Garage, Menlo Park, California.

About Author (s):

Titus Iulian Ciocoiu is a second year doctoral student active in the System Engineering research field at the "Transilvania" University of Brasov, Romania. The working title of her thesis is "Security in digital electronic systems". Her research is focused on digital image processing and active vision systems. He holds a diploma in Automatic and Applied Informatics and a master in Advanced Automatic Systems and Information Technology both from the "Transilvania" University of Brasov.

Florin Moldoveanu received the B. Sc., and Ph. D. degrees in electrical engineering from Transilvania University of Brasov, Romania, in 1975 and 1998, respectively. He is currently Professor as part of the Department of Automation and Information Technology, Faculty of Electrical Engineering and Computer Science, Transilvania University of Brasov. His main research interests focus on digital circuits, discret event systems, sliding mode control, digital image processing