The Registration System for the Evaluation of Indoor Visual SLAM and Odometry Algorithms

Received 10th October 2012; accepted 22nd November 2012.

Adam Schmidt, Marek Kraft, Michał Fularz, Zuzanna Domagała

Abstract:
This paper presents the new benchmark data registration system aimed at facilitating the development and evaluation of the visual odometry and SLAM algorithms. The WiFiBOT LAB V3 wheeled robot equipped with three cameras, XSENS MTi attitude and heading reference system (AHRS) and Hall encoders can be used to gather data in indoor exploration scenarios. The ground truth trajectory of the robot is obtained using the visual motion tracking system. Additional static cameras simulating the surveillance network, as well as artificial markers augmenting the navigation are incorporated in the system. The datasets registered with the presented system will be freely available for research purposes.

Keywords: *SLAM, Visual Odometry, Benchmark*

1. Introduction

In recent years, the image processing algorithms have been playing an increasing role in the autonomous robot navigation. The fast and reliable image detectors and descriptors of prominent image features and the accurate algorithms for ego-motion estimation facilitated the development of successful visual odometry [1, 2] and simultaneous localization and mapping (SLAM) systems [3, 4].

However, the further progress of the visual robot navigation is hindered by the lack of the widely available benchmark datasets with a reliable ground truth data necessary to evaluate and compare the performance of the developed algorithms. Over the last years, there have been several attempts to gather such data. During the Rawseeds Project, [5] a mobile robot was used to gather multi-sensory data from both the indoor and outdoor runs. The data included video sequences from the onboard cameras, measurements from the inertial measurement unit (IMU), laser scans and the ground truth (GT) of the robot's trajectory. In [6] the authors presented benchmarking data for the outdoor SLAM. The video sequence was recorded by an outdoor robot and the approximate GT trajectory was based on the GPS measurements. Unfortunately, both the datasets suffer from the simplifying assumption that the robot's movement was purely planar.

The sequences gathered at the Freiburg University [7] consist of the RGB-D data recorded with the Kinect sensor and the corresponding 3D GT trajectory of the sensor. Unfortunately, the video sensor of the Kinect has low resolution and only a single camera was used in the experiments.

An alternative approach to the SLAM benchmarking was presented in [8]. The authors proposed a system rendering completely artificial video sequences. Such an approach provides perfect GT trajectory of the camera, however the environment is considerably less complex than in the realistic scenarios.

This paper presents the new environment setup for registration of data allowing the evaluation of the visual odometry and SLAM algorithms. The system allows to gather the following data:
- video sequences recorded by three cameras placed on the mobile robot,
- measurements from the onboard AHRS,
- video sequences from the five high resolution cameras observing the environment,
- video sequences from three static surveillance cameras placed in the environment,
- ground truth positions of all the cameras and their intrinsic parameters,
- ground truth trajectory of the mobile robot.

The additional surveillance cameras are used to simulate SLAM scenarios in which parts of the environment is observed by static cameras. Moreover, artificial markers were placed in the environments to facilitate research on augmenting the visual SLAM by enhancing the environment.

The paper is organized in the following fashion. The section 2 presents the environment setup and the calibration procedures, the section 3 shows the registration and the calculation of the GT trajectory, final conclusions and the future work are presented in the section 4.

The data registered with the system will be made freely available to facilitate research on the visual odometry and SLAM systems.

2. Environment Setup and Calibration

2.1. Robot

The WiFiBOT LAB V3 [9] robot (Figure 3) is used in the registration system. The robot is equipped with a Logitech Sphere high resolution pan-tilt camera, two Basler ACE aca640-100gc cameras with 3.5 mm lens for high speed (100 fps) video registration at the 640×480 resolution, an XSENS MTi 10 attitude and heading reference system (AHRS) and two Hall encoders. An asymmetric, circular grid was attached to the robot allowing to track the robot's position using the overhead cameras. The onboard cameras were calibrated using the standard calibration functions distributed with the OpenCV library [11].

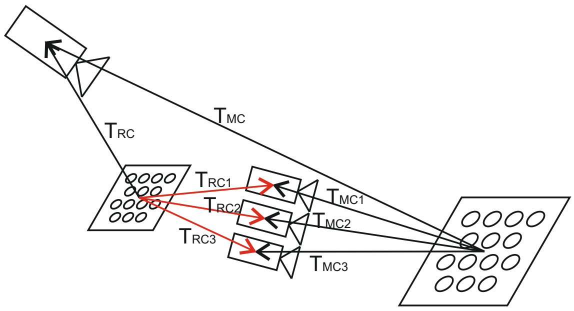

Fig. 1. *The robot cameras calibration*

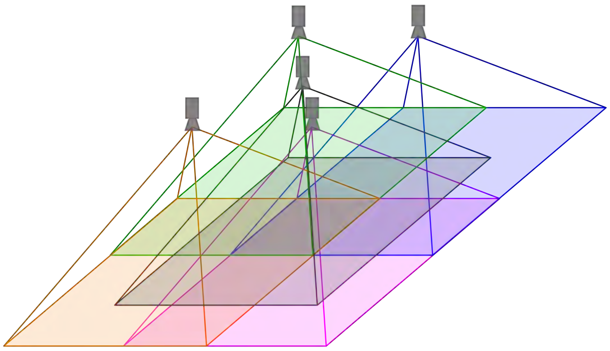

Fig. 2. *The arrangement of the cameras array and their corresponding fields of view*

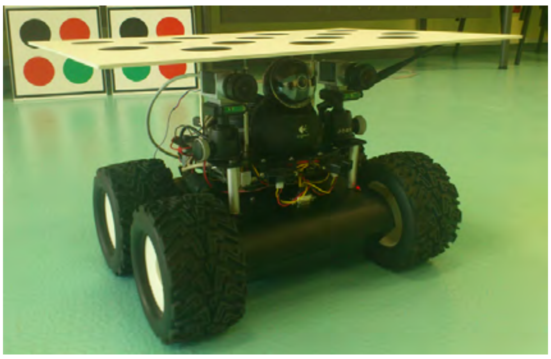

Fig. 3. *The WiFiBOT LAB V3 robot*

It is assumed that the robot's coordinate system is located in the geometrical center of the onboard marker. The AHRS system is attached under the the center of the marker. As the direct measurement of the cameras pose relative to the robot's frame would be cumbersome and unreliable an indirect, visual approach is used. A calibration marker is placed in front of the robot and an external camera is used to capture an image with the calibration marker and the onboard marker simultaneously visible. The relative pose of the camera with regard to the calibration marker (T_{MC}) and the onboard marker (T_{RC}) is calculated using the robust homography estimation algorithm. The pose of the onboard cameras relative to the calibration marker ($T_{MC1}, T_{MC2}, T_{MC3}$) is determined using the same method (Figure 1).

Afterwards, the cameras pose in the robot's frame is calculated according to:

$$T_{RCi} = T_{RC} {T_{MC}}^{-1} T_{MCi} \quad (1)$$

where i is the robot camera's number.

2.2. Tracking System

The movement tracking system consists of five Basler Ace aca1600-20gm cameras with low-distortion, aspherical 3.5 mm lens. The cameras are mounted under the ceiling of the laboratory, forming an array with partially overlapping fields of view (Figure 2). The cameras were calibrated using the functions available in the OpenCV library [11].

The tracking system's reference frame is assumed to be coincident with the coordinate system of the central camera (T_{CC}). The pose of the peripheral cameras relative to the tracking system's reference coordinates ($T_{PC1}, T_{PC2}, T_{PC3}, T_{PC4}$) was calculated by observing the calibration patterns placed on the laboratory floor in the overlapping part of the fields of view (Figure 4):

$$T_{PCi} = {T_{Mi1CC}}^{-1} T_{MPCi} \quad (2)$$

where i stands for the number of the peripheral camera.

The tracking system uses observations of the asymmetric, circular pattern attached on top of the robot. The marker is used to calculate the position of the robot in camera's coordinate systems and thus provides the ground truth trajectory of the robot. OpenCV library includes functions allowing for robust detection of predefined patterns in images. Moreover, the image coordinates of characteristic points of these patterns can be established with subpixel accuracy. The image coordinates of characteristic points on the observed marker can therefore be used as an input to the perspective-n-point solver [10], as the points are coplanar. An iterative, Levenberg-Marquardt method is used for optimization. The function tries to find a pose that minimizes the reprojection error, defined as the sum of squared distances between the observed projections of object points and a representation of object points tied to the camera coordinate system.

2.3. Additional Static Cameras

Three additional Basler Ace aca640-100gm cameras are installed on photographic tripods and placed in the environment. The main purpose of those cameras is to register additional video sequences covering parts of the robot's trajectory which would not be used to calculate the GT trajectory. Therefore, the sequences will not be correlated with the estimated GT trajectory of the robot and can be used to simulate the presence of the surveillance cameras in the robot's environment as well as to develop algorithms for static sensors discovery in SLAM systems.

The additional cameras were also calibrated using the OpenCV library. Similarly to the calibration of the tracking system, their position relative to the tracking system's reference frame ($T_{SC1}, T_{SC2}, T_{SC3}$) was calculated by registering images of the calibration mark-

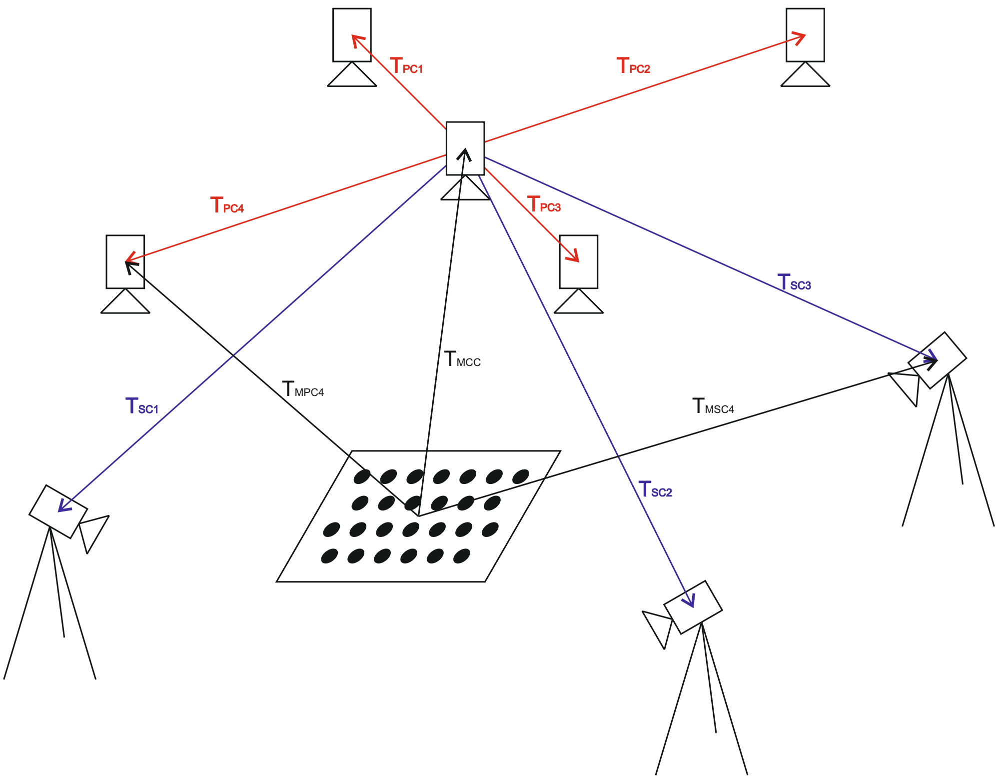

Fig. 4. The tracking system and static cameras calibration

ers visible by both the tracking system and the additional cameras (Figure 4):

$$T_{SCi} = T_{Mi1CC}{}^{-1} T_{MSCi} \tag{3}$$

where i stands for the number of the static camera.

2.4. Artificial Landmarks

Several artificial markers are placed in the environment to facilitate research on augmenting the environment for navigation purposes [12]. Each marker consists of four circles printed on a white square with 1 cm thick black border. One of the circles is black in order to provide the information on the marker orientation. The other three circles are used to color-code the markers allowing their recognition (Figure 6).

3. Registration Procedure

3.1. Synchronization

The eleven cameras used in the experiment register relatively large amount of data, therefore three independent computers are used to capture and save the images from the cameras. The first computer is capturing video streams from the robot's onboard cameras, while the other two are capturing the data from the overhead and surveillance cameras.

The images registered in such configuration require a timestamp in order to allow synchronization of the data streams and correct matching. The Precise Time Protocol (PTP) is used to synchronize the registering computers. PTP is a time transfer protocol defined in IEEE 1588-2002 (original version) and IEEE1588-2008 (PTPv2) standards [13]. It allows clock synchronization across a computer network. It offers very high accuracy level making it suitable for measurement and control systems. The Domain Time II application suite was used for the PTP server and clients. The synchronization system consists of the master computer serving as a PTP server and controlling the whole data registration process and three computers capturing images from the cameras (Figure 7).

4. Conclusions

The system for registration of mobile robot's trajectories in an indoor environment was designed, assembled and calibrated. The system allows precise tracking of the robot's trajectory as well as registration of the additional data such as the AHRS measurements or video sequences from static cameras. The system will be used to register benchmarking data for the evaluation of the visual SLAM and odometry systems. An exemplary snapshot from the system's cameras is presented on the Figure 5. The data registered with the presented system will be made publicly available with accordance to the Open Access research model.

In the future the benchmarking system will be extended to allow simultaneous tracking of multiple robots exploring a shared environment. Moreover, the system will be adapted to track the movement of a quadrocopter robot.

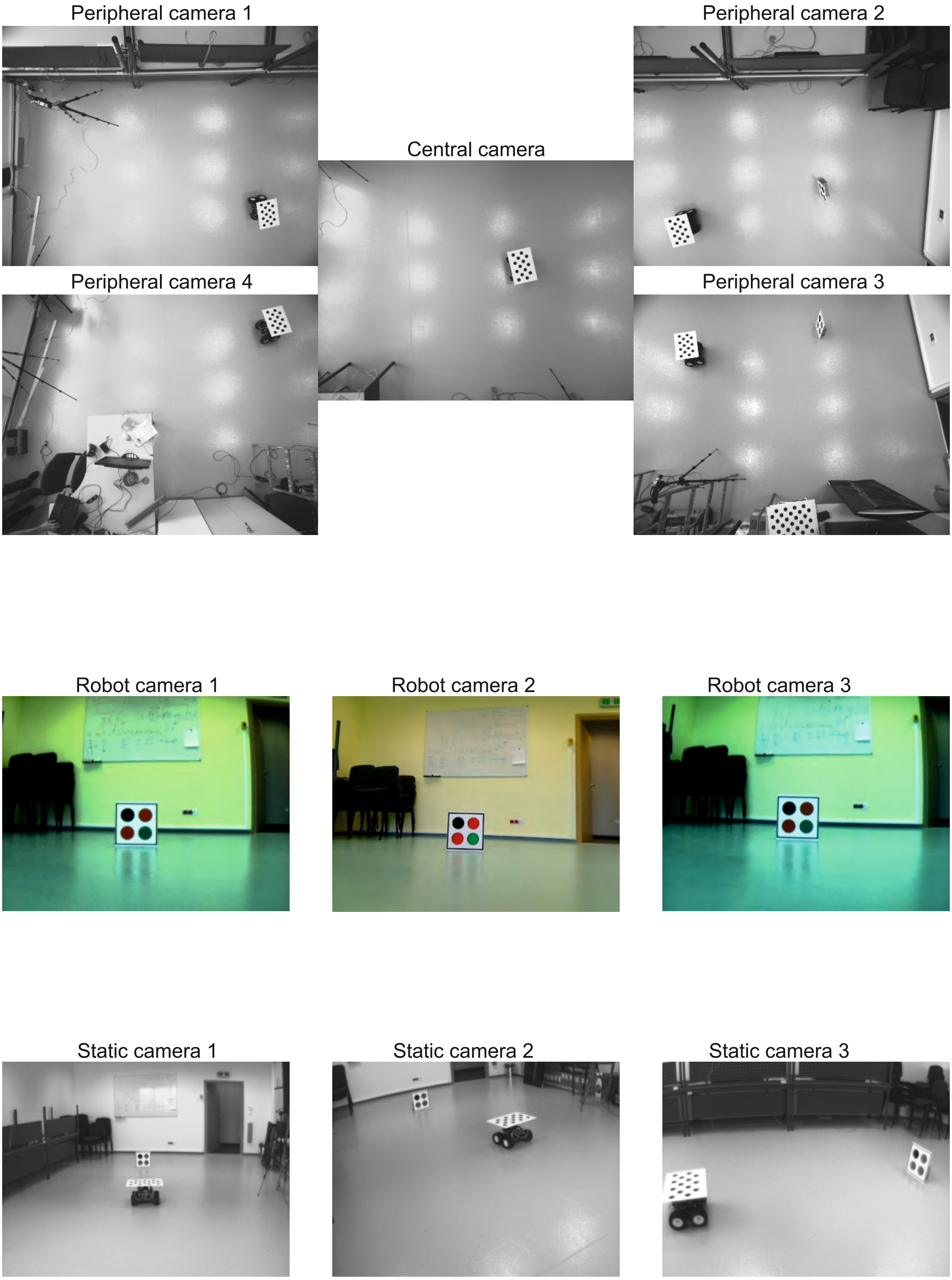

***Fig. 5.** Exemplary snapshot from the system cameras. Tracking system cameras - top, robot cameras - middle, static cameras - bottom*

Name	f_x	f_y	c_x	c_y	k_1	k_2	p_1	p_2	k_3	k_4	k_5	k_6
RC1	640.20	640.67	332.47	258.15	-0.112	0.210	0.002	-0.001	-1.810	0.288	0.011	-1.873
RC2	542.57	542.59	319.18	234.60	0.055	-0.584	0.004	0.001	-1.187	-0.051	-0.266	-1.436
RC3	651.57	652.25	355.39	226.25	-0.106	2.148	0.001	-0.001	-1.358	0.290	1.986	-0.714
CC	845.49	845.05	806.61	588.19	0.014	0.044	0.001	0.001	0.082	0.052	-0.058	0.132
PC1	827.56	827.60	803.00	588.28	1.555	0.102	0.001	-0.001	0.029	1.622	-0.016	0.071
PC2	835.75	836.15	791.47	609.57	1.022	0.224	0.001	0.001	-0.245	1.086	0.082	-0.180
PC3	829.48	830.40	799.16	606.89	-0.015	0.097	-0.001	-0.001	0.017	0.035	-0.023	0.079
PC4	829.14	828.61	812.27	632.35	-0.043	0.076	0.001	-0.001	0.070	0.011	-0.057	0.144
SC1	651.19	651.98	340.51	257.77	-0.188	1.113	0.001	-0.001	0.220	0.229	0.769	0.975
SC2	650.28	651.30	342.45	234.85	-0.179	1.373	-0.001	0.001	0.568	0.203	1.233	1.089
SC3	647.69	647.60	326.33	259.68	-0.185	-0.218	0.001	0.001	5.085	0.226	-0.720	6.399

Tab. 1. *Intrinsic cameras parameters and distortion coefficients*

Name	r. e. min	r. e. max	r. e. mean
RC1	0.051	0.350	0.162
RC2	0.046	0.403	0.138
RC3	0.052	0.302	0.158
CC	0.080	0.375	0.189
PC1	0.066	0.240	0.132
PC2	0.077	0.312	0.168
PC3	0.113	0.304	0.218
PC4	0.057	0.255	0.142
SC1	0.070	0.257	0.132
SC2	0.049	0.300	0.117
SC3	0.053	0.382	0.149

Tab. 2. *Minimum, maximum and mean reprojection error of the calibration procedure*

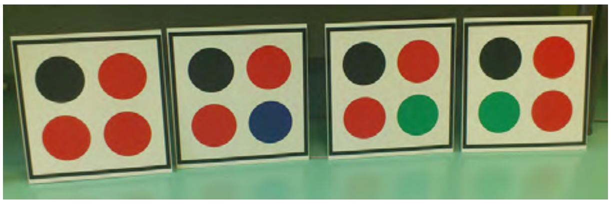

Fig. 6. *Exemplary artificial markers*

AUTHORS

Adam Schmidt* - Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland, e-mail: Adam.Schmidt@put.poznan.pl.

Marek Kraft - Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland, e-mail: Marek.Kraft@put.poznan.pl.

Michał Fularz - Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland, e-mail: Michal.Fularz@put.poznan.pl.

Zuzanna Domagała - Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland, e-mail: Zuzanna.Domagala@cie.put.poznan.pl.

*Corresponding author

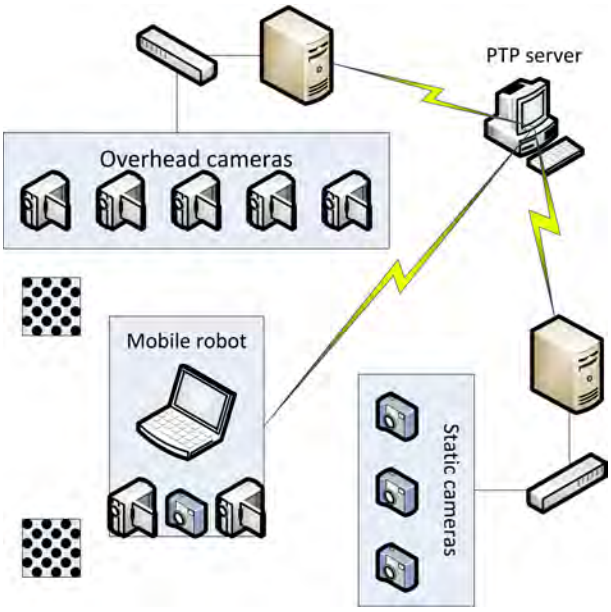

Fig. 7. *System configuration for precise time synchronization*

ACKNOWLEDGEMENTS

Adam Schmidt, Marek Kraft and Michał Fularz are scholarship holders within the project „Scholarship support for PH.D. students specializing in majors strategic for Wielkopolska's development", Sub-measure 8.2.2 Human Capital Operational Programme, co-financed by European Union under the European Social Fund.

This research was financed by the National Science Centre grant funded according to the decision DEC-2011/01/N/ST7/05940, which is gratefully acknowledged.

REFERENCES

[1] D. Scaramuzza, F. Fraundorfer, „Visual Odometry: Part I - The First 30 Years and Fundamentals", *IEEE Robotics and Automation Magazine*, vol. 18(4), 2011, pp. 80-92.

[2] F. Fraundorfer, D. Scaramuzza, „Visual Odometry: Part II - Matching, Robustness and Applica-

tions", *IEEE Robotics and Automation Magazine*, vol. 19(2), 2012, pp. 78-90.

[3] A. J. Davison, I. Reid, N. Molton, O. Stasse, „MonoSLAM: Real-Time Single Camera SLAM", *IEEE Trans. PAMI*, vol. 29(6), 2007, pp. 1052-1067.

[4] A. Schmidt, A. Kasińsksi, „The Visual SLAM System for a Hexapod Robot", *Lecture Notes in Computer Science*, vol. 6375, 2010, pp. 260-267.

[5] S. Ceriani, G. Fontana, A. Giusti, D. Marzorati, M. Matteucci, D. Migliore, D. Rizzi, D. G. Sorrenti, P. Taddei, „RAWSEEDS ground truth collection systems for indoor self-localization and mapping", *Autonomous Robots Journal*, vol. 27(4), 2009, pp. 353-371.

[6] S. M. Abdallah, D. C. Asmar, J. S. Zelek, „Towards benchmarks for vision SLAM algorithms", In: *Proceedings 2006 IEEE International Conference on Robotics and Automation*, 2006, pp. 1542-1547.

[7] J. Sturm, N. Engelhard, F. Endres, W. Burgard, D. Cremer, „Towards a benchmark for RGB-D SLAM evaluation", In: *Proc. of the RGB-D Workshop on Advanced Reasoning with Depth Cameras at Robotics: Science and Systems Conf. (RSS)*, 2011.

[8] J. Funke, T. Pietzsch, „A Framework for Evaluating Visual SLAM", In: *Proc. British Machine Vision Conference*, 2009.

[9] http://www.wifibot.com

[10] R. M. Haralick, Ch.-N. Lee, K. Ottenberg, M. Noelle, „Review and analysis of solutions of the three point perspective pose estimation problem", *International Journal of Computer Vision*, vol. 13(3), 1994, pp. 331-356.

[11] http://opencv.org

[12] R. Baczyk, „Wizyjny system nawigacyjny oparty na zmodyfikowanej metodzie samolokalizacji i tworzenia modelu otoczenia robota", PhD Thesis, Poznan University of Technology, 2007.

[13] K. Lee, J. Eidson, „IEEE-1588 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", In: *34th Annual Precise Time and Time Interval (PTTI) Meeting*, 2002, pp. 98-105.