Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 69 (2014) 968 – 973

Procedia Engineering

www.elsevier.com/locate/procedia

24th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2013

Object Tracking with a Multiagent Robot System and a Stereo Vision Camera

Filip Šuligoj *, Bojan Šekoranja, Marko Švaco, Bojan Jerbić

University of Zagreb, Faculty of Mechanical Enigineering and Naval Arhitecture, Department of Robotics and Production System Automation,
Ivana Lučića 5, 10000 Zageb, Croatia

Abstract

When working with a robot in terms of object manipulation the essential information is relative position between robot's tool center point (TCP) and the object of interest. This paper proposes a method of frame relative displacement and describes a working multiagent robot application that can be used for tracking, tooling or handling operations with the use of stereo vision in unstructured laboratory environment. Robot system is composed of two Fanuc robot arms, one of which carries a stereo vision camera system and the other which is guided in relation to object of interest. The latter robot has a marker that is used for navigation between the robot and the object of interest. Image processing, marker detection, 3-D coordinates extraction, coordinate system transformations, offset coordinates calculation and communication are handled using c++ multithread program and TCP/IP protocol.

Keywords: object tracking; robotic manipulation; stereo vision; image processing; object recognition

1. Introduction

Robots in production, service, movie or other facets of industry are mostly faced with operations which include interaction with solid objects. To perform any kind of operation with the object, its position in relation to the robot must be known. Method for obtaining the required position between object of interest and robot's TCP, without the use of sensors, includes placing an object at the same position in reference to the robot base. Today, machine vision is often used for locating objects and extracting position coordinates. Determining object's position in a 2-D image

---

* Corresponding author. Tel.: +385981614645; fax: +0-000-000-0000 .
E-mail address: filip.suligoj@fsb.hr

doi:10.1016/j.proeng.2014.03.077

with feature detection consists of using an input image of the object to recognize its features in the scene image, as seen in [1, 2, 3]. Extracted features are then compared with the features from scene image by using mathematical methods, such as Random Sample Consensus (RANSAC) [4], to obtain affine projection matrix which best fits. Objects orientation is given in terms of affine projection. Another approach for determining object's position and orientation is model-based object recognition which matches model of an object with the 3-D data of the scene [5, 6]. The prerequisite for model based object recognition is a device which generates range images. Stereo vision system extracts 3-D data of the scene by comparing images taken from two vantage points. Position of the robot in the coordinate system in which the object was found gives the final needed information. With service robotics undergoing intensive development in recent years, the main task for mobile robots is manipulation with household objects in cluttered environments. This presents new challenges in object detection and grasping point identification. State of the art object recognition techniques use both range and 2-D image data [7, 8, 9, 10].

System described in this paper uses stereo vision based system to extract 3-D coordinates of three points in space from marker on a robot tool and from a marker on the object. This makes the setup very flexible because no predetermined robot position is necessary. The only condition for the application to work is that the stereo vision is pointed towards the two markers and that the marker on the robot is calibrated as TCP. When using stationary cameras the limitations of workspace size come from disparity range of stereo vision system. Disparity range limits the distance from which the camera is able to extract position coordinates. Software used with presented stereo vision system has a disparity range limit of 240, meaning that it can produce range data from distance of 0.45 to 1 meter with the stereo resolution setting of 1080x768 and a 240 mm baseline. Using multiagent collaboration models [11, 12], workspace depth and accuracy are increased by including additional robot to carry the stereo vision system. Additional benefit of using multiagent concepts is demonstrated in [13], enabling motion optimization of two six DOF robots working together in a shared environment.

2. System overview

Main components of system described in this paper are: two Fanuc LR Mate 200iC 5L robot arms, Bumblebee XB3 stereovision camera system and a personal computer (PC), shown in Fig. 1. Communication between robots and a PC is handled using socket communication and developed client-server programs.

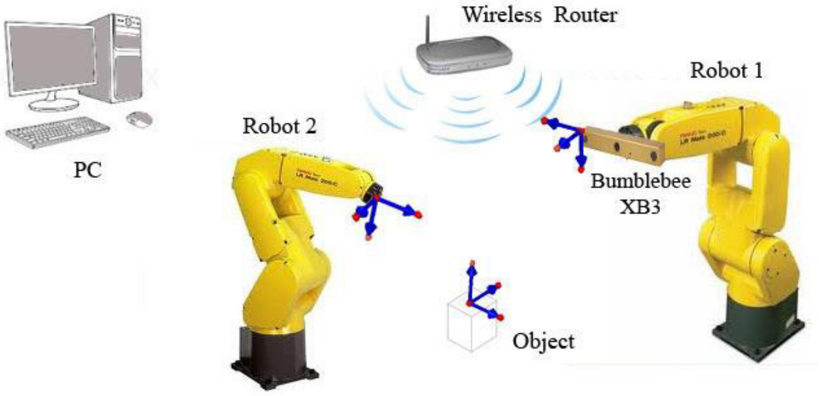

Fig. 1. Main components.

Multi-thread program in c++ uses the following software packages:

- Flycapture and Triclops from Point grey research company to manage image capture and 3D coordinates extraction from Bumblebee XB3
- Open source computer vision library (OpenCV) for image processing

- Eigen open source template library for linear algebra

Program run on PC initiates the stereo vision system and captures a 2-D image of the scene as well as the disparity image. Scene image is processed with the purpose of finding markers by using 2 different methods which will be described in detail later. Pixel coordinates of markers determined in a 2-D image are used to extract 3-D coordinates from disparity images. Extracted 3-D points are checked based on the previous knowledge of distance between points of each marker. Robot programs are started as well and they are running in a loop while waiting a message from the server. Robot 1, which carries the stereo vision system, follows the marker positioned on the second robot's tool at a fixed distance without changing its orientation and therefore only needs *x, y* and *z* coordinates. Robot 2 changes position and orientation based on relative position between the two coordinate systems defined with three points on each marker. Distance between coordinate systems and angles between its axes are then used to set a new TCP on a robot or to determine the offset between the robots TCP and the object. Choice between the two can be set in a robot program or from user input. When sent and received the data is written in the robot's registers and read when needed.

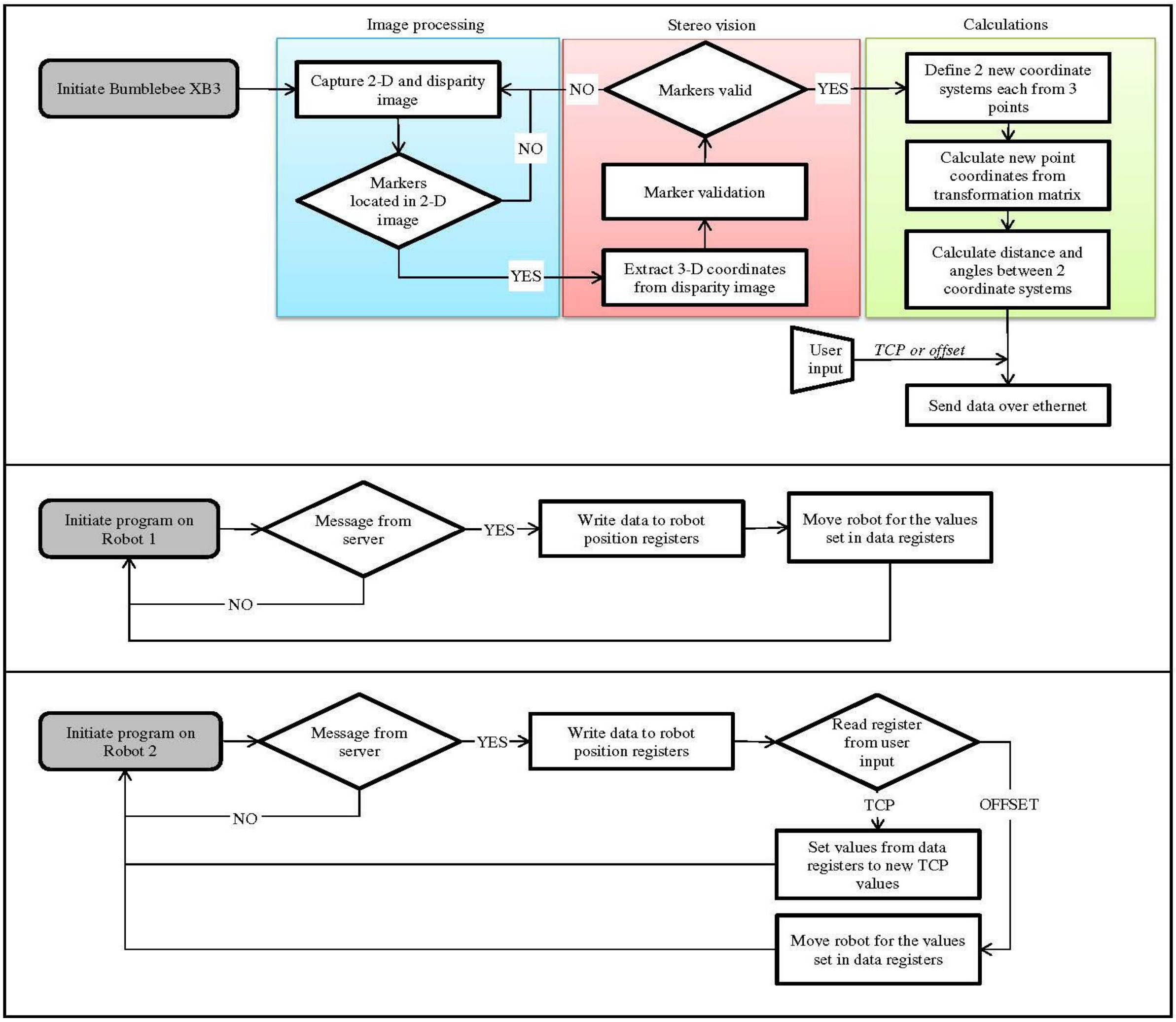

Fig. 2. System flowchart.

System flowchart shown in Fig. 2. presents a developed application which can run independently or be called from within any robot program. The relative position can be learned by guiding a robot, choosing to memorize the position or to manually program the wanted offset on a robot. When the robot needs to be aligned towards the object or moved around it in a specific way, a relative displacement method can be applied. Current distance and orientation of tool's marker coordinate system in object's marker coordinate system is memorized as a robot point. If the object is moved, the robot's user frame is placed at the object's new position and the robot is send to the memorised point in that coordinate frame.

2.1. Image processing with OpenCv

Image processing with the main goal of locating markers in 2-D image has been solved by using OpenCv computer vision library. Two different methods have been tested. First method is based on circle detection on the image with the use of feature extraction method: Hough circles. Each marker is composed of three red spheres (Fig. 3a) which project circles on the image regardless of the viewing direction. To eliminate many false readings and at the same time make lighting impact, the image is filtered based on the color (hue), saturation and value. First the color image is transformed from RGB (red, green, blue) to HSV (hue, saturation, value) color representation. With the use of range filters all colors in the image except red are removed. Hough circles algorithm then searches the image for all the circles and saves 2-D coordinates of their center in an array.

Second method uses two planar templates (Fig. 3b) whose features are extracted, with Speeded Up Robust Features algorithm (SURF), and matched with features on the greyscale scene image. Algorithm returns 2-D coordinates of four template corners in the scene image for every template, of which only three are used.

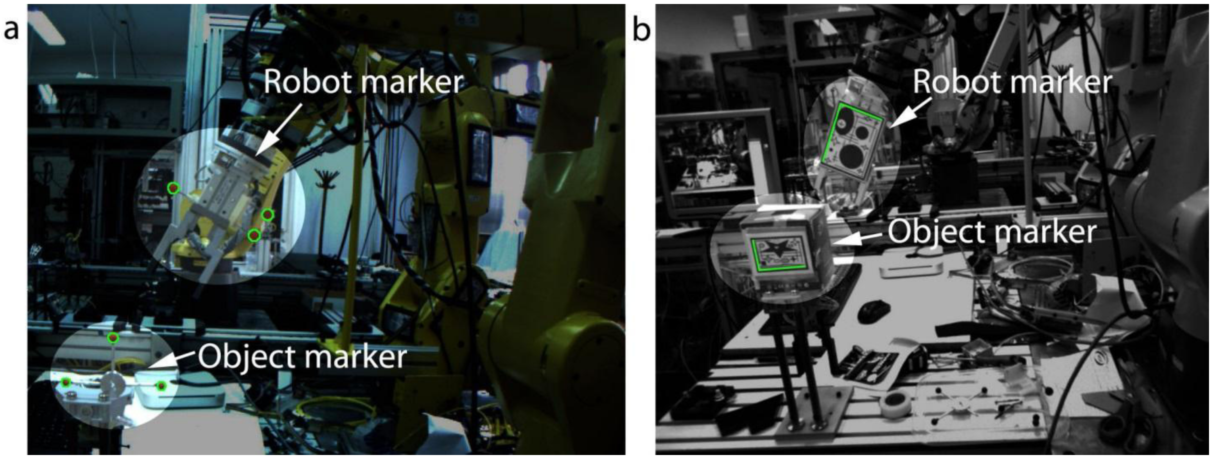

Fig. 3. (a) Hough circles; (b) SURF matching (markers are highlighted in green).

Both methods used are based on feature extraction from 2-D images and better results were obtained with Hough circles matching. Surf method is not able to find the marker correctly when it is strongly tilted away from the image capture plane. This can be contributed with the fact that sphere markers are used for Hough circles method, in comparison with planar template marker used with Surf whose position has to be mathematically approximated.

2.2. Distance and orientation calculations

First reduction of wrong readings from the stereo vision system is done with the use of validation methods for verifying correct matches between images. Validation methods are built into Triclops software package. After Triclops extracts 3-D marker coordinates, Euclidian distances between marker points are compared with their actual measured distance. In case that the values do not fit the requested precision range they are erased and the program is directed back to capture image phase.

Marker coordinates are received in the camera's coordinates system ($\mathbf{C_B}$), shown in Fig. 4. Using the three marker points on a robot tool the transformation matrix is calculated and all point coordinates are transformed to robot marker coordinate system ($\mathbf{C_R}$). Object's coordinate system ($\mathbf{C_O}$) is calculated from its three marker points. Distance between the two coordinate systems and three angles between their corresponding axes are used to set new user frame or to determine the difference in current and wanted position between the robot tool and the object.

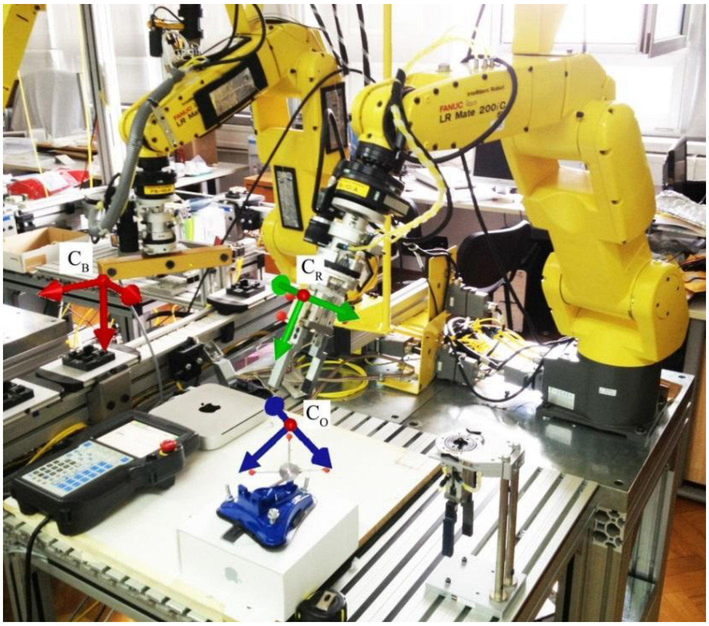

Fig. 4. Coordinate systems.

2.3. User frame displacement method and its application

Once the distance and angles between $\mathbf{C_O}$ and $\mathbf{C_R}$ are calculated, this information can be used to set new user frame. When setting user frame in such way all further movement of robot is done in $\mathbf{C_O}$, what can be very useful in many applications such as object handling, camera tracking or tooling.

In handling applications when the object needs to be moved in a specific way or through tight spaces and rotation or translation around one of its axes is needed, a new TCP can be set. Once the robot gripper has grasped the object, TCP can be set based on position of $\mathbf{C_O}$. For a simple pick and place program the wanted grip position needs to be saved. When the object is moved in workspace, the difference between its current position and the saved position is used as an offset to robots current position to bring the tool back to gripping position.

In tooling applications fast setup can be achieved if the tool itself is the object with marker. When the tooling point is known in relation with the marker, no tool calibration on the robot is needed. Such application can be very useful in robot multiagent systems. Multiple robots can exchange same tools while using different mounting points with no manual calibration needed.

Camera tracking can be used as a part of control process in the industry or for filming a certain object. With a camera mounted on a robot and aimed at the object of interest, current view can quickly be saved and regardless of object's movement in space it can follow it at the same distance and the same angle. Since robot's TCP is virtually set on the object, programming rotations of camera around object's axis while keeping the object in the center of the frame is simple. The same programmed point on the robot can be used and only the angle needs to be changed.

3. Conclusion and further work

Application presented in this paper utilises sensors, developed tracking markers and software in order to achieve greater robot autonomy when working in unstructured workspaces [14]. Robot system based on stereo vision guidance through 3-D space has been proven to work in laboratory environment. System has been tested with object tracking and pick and place applications. Two feature based marker detection approaches have been compared in terms of reliability and precision.

Future work will include improvements in systems precision and repeatability which are not at the industrial level. Such improvements could be achieved with model-based object recognition which would fully exploit the benefits of stereo vision 3-D data. Concerning multiagent collaboration, improvements regarding behaviour of Robot 1 could be introduced through intelligent repositioning in cases when the markers can't be found or when precision requirements are not satisfied. Future application of stereo vision system mounted on a robot can also include welding in terms of seam finding and automatic route or welding parameters corrections based on the welding joint type, size and location.

References

[1] H. Bay, A. Ess, T. Tuytelaars, Luc Van Gool, Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding Vol. 110, Issue 3 (2008), 346-359.

[2] David G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision Vol. 60, No. 4 (2004), 91-110.

[3] Luo Juan, Oubong Gwun, A comparison of SIFT, PCA-SIFT and SURF, International Journal of Image Processing Vol. 3, Issue 4 (2009), 143-152.

[4] M. A. Fischler, R. C. Bolles, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Communications of the ACM 24 (1981) .

[5] B. Drost, M. Ulrich, N. Navab, S. Ilic, Model globally, match locally: Efficient and robust 3D object recognition, The Twenty-Third IEEE conference on Computer vision and Pattern Recognition (CVPR), San Francisco (2010), 998-1005.

[6] Y. Sumi, Y. Kawai, T. Yoshimi, F. Tomita, 3D Object recognition in Cluttered Environments by Segmented-Based Stereo Vision, International Journal of Computer Vision Vol. 46, No. 1 (2002), 5-23.

[7] F. Rigual, A. Ramisa, G. Alenya, C. Torras, Object detection methods for robot grasping: Experimental assessment and tuning, CCIA 15 (2012), 123-132.

[8] A. Ramisa, D. Aldavert, S. Vasudevan, R. Toledo, R. Lopez de Mantaras, Evaluation of three vision based object perception methods for a mobile robot, Journal of Intelligent & Robotic Systems Vol. 68, Issue 2 (2012), 1–24.

[9] A. Collet, D. Berenson, S. S. Srinivasa, and D. Ferguson, Object recognition and full pose registration from a single image for robotic manipulation, International Conference on Robotics and Automation (ICRA) (2009), 48–55.

[10] D. Pangercic, V. Haltakov, and M. Beetz, Fast and robust object detection in household environments using vocabulary trees with sift descriptors, in International Conference on Intelligent Robots and Systems (IROS), Workshop on Active Semantic Perception and Object Search in the Real World (2011).

[11] M. Svaco, B. Sekoranja, B. Jerbic: A Multiagent Framework for Industrial Robotic Applications, Cihan H. Dagli (ed.), Procedia Computer Science, Vol 6, Elsevier B.V. (2011), 291-296.

[12] M. Svaco, B. Sekoranja and B. Jerbic (2011) Autonomous Planning Framework for Distributed Multiagent Robotic Systems, Proc. of The 2nd Doc. Conf. on Computing, Electrical and Industrial Syst., Caparica, Portugal, Luis M. Camarinha-Matos (Ed.), 147–154, Springer

[13] P. Ćurković, B. Jerbić, T. Stipančić: Coordination of Robots With Overlapping Workspaces Based on Motion Co-Evolution, International journal of simulation modelling, 12 (2013) , 1, 27-38.

[14]T. Stipancic, P. Curkovic, B. Jerbic, Robust autonomous assembly in environment with relatively high level of uncertainty, MITIP 2008 - 10th International Conference on The Modern Information Technology in the Innovation Processes of the Industrial Enterprises / J. Han, P. Holejsovska(ur.), Bily slon, s.r.o. Press, (2008), 193-198.