Multi-camera 3D Object Reconstruction for Industrial Automation

Malamati Bitzidou, Dimitrios Chrysostomou, Antonios Gasteratos

HAL Id: hal-01472287

https://hal.inria.fr/hal-01472287

Submitted on 20 Feb 2017

Multi-camera 3D object reconstruction for industrial automation

Malamati Bitzidou, Dimitrios Chrysostomou, and Antonios Gasteratos

Laboratory of Robotics and Automation,
Department of Production and Management Engineering,
Democritus University of Thrace,
PME Building, Vasilissis Sophias 12, GR-671 00 Xanthi, Greece
{malabitz,dchrisos,agaster}@pme.duth.gr
http://robotics.pme.duth.gr

Abstract. In this paper, a method to automate industrial manufacturing processes using an intelligent multi-camera system to assist a robotic arm on a production line is presented. The examined assembly procedure employs a volumetric method for the initial estimation of object's properties and an octree decomposition process to generate the path plans for the robotic arm. Initially, the object is captured by four cameras and its volumetric representation is produced. Thereafter, a quality check with its respective CAD model is performed and the final details of the 3D model are refined. An octree decomposition technique is utilized afterwards to facilitate the automatic generation of the assembly path plans and translate them to a sequence of movements for the robotic arm. The algorithm is fast, computationally simple and produces an assembly sequence that can be translated to any major robotic programming language. The proposed algorithm is assessed and preliminary experimental results are discussed.

Keywords: object assembly, multi-camera system, octree decomposition, 3D reconstruction, industrial automation

1 Introduction

The use of traditional manufacturing machinery such as robotic systems as a way of rapid manufacturing has attracted growing interest in recent years. Intelligent automation is one of the outcomes of this interest and its major impact enables a significant reduction in production time and cost, machining work and product quality. One of the manufacturing processes, in which intelligent automation is applied, is assembly, which constitutes an important stage in product development and accounts for a large proportion of the manufacturing costs. However, assembly still remains one of the least understood manufacturing processes [1]. In this paper an object assembly sequence is planned that incorporates the production of its volumetric model by a multi-camera system, its three-dimensional representation with octrees and its construction implemented by a robot arm of

5 degrees-of-freedom and a gripper. The final goal is to plan a robot arm path consisting of predetermined motions for the assembly of everyday objects.

The rest of this paper is organized as follows. Section 2 provides a review of the related works regarding multi-camera vision systems, automatic assembly planning processes and 3D object representation methods. In section 3 the overview of our system is described and the details about each one of the steps are discussed. Finally, the paper concludes in section 4 where the final remarks about the performance of the algorithm are given along with discussion for future research endeavors.

2   Related work

Technological advances in sensor design, communication, and computing are stimulating the development of new applications that will transform traditional visual systems into pervasive intelligent camera networks. Applications enabled by multi-camera networks are very common in industry. [2]. Systems utilizing multiple visual sensors are required in many manufacturing applications. Over the last few years the state of the art concerning multi-view stereo is improving rapidly. The goal of multi-view stereo is to reconstruct a complete three-dimensional object model from a collection of images taken from known camera viewpoints. A number of high-quality multi-view stereo algorithms have been developed, compared and evaluated [3]. In order to achieve multi-view stereo, systems consists of many visual sensors are usually needed. The uses of multi-camera systems in a variety of tasks premise the determination of optimal sensors configurations. Each system has its own set of requirements, but there is a generic formulation that can be customized to find optimum sensor configurations for all the systems [4].

Sensor planning has been researched quite extensively and several approaches have been developed in the last decades, so there are many different variations depending on the application. A popular set of methods, called next-view planning, attempt to built a model of the scene incrementally by successively sensing the unknown world from effective sensor configurations using the information acquired about the world up to this point [5], [6], [7], [8], [9], [10]. A related set of methods [11] have focused on finding the ideal sensor positions for capturing a static scene from desirable viewpoints assuming that some geometric information about the scene is available. Bordering on the field of graphics, the main contribution of such methods is to develop efficient methods for determining the view of the scene from different viewpoints. Methods that are directly related to ours are those that assume that complete geometric information is available and determine the location of static cameras so as to obtain the best views of an object. This problem was originally posed in the computational geometry literature as the art-gallery problem [12]. The traditional formulation of such problem assumes that two points are called visible if the straight line segment between them lies entirely inside the polygon. Even with such simple definition of visibility, the problem is NP-complete.

The continuous demand for reduced costs but increasing speed and quality in product manufacturing has led to the requirement for the rapid prototyping of solid models prior full-scale production. This is the reason why automatic assembly planning is essential. The series of technical issues addressed in the process of automated assembly planning are the assembly sequence representation at first, the generation and evaluation provided by the assembly assumptions, the calculation of the accuracy and efficiency of the planning process, the integration of the CAD program and finally the motion planning [13]. Numerous algorithms have been exploited for the rapid and effective production of feasible assembly tasks. The vast majority of assembly sequence generators, namely ASG, convert the problem of employing assembly sequences into disassembly ones. During the testing of their feasibility several issues must be considered such as geometrical and mechanical feasibility, manipulability, stability, visibility and used material.

An essential preprocess step of the assembly planning procedure is the three-dimensional representation of the object. The state of the art has changed dramatically over the last years, warranting a new overview in the field. Multiple approaches for assembly, utilize 3D geometry on a regularly sampled three-dimensional grid (volume), either as a discrete occupancy function [14] or as a function encoding distance to the closest surface [15]. The popularity of three-dimensional grids is increasing due to their simplicity and their ability to estimate various kinds of surfaces. A surface can be represented as a set of connected planar facets using polygon meshes, whilst in case of multiple views one can represent the examined scene as a set of depth maps. Among the multitude of existing methods, one can distinguish the use of octree approximation as an approach of compact visualization of computer-aided design models, which was recently reported [16].

3   System Overview

The proposed method lies within the category of intelligent assembly path planning methods. The main concept of the algorithm is to produce the volumetric model of the object, represent it using octrees and then create an assembly sequence of predetermined motions. The block diagram of the proposed algorithm is shown in Fig. 1 and the details for every step are given below.

3.1   Image Acquisition and Camera Calibration

The system's cameras topology is an application depended task and, therefore, it is based on the desired coverage of the working space taking into account the best combination of number of visual sensors and available bandwidth. In order to provide visual data for accurate 3D object assembly, a four-camera system was assembled inside our lab. The visual sensors are thus, mounted on the four corners of a cube processing $90°$ rotation along their optical axis, $60°$ rotation on the x axis and $150cm$ translation between them. Fig. 2 depicts the system's architecture and shows the position of the cameras inside the cubical structure.
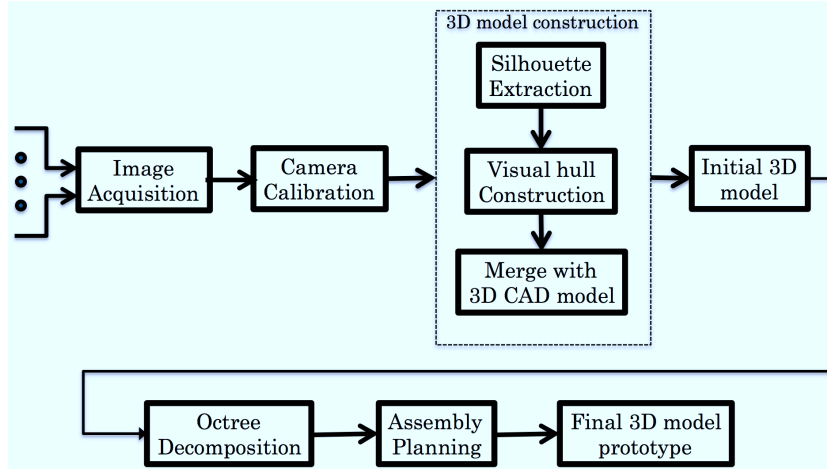
Fig. 1. Block diagram of the proposed method.

The first step of the whole process comprises an initialization phase where the calibration of the multi camera system occurs. The calibration procedure is fast, straightforward and results the registration of the cameras to the room coordinate system, employing a classic method [17] realized within the Camera Calibration Toolbox for Matlab [18]. The multi-camera rig acquires static snapshots of the object in a relatively controlled environment. The cameras viewpoints are carefully chosen in a way that will let us acquire the maximum possible shape information from a set of objects images.
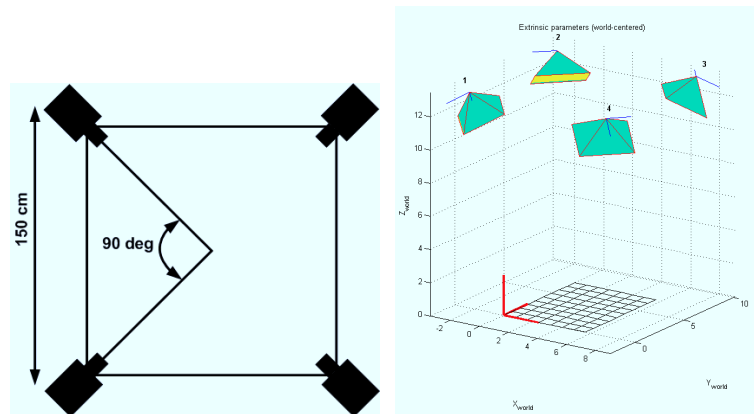

Fig. 2. The multi-camera system's architecture.

Hardware configuration. The proposed vision sensorial framework consists of a four *Flea*2 cameras, manufactured by Point Grey Research Inc [19] and are able to capture images of up to 1288×964 pixels resolution transmitted at a frame rate of 30Hz. They are manufactured using the industrial form factor of similar analog cameras and their size is exceptionally small. All cameras are equipped with $6mm$ lenses, which is equivalent to $34.25mm$ in the $35mm$ format, resulting in a working $61.36°$ angular field of view. All four cameras are connected through IEEE 1394b interface to a dedicated master PC, whilst the data transmission is established with the IEEE 1394b transfer protocol using two Point Grey Research PCI Express cards.

3.2   Initial 3D model construction

Following the camera calibration and image acquisition steps, the obtained image set is then processed in order to produce the first estimation of the 3D model of our examined object. Initially, each image is projected into an array of pixels and an accurate image thresholding approach is employed to result the first black-and-white silhouette of each source photograph as shown in Fig. 3. Having the introductory 2D approximation of the shape of the model consisted of pixels, a background subtraction algorithm isolates the object from its background information. Thereafter, the maximal shape that yields the same silhouettes as the actual object for all views, the so called Visual Hull [20] of the object, is approximated. The Visual Hull is the maximum possible shape that can be retrieved by the object's views and is consistent with its respective silhouettes as seen from any viewpoint in a given region. Once the visual hull of the test object is created it is then merged with its respective 3D CAD model to further refine any discrepancies appeared in the surface of the visual hull. The outcome of this algorithmic step is an initial approximation of the 3D model of the object that will be used as an input for the OcBlox system, analyzed in the next section.

3.3   Octree decomposition

In order to achieve the assembly of the model, a system heavily inspired by the OcBlox one, introduced by [16] is used. It also consists of the octree decom-
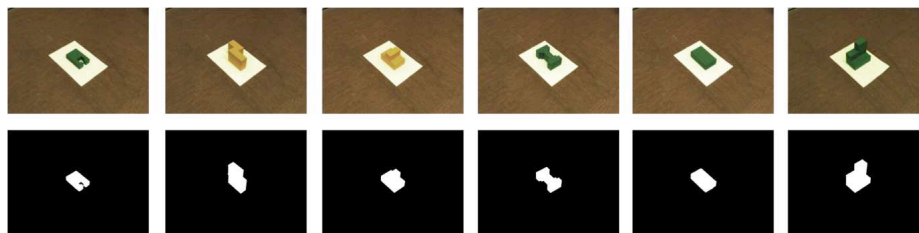

Fig. 3. The everyday objects used as an input for the assembly process and their respective black-and-white silhouettes.
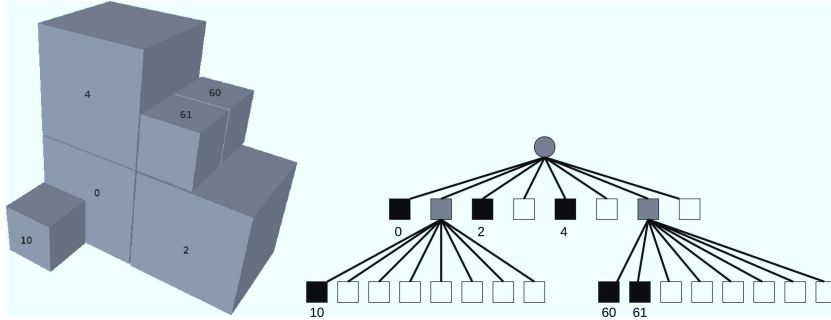
Fig. 4. Example of an octree representation.

position technique, the assembly planning and an assembly cell. The proposed system applies octree decomposition to the 3D visual model in order to convert it into approximate octree models. The octree decomposition technique results to hierarchical structure, that subdivides a volume into cubes of varying sizes, named octants. Among the several cubes a hierarchical tree relationship holds, whereat, each branch is associated to the relative position of the octant in the respective parent node. Octants are distinguished as full, empty or boundary, depending on their relative location in the 3D CAD model: inside, outside or partially inside, respectively. For maximum approximation, full and boundary octants are included in the rough model. The final octree is checked against different criteria, such as the maximum level of decomposition or the minimum size of octant. By concluding the decomposition process, the list of octants form an approximate representation of a 3D CAD model. In this system, octant sizes are limited to particular sizes according to the capability of the manufacturing process. An example of such an octree representation is depicted in Fig. 4.

3.4   Assembly planning

An assembly cell that is used to build cube-based models is developed as shown in Fig. 5 comprising:

1. The SCORBOT-ER 9 Pro (5(a)) industrial robot arm with great accuracy of nominal repeatability of $\pm 0.05 mm$, maximum operation speed: base rotation $140°/sec$, shoulder rotation $123°/sec$, elbow rotation $140°/sec$, wrist pitch $166°/sec$, wrist roll $300°/sec$ and a USB Controller-B with 16 I/O that can be read and generated via an Advanced Control Language (ACL).
2. A workspace where the configuration of the assembly cell is maintained and the cubes that constitute the prototype models are stored. (5(b)).
3. Raw material. Solid cubes of 10, 20 and 40 mm, made of wood with texture additions are used.
4. A camera with frame rate of 25 fps and resolution of 640x480 pixels is also used to recognize each wooden cube's coordinates and project them to the robot arm's coordinate system.
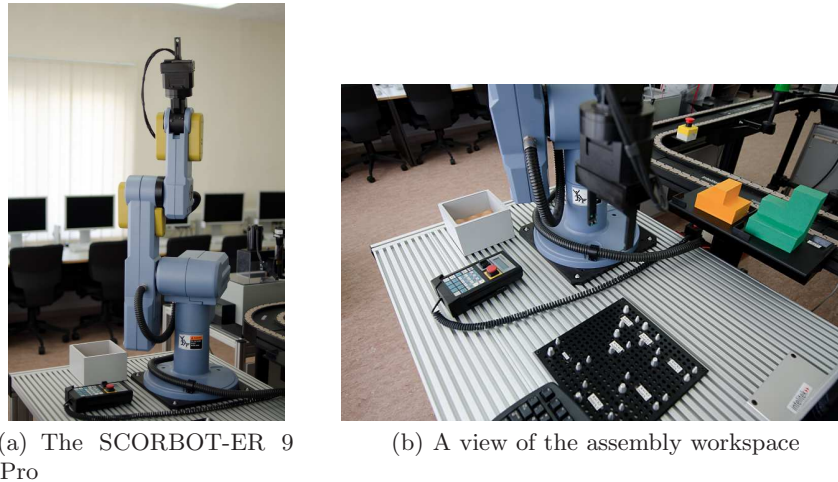
(a) The SCORBOT-ER 9 Pro

(b) A view of the assembly workspace

Fig. 5. The robot arm and the workspace used for the assembly cell.

The 3D data extracted from the previous step are the input for the assembly planning system. Thereafter, an estimation of the quantity and dimensions of the wooden cubes needed to assemble the real object is calculated. Their respective coordinates for their correct positioning inside the scene are then computed as well. The outcome of this process is the assembly instructions for the construction of the real prototype from the robotic arm.

4   Conclusion

In this paper, a method that performs object assembly using an intelligent vision network is presented. The proposed procedure is based on a volumetric approach for the objects reconstruction resulting to an initial 3D model. After the merging of this 3D model with its 3D CAD design, it can be represented using octree decomposition technique. The final 3D model prototype is constructed by a block of solid cubes positioned inside the assembly cell by the robotic arm. An assembly planning process that generates a sequence of predetermined motions specifies the exact positions of the cubes. The main contribution of the algorithm is the ability to reconstruct an ordinary object using only a system of four cameras. The algorithm is fast, computationally simple and the assembly sequence that is produced can be translated to any robotic programming language. Besides, the computational burden is directly linked to the complexity of objects geometry. The proposed technique is applicable to industrial smart manufacturing systems, to rapid prototyping cells and to standard robotic and machine vision applications.

References

1. Medellín, H., Corney, J., Ritchie, J., Lim, T.: Automatic generation of robot and manual assembly plans using octrees. Assembly Automation **30**(2) (2010) 173–183
2. Aghajan, H.K., Cavallaro, A.: Multi-camera networks: principles and applications. Academic Press (May 2009)
3. Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Volume 1., Ieee (2006) 519–528
4. Mittal, A.: Generalized multi-sensor planning. In: ECCV'06: Proceedings of the 9th European conference on Computer Vision, Springer-Verlag (2006) 522–535
5. Hager, G., Mintz, M.: Computational methods for task-directed sensor data fusion and sensor planning. The International journal of robotics research **10**(4) (1991) 285–313
6. Cameron, A., Durrant-Whyte, H.: A bayesian approach to optimal sensor placement. The International Journal of Robotics Research **9**(5) (1990) 70–88
7. Kutulakos, K., Dyer, C.: Recovering shape by purposive viewpoint adjustment. International Journal of Computer Vision **12**(2) (1994) 113–136
8. Maver, J., Bajcsy, R.: Occlusions as a guide for planning the next view. Pattern Analysis and Machine Intelligence, IEEE Transactions on **15**(5) (1993) 417–433
9. Ye, Y., Tsotsos, J.: Sensor planning for 3d object search. Computer Vision and Image Understanding **73**(2) (1999) 145–168
10. Miura, J., Ikeuchi, K.: Task-oriented generation of visual sensing strategies. In: Computer Vision, 1995. Proceedings., Fifth International Conference on, IEEE (1995) 1106–1113
11. Sing, B., Seitz, S., Sloan, P.P.: Visual tunnel analysis for visibility prediction and camera planning. In: Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on. Volume 2. (2000) 195 –202 vol.2
12. O'Rourke, J.: Art gallery theorems and algorithms. Volume 57. Oxford University Press Oxford (1987)
13. de Mello, L., Lee, S.: Computer-aided mechanical assembly planning. Volume 148. Springer (1991)
14. Slabaugh, G., Culbertson, W., Malzbender, T., Stevens, M., Schafer, R.: Methods for volumetric reconstruction of visual scenes. International Journal of Computer Vision **57**(3) (2004) 179–199
15. Jin, H., Soatto, S., Yezzi, A.: Multi-view stereo reconstruction of dense shape and complex appearance. International Journal of Computer Vision **63**(3) (2005) 175–189
16. Medellín, H., Corney, J., Davies, J., Lim, T., Ritchie, J.: An automated system for the assembly of octree models. Assembly Automation **24**(3) (2004) 297–312
17. Tsai, R.: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. Robotics and Automation, IEEE Journal of **3**(4) (1987) 323–344
18. Bouguet, J.Y.: Camera calibration toolbox for Matlab (2008)
19. Research, P.G.: Point grey research cameras
20. Laurentini, A.: The visual hull concept for silhouette-based image understanding. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (February 1994) 150–162