A Fast and Accurate Recognition System for Flexible Grasping of Electronic Goods

Ch. Woegerer[1*], M. Rooker[1], A. Angerer[1], Ch. Kopf[1], and A.Pichler[1]

[1]Robotics and Adaptive Systems
PROFACTOR GmbH
Steyr, 4407, Austria

Abstract

Automated packaging is becoming more and more interesting for various production sites even in Europe. However, in many companies the packaging process can only handle a limited amount of products. This leads to a very inflexible packaging and a bottle neck in the production. Most of the time for new products, the packaging system needs to be modified. One of the main components in an automated packaging process is the recognition and then the grasping of the objects that need to be packed. For different products mostly a different gripping system is required. In this paper, solutions from a research project are presented which enables the recognizing as a basis for flexible grasping of different products. These solutions includes a flexible pose recognition for different objects by a low cost system based on Kinect or or Asus Xtion cameras.

1. Introduction

Electronic consumer goods, e.g. dish washers, TV sets, toasters and microwaves have a large number of variants and are packaged manually at most production sites. Only in single-variant production lines with high lot sizes, an automation of the packaging step has been introduced. However, automating the packaging process will decrease the production cycle time (and thus costs) also for mixed variant production lines, thus allowing that several production lines can be merged to a reduced number of flexible packaging stations. This also allows an optimization with regard to the actual demands of the (various) goods (i.e. number of items produced per day).

A key system is the recognition system of the different goods. The object pose recognition relies on a 3D surface scan of the object to be grasped. Therefore, the software ReconstructMe is used to generate an accurate 3D surface model of the object to be grasped. The ReconstructMe system consists of a depth image input device, such as a Microsoft Kinect or Asus Xtion, a computer for the calculation of the 3D surface and visual feedback to the user. For 3D surface reconstruction, the user takes the depth image capture device in his hand and films the object from different viewpoints. The 3D surface model is captured in real-time, at approximately 30 frames per seconds (fps). The real-time capability is achieved through calculation on the GPU, which provides data parallel execution.

The paper presents the theoretical background of the reconstruction system, the development steps and also the results of different trails within the European funded project called CustomPacker.

2. Technologies

Within the CustomPacker project [1], the goal is to develop a flexible packaging system. The system consists out of different components that will optimize the packaging process. In the upcoming sections, three different components of the proposed solution will be presented, which will cover the topics of object pose recognition for detecting where the object to be grasped is located and in which pose, the collision detector for detecting possible collisions between the planned path of the robot system and the work cell and finally, the flexible gripper that is developed within the project.

* Corresponding author: Tel.: (+43) 664 6207675, E-mail: christian.woegerer@profactor.at

2.1. Object Pose Recognition

The object pose recognition relies on a 3D surface scan of the object to be grasped. Therefore, the software ReconstructMe [12] is used to generate an accurate 3D surface model of the object to be grasped. The ReconstructMe system consists of a depth image input device, such as a Microsoft Kinect or Asus Xtion, a computer for the calculation of the 3D surface and visual feedback to the user. For 3D surface reconstruction, the user takes the depth image capture device in his hand and films the object from different viewpoints. The 3D surface model is captured in real-time, at approximately 30 frames per seconds (fps). The real-time capability is achieved through calculation on the GPU, which provides data parallel execution.

The theoretical background of ReconstructMe is to construct a copy of the surface via a truncated signed distance function (TDSF), represented by a union volume grid with a configurable size in all dimension extensions (x, y, z) and grid size. Thus, the 3D reconstruction is limited to the size of the volume, and the accuracy can depend on the grid size. Each depth image frame of the input device is integrated via the TDSF into the volume. Since the camera can move around freely, the position of the camera has to be tracked relative to the volume. Therefore, the previous depth images and the current one are transformed to point clouds in the camera coordinate system and aligned by a high speed version of the Iterative Closest Point (ICP) algorithm. Once the reconstruction is finished, a triangulated model can be exported by the Marching Cube algorithm [13].

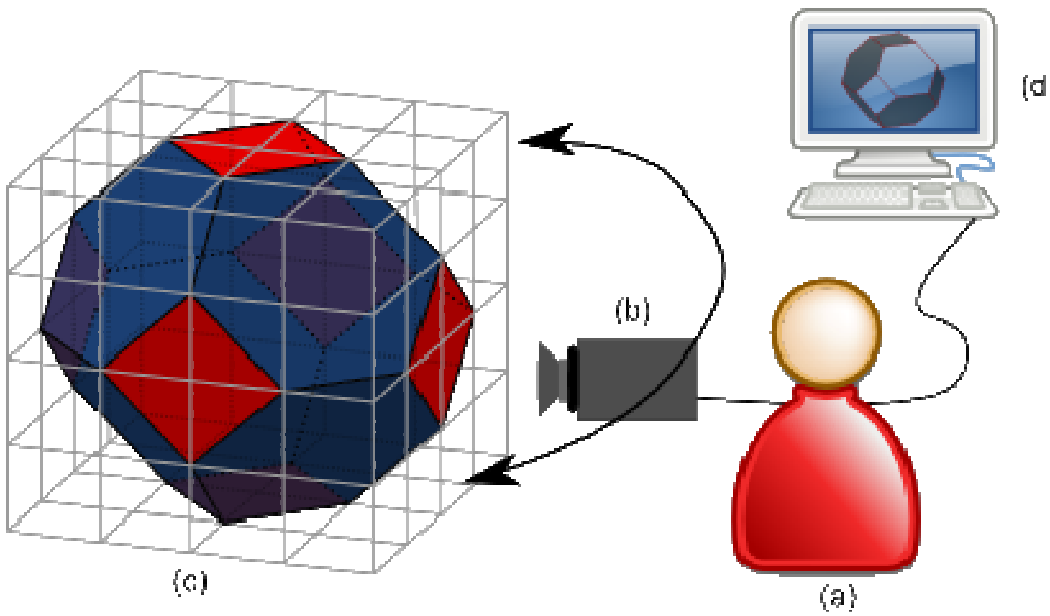

Fig. 1. System components. (a) User, (b) Sensor, (c) Global model, (d) Computer / real-time visualization.

The sensor system defines the global model in the world coordinate system and the global model is defined within a discrete volume. Since the volume is axis aligned to the world coordinate system, it can be defined by two 3D points. These points are the minimum and the maximum corner, both described in the world coordinate system (Fig. 2). Assuming a setup with one sensor, the sensor moves within the world coordinate system and the sensor position has to be tracked.

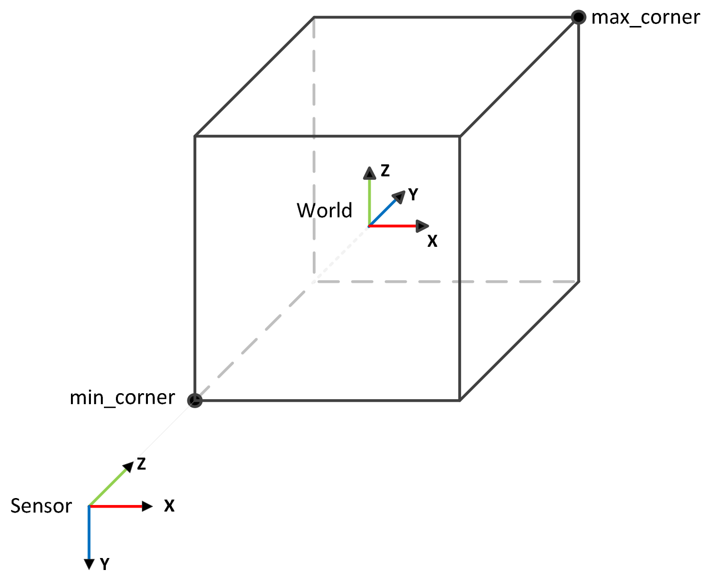

Fig. 2. Coordinate Systems.

Fig. 2 shows the initial sensor position in the world coordinate system. It is aligned and the z-axis points across the y-axis of the world coordinate system and is positioned 400 millimeters in front of the volume. Thus the initial sensor position $t_{initial}$ is a translation along the y-axis, depending on the volume definition. When the user starts the scanning process, the world coordinate system is fixed and the sensor now has to estimate its position within the world coordinate system. The sensor tracking is based on an efficient variant of the ICP algorithm [14]. For the ICP registration i, the last sensor point cloud $\mathbb{F}_{i-1}$ described in the sensor coordinate system and the current sensor point cloud $\mathbb{F}_i$ defined in the sensor coordinate system as well gets aligned and a rigid transformation t_i can be calculated. Thus, the sensor position in iteration i can be defined as

$$T = (\prod_{i}^{1} t_i) t_{initial} \qquad (1)$$

This kind of tracking is defined as local tracking. See Fig. 3 for a recorded sensor path.

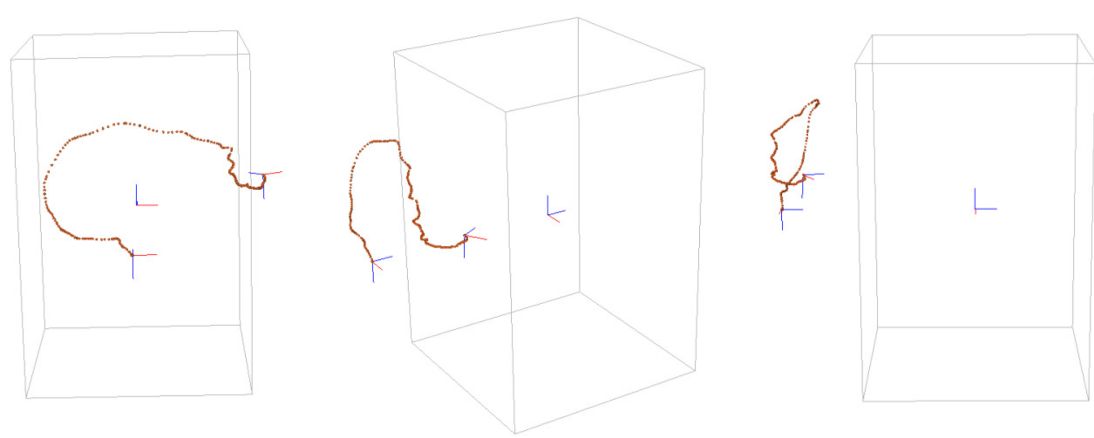

Fig. 3. A recorded sensor path from different views, tracked by the ICP algorithm.

This approach uses an optimized ICP algorithm which assumes small sensor movement between the iterations. If the movement is too big, the ICP algorithm produces erroneous results and the sensor tracking is lost. The system is able to detect such tracking failures. In case of tracking failure detection, global tracking comes into account. Fig. 4 shows a point cloud and the reconstruction of the scanned TV set.

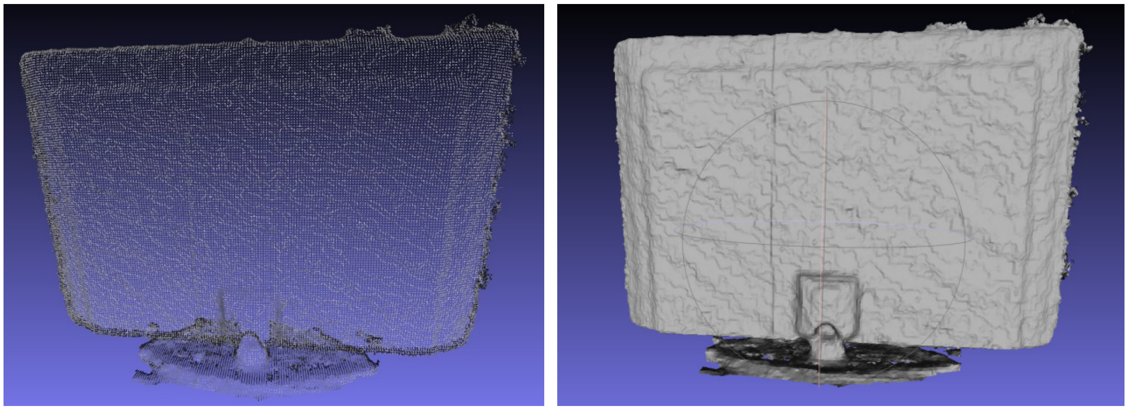

Fig. 4. Scan of a TV set. Left: Point Cloud, Right: Reconstruction.

The point cloud recorded with the ReconstructMe software is used as input for the 3D object pose recognition.

For identifying new objects a CAD file or a ReconstructMe scan of the object is needed. Furthermore, it is important as the grip points for each object needs to be defined via the available CAD models or the in advanced reconstructed objects. All these offline prepared objects are stored into a database, which will then contain a collection of all available objects that need to be grasped (right side of Fig. 5).

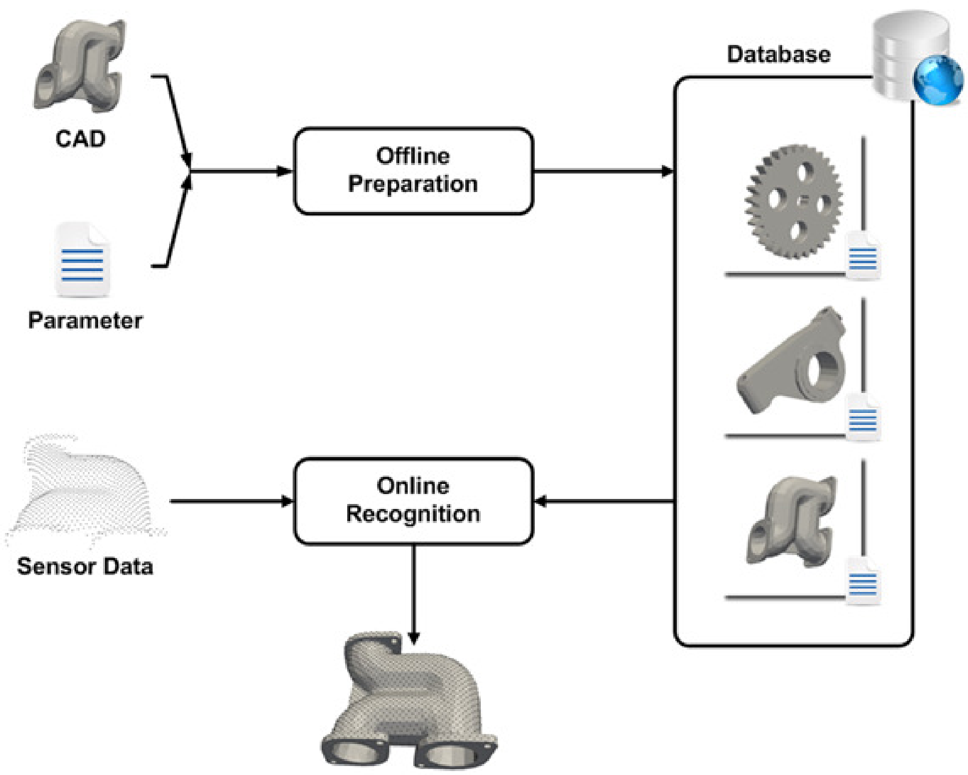

Fig. 5. Preparation of objects for online object and pose recognition.

For the actual 3D object pose recognition, a randomized global object localization algorithm (RANGO) is used, which is based on random sampling algorithm (RANSAC) [15]. The implementation is currently limited to finding one or more instances of one object in a scene. The large advantage is a very fast detection rate of objects when e.g.

compared to [16]. The main contribution is the replacement of K-nearest neighborhood search for inlier detection with a probabilistic grid based approach. This has as a result that the time complexity for the evaluation of a hypothesis (acceptance function) is reduced from $O(n * \log(m))$ where n is the number of model points and m the number of points in the scene, too $O(n)$. Additionally, the evaluation of the number of model points that fit the hypothesis is stopped early as soon as the probability of finding a good match is lower than a given threshold.

The final recognition of the television set can be seen in Fig. 6. The blue parts are the best identified parts, whereas the red parts and the unidentified parts.

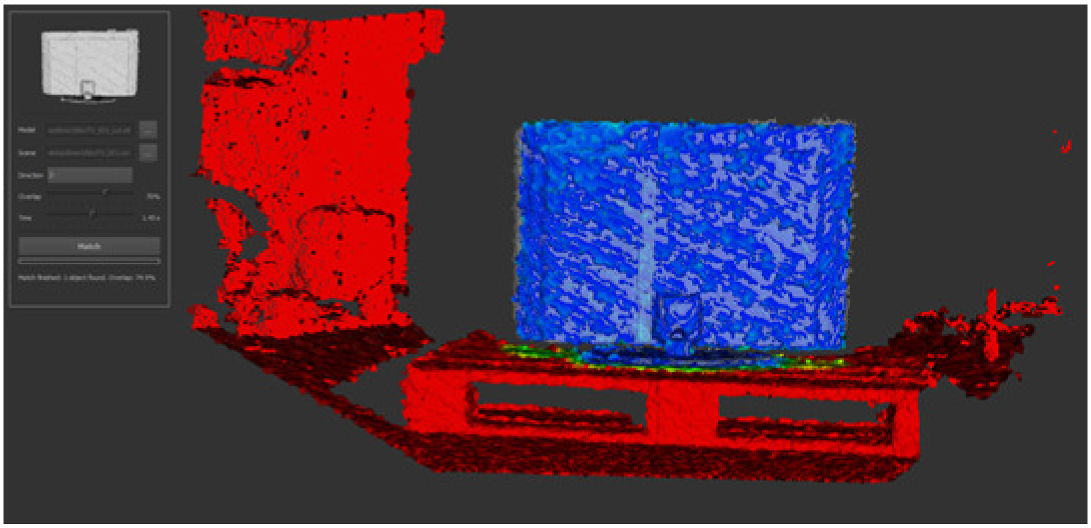

Fig. 6. Recognition of TV Set.

3. Future Work

The CustomPacker EU research project has finished in 2013 including the work on the individual assistive technologies for flexible grasping of electronic consumer goods. The developments of the technologies are now in such a stable state that they are ready for industrializing. Within the final phase of the project, a market study has been performed and a business plan has been defined on how to bring the technologies to the market.

The results of the system are not only relevant for manufacturers of electronic consumer goods (especially for heavy and/or fragile products, which are up-to-now packaged manually), but also for other industry sectors and scenarios. Even at some time within the automotive industries, there are a large number of variants that are packaged manually at most production sites.

The results of the business plan clearly defines that a company needs to be identified that is willing to take the risk of exploiting the overall results of the CustomPacker project, which is a flexible packaging cell. The individual components (including the flexible gripper, object reconstruction and detection and the collision check algorithms) are either exploited by the individual partners or an external company is required that is interested and capable of exploiting these results.

4. Conclusions

The packaging industry is looking more and more at approaches for optimizing the packaging process and providing humans with ergonomic enhancements. Within this paper different assistant technologies are presented with the goal to create a highly customizable and flexible packaging station. A cheap but effective object pose recognition solution is introduced where the object to be grasped is digitally reconstructed and detecting using low-cost sensors (e.g. the Microsoft Kinect or Asus Xtion). With regards to the developed gripper, it is shown that it performs appropriate with very different sized parts, and that it can handle strong and heavy parts as well as delicate and light weight ones. The adaptability is provided by a hybrid mechanical solution comprising of three linear stages and compliant pads supported by a double mode control strategy consisting of force mode and position mode. Finally, an improved approach for collision detection of autonomous robot trajectories is introduced.

The individual results are all merged in a final prototypical packaging demonstrator, which is now in a state that it can be industrialized for employment into industrial companies.

ACKNOWLEDGEMENTS

The work was supported by the CustomPacker FP7 EU Project – "Highly Customizable and Flexible Packaging Station for mid- to upper sized Electronic Consumer Goods using Industrial Robots", http://www.custompacker.eu

REFERENCES

[1] *CustomPacker FP7 EU Project – "Highly Customizable and Flexible Packaging Station for mid- to upper sized Electronic Consumer Goods using Industrial Robots", http://www.custompacker.eu.*

[2] *World Robotics 2011 Industrial Robots, ISBN 978-3-8163-0635-1. Published by the IFR Statistical Department, hosted by VDMA Robotics + Automation, Germany.*

[3] *H. Brantmark, E. Hemmingson, "FlexPickerTM with PickMasterTM revolutionizes picking operations", ABB Review, 1/2001, 2001.*

[4] *A. Jordt, A.R. Fugl, M. Willatzen, R. Koch, H.G. Petersen, K.A. Andersen, M.M. Olsen and N. Krüger, „An Outline for an Intelligent System performing Peg-in-Hole Actions with flexible Objects", The International Conference on Intelligent Robotics and Applications (ICIRA), 2011.*

[5] *A. Pichler, M. Ankerl, "User Centred Framework for Intuitive Robot Programming", International Workshop on Robotics and Sensors (ROSE) 2010, October 15-16, 2010, Phoenix, Arizona, USA.*

[6] *A. Pichler, C. Wögerer, "Towards Robot Systems for Small Batch Manufacturing", IEEE International Symposium on Assembly and Manufacturing (ISAM 2011), Tampere, Finland, 25-27 May 2011.*

[7] *FESTO Didactic – The Bionic Tripod. [Online]. http://www.festo.com/cms/en_corp/9653.htm*

[8] *E. Brown, N. Rodenberg, J. Amend, A. Mozeika, E. Steltz, M. Zakin, H. Lipson, H. Jaeger, „Universal robot gripper based on the jamming of granular material", Proceedings of the National Academy of Sciences (PNAS), Vol. 107, no. 44, pp. 18809–18814.*

[9] *J.R. Amend, E. Brown, M. Rodenberg, H. Jaeger, H. Lipson, „A Positive Pressure Universal Gripper Based on the Jamming of Granular Material", IEEE Transaction of Robotics, Vol. 28, pp. 341–350, April 2012.*

[10] *Robotiq Adaptive Robot Gripper. [Online]. http://www.robotiq.com*

[11] *Schunk Dextrous Hand (SDH). [Online]. http://www.schunk.com*

[12] *C. Heindl and C. Kopf, "ReconstructMe - Digitize Your World": http://www.reconstructme.net/*

[13] *W.E. Lorensen, and H.E. Cline, "Marching Cubes: A high resolution 3D surface construction algorithm", In Proceedings of the 14th annual conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96), pp. 303–312, 1987.*

[14] *S. Rusinkiewicsz and M. Levoy, "Efficient Variants of the ICP Algorithm", in: Proceedings of the Third Intl. Conf. on 3D Digital Imaging and Modeling, pp. 145–152, 2001.*

[15] *C. Papazov and D. Burschka, "An efficient RANSAC for 3D object recognition in noisy and occluded scenes", Computer Vision-ACCV 2010, pp. 135–148, 2011.*

[16] *S. Winkelbach, S. Molkenstruck and F. Wahl, „Low-cost laser range scanner and fast surface registration approach", Pattern Recognition, pp. 718–728, 2006.*

[17] *Ch. Woegerer, A. Pichler, Martjin Rooker, A. Angerer, J. Capco, Ch. Heindl, A Olarra and E. Fuentes - "Flexible Grasping of Electronic Consumer Goods". Robotics in Smart Manufacturing P. Neto and A.P. Moreira (Eds.): WRSM 2013, CCIS 371, pp. 158–169, 2013. © Springer-Verlag Berlin Heidelberg 2013, ISBN 978-3-642-39222-1.*

[18] *Ch. Wögerer, M. Rooker, A. Pichler, A. Angerer, F. Wallhoff, J. Blume, A. Bannatt, P. Ferrara, A. Olarra J. Kiirikki: A Robotic Assistance System for flexible Packaging Consumer Goods using in Electronic Industry, RAAD2013, 22nd International Workshop on Robotics in Alpe-Adria-Danube Region Portorož, Slovenia, September 11–13, 2013.*