

A Visual Servoing System for Interactive Human-Robot Object Transfer

Ying Wang, Daniel Ewert, Rene Vossen, and Sabina Jeschke

Institute Cluster IMA/ZLW & IfU, RWTH Aachen University, Aachen, Germany

Email: {ying.wang, daniel.ewert, rene.vossen, sabina.jeschke}@ima-zlw-ifu.rwth-aachen.de

Abstract—As the demand for close cooperation between human and robots grows, robot manufacturers develop new lightweight robots, which allow for direct human-robot interaction without endangering the human worker. However, enabling direct and intuitive interaction between robots and human workers is still challenging in many aspects, due to the nondeterministic nature of human behavior. This work focuses on the main problems of interactive object transfer between a human worker and an industrial robot: the recognition of the object with partial occlusion by barriers including the hand to the human worker, the evaluation of object grasping affordance, and coping with inaccessible grasping points. The proposed visual servoing system integrates different vision modules where each module encapsulates a number of visual algorithms responsible for visual servoing control in human-robot collaboration. The goal is to extract high-level information of a visual event from a dynamic scene for recognition and manipulation. The system consists of several modules as sensor fusion, calibration, visualization, pose estimation, object tracking, classification, grasping planning and feedback processing. The general architecture and main approaches are presented as well as the future developments planned.

Index Terms—visual servoing, human-robot interaction, object grasping, visual occlusion

I. INTRODUCTION

Robots are a crucial part of nowadays industrial production with applications including e.g. sorting, manufacturing as well as quality control. The afflicted processes gain efficiency owing to the working speed and durability of robotic systems, whereas product quality is increased by the exactness and repeatability of robotic actions. However, current industrial robots lack the capability to quickly adapt to new tasks or improvise when facing unforeseen situations, but must be programmed and equipped for each new task with considerable expenditure. Human workers, on the other hand, quickly adapt to new tasks and can deal with uncertainties due to their advanced situational awareness and dexterity.

Current production faces a trend towards shorter product life cycles and a rising demand for individualized and variant-rich products. To be able to produce small

batches efficiently, it is desirable to combine the advantages of human adaptability with robotic exactness and efficiency. Such close cooperation has not yet been possible because of the high risk of endangerment caused by conventional industrial robots. In consequence, robot and human work areas had strictly been separated and fenced off. To enable a closer cooperation, robot manufacturers now develop lightweight robots for safe interaction. The light-weight design permits mobility at low power consumption, introduces additional mechanical compliance to the joints and applies sensor redundancy, in order to ensure the safety of humans in case of robot failure. These robots allow for seamless integration of the work areas of human workers and robots and therefore enable new ways of human-robot cooperation and interaction. Here, the vision is to have human and robot workers work side by side and collaborate as intuitively as human workers would among themselves [1]-[4].

Among all forms of human-robot cooperation, interactive object transfer is one of the most common and fundamental tasks and it is also a very complex and thus difficult one. One major problem for robotic vision systems is visual occlusion, as it dramatically lowers the chance to recognize the target out of a group of objects and then perform successive manipulations on the target. Even without any occlusion, objects in a special position and orientation or close to a human, make it difficult for the robot to find accessible grasping points. Besides, in the case of multiple available grasping points, the robot is confronted with the challenge of deciding on a feasible grasping strategy. When passing the object to the human coworker, the robot has to deal with the tough case of offering good grasping options for the human partner.

A visual servoing system is proposed to address the above-mentioned concerns in human-robot cooperation. Our work considers an interaction task where the robot and human hand over objects between themselves. Situational awareness will be greatly increased by the vision system, which allows for the prediction of the work area occupation and the subsequent limitation of robotic movements in order to protect the human body and the robotic structure from collisions. Meanwhile, the visual servoing control enhances the abilities of robotic systems to deal with the unknown changing surroundings and unpredictable human activities.

Recognition of partially occluded objects will be solved by keeping records of the object trajectory. Whenever the object recognition fails, the last trajectory information of the object will be retrieved for estimating the new location. Reconstruction of the object from its model eliminates the effect of the presence of partial occlusion, and thus enables the successive grasping planning. To equip the robot partner with human-like perception for object grasping and transferring, a planning module will be integrated into the visual servoing system to perform grasping planning, including the location and evaluation of possible grasping points. Thus, the robot, due to its awareness of the object to hand over, will be able to detect, recognize and track the occluded object. Fig. 1(a) considers the occlusion by the human hand, which is one unavoidable barrier among all the possible visual occlusions we are dealing with.

Addressing inaccessible grasping points for the robot, the visual servoing system analyses and evaluates the current situation. The robot adjusts its grippers to a different pose for a new round of grasping affordance planning, as shown in Fig. 1(b, c, d). In some case, this method might fail due to mechanical limitations of the robot. As an alternative, the human coworker will be requested to assist the robot with the unreachable grasping points by presenting the object in other way.

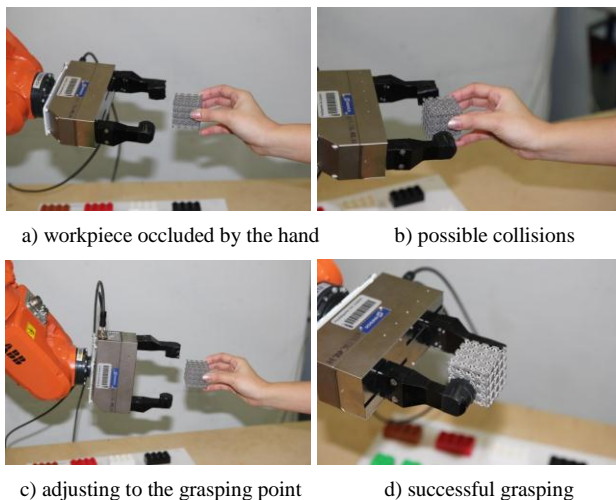


Figure 1. Interactive object human-robot transfer.

The remainder of the paper is organized as follows: section II presents a brief review of the recent literature regarding development of the visual servoing control and state-of-the-art approaches to human-robot interactive object handling. The system architecture and workflow of our proposed visual servoing system are discussed in section III. The key tools and methods for developing the proposed vision system are presented in section IV. Since the visual servoing system has not yet been completely implemented, section V summarizes the research results of this paper and plans on future work.

II. RELATED WORK

A. Visual Servoing

In robotics, the use of visual feedback for motion coordination of a robotic arm is termed visual servoing

[5]. A few decades ago, technological limitations (the absence of powerful processors and the underdevelopment of digital electronics) failed some early works in meeting the strict definition of visual servoing. Traditionally, visual sensing and manipulation are combined in an open-loop fashion: first acquire information of the target, and then act accordingly. The accuracy of operation depends directly on the accuracy of the visual sensor, the manipulator and its controller. The introduction of a visual-feedback control loop serves as an alternative to increasing the accuracy of these subsystems. It improves the overall accuracy of the system: a principle concern in any application [6].

There have been several reports on the use of visual servoing for grasping moving targets. The earliest work has been reported by SRI in 1978 [7]. A visual servoing robot is enabled to pick items from a fast moving conveyor belt by the tracking controller conceived by Zhang *et al.* [8]. The hand-held camera worked at a visual update interval of 140ms. Allen *et al.* [9] used a 60Hz static stereo vision system, to track a target which was moving at 250mm/s. Extending this scenario to grasping a toy train moving on a circular track, Houshangi *et al.* [10] used a fixed overhead camera, and a visual sample interval of 196ms, to enable a Puma 600 robot to grasp a moving target.

Papanikolopoulos *et al.* [11] and Tendick *et al.* [12] carried out research in the application of visual servoing in tele-robotic environment. The employment of visual servoing makes it possible for human to specify the task in terms of visual features, which are selected as a reference for the task. Approaches based on neural networks and general learning algorithms have been used to achieve robot hand-eye coordination [13]. A fixed camera observes objects as well as the robot within the workspace, and learns the relationships between robot joint angles and 3D positions of the end-effector. At the price of training efforts, such systems eliminate the need for complex analytic calculations of the relationships between image features and joint angles.

B. Human-Robot Interactive Object Handling

Transferring the control of an object between a robot and a human is considered a highly complicated task. Possible applications include but are not limited to preparing food, picking up items, and placing items on a shelf [14]-[17]. Related surveys present some research achievements concerning robotic pick-up tasks in the recent years. Jain and Kemp [18] demonstrate their studies in enabling an assistive robot to pick up objects from flat surfaces. In their setup a laser range camera is employed to reconstruct the environment out of the point clouds. Various segmentation processes are then performed to extract flat surfaces and retrieve point sets corresponding to objects. The robot uses a simple heuristic to grasp the object. The authors present a complete performance evaluation towards their system, revealing its efficiency in real conditions.

Other approaches follow image-based methods for grasping novel objects, considering grasping on a small region. Saxena *et al.* [19] create the prediction model for

novel object grasping from supervised learning. The idea is to estimate the 2D location of the grasp based on detected visual features on an image of the target object. From a set of images of the object, the 2D locations can then be triangulated to obtain a 3D grasping point. Obviously, given a complex pick-and-place or fetch-and-carry type of task, issues related to the whole detect-approach-grasp loop [6] have to be considered. Most visual servoing systems, however, only deal with the approach step and disregard issues such as detecting the object of interest in the scene or retrieving its 3D structure in order to perform grasping.

In many robotic applications, manipulation tasks involve forms of cooperative object handling. Papanikolopoulos and Khosla [11] studied the task of a human handing an object to a robot. The experimental results show how human subjects, with no particular instructions, instinctively control the objects position and orientation to match the configuration of the robots hand while it is approaching the object. The human spontaneously tries to simplify the task of the robot. Recent research developments with the NASA Robonaut [20], the AIST HRP-2 [21], and the HERMES [22] robot also address handing over objects between a humanoid robot and a person. Nevertheless, none of these projects have carried out in-depth discussion on object transfer. Our proposed system focuses on planning and implementing interactive human-robot object transfer, addressing the main challenges: visual occlusion and grasping affordance evaluation.

III. SYSTEM ARCHITECTURE

A. Overview

The visual servoing system constitutes the following modules: sensor fusion, calibration, visualization, pose estimation, object tracking, object classification, grasping planning and feedback processing, as shown in Fig. 2. The most primary inputs for the system are sensory data of the targets and the visual feedback. Feature sets and 2D/3D models of the targets are provided beforehand in the forms of 2D images or point clouds and serve as a knowledge base for tracking and classifying of the targets, as well as for the visualization. Physical constraints for the sensing configuration are crucial elements for the system to handle the acquired image data, such as data registration, alignment and object pose estimation.

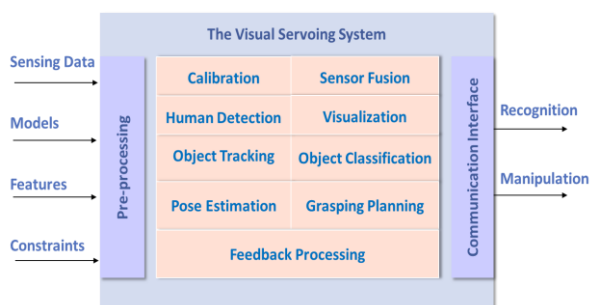


Figure 2. The visual servoing system.

The system will make excessive use of 2D/3D vision processing libraries, such as *PCL* (point cloud library) [23], *OpenCV* (open source computer vision library) [24], *ViSP* (visual servoing platform) [25] within the above-mentioned visual functional modules, including the pre- and post-processing of the image data. For human-robot interactive object grasping the library *GraspIt!* [26], a tool for grasping planning, will be integrated in this system to evaluate each grasp with numeric quality measures. Additionally it also provides simulation methods to allow the user to evaluate the grasp and create arbitrary 3D projections of the 6D grasp wrench space.

To implement the proposed visual servoing system, in our laboratory an experimental platform has been established as shown in Fig. 3. It comprises two ABB IRB120 robots, two Kinect sensors and Lego sets. With the static configuration of Kinect sensors in the platform, the following functions are already realized: multiple sensor calibration and fusion, visualization, object tracking, pose estimation and camera self-localization.

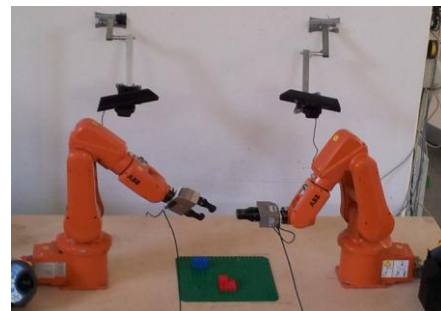


Figure 3. The experimental platform

B. Module Description and Workflow

The main workflow our proposed visual servoing system is depicted in Fig. 4. The workflow comprises four major processes which are noted as follows.

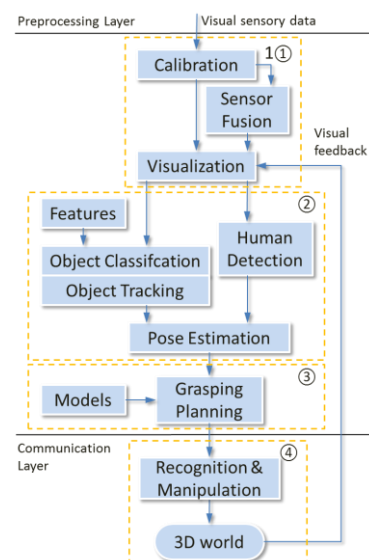


Figure 4. The workflow of the visual servoing system.

- 1) The Calibration module estimates intrinsic and extrinsic camera parameters from several views of a reference pattern, and computes the rectification

transformation that makes the camera optical axes parallel. In many cases, a single view may not pick up sufficient features to recognize an object unambiguously. In various applications this process is extended to a complete sequence of images, usually received from multi-sensors at several viewpoints. If more than one camera in the system, the module calculates the relative position and orientation between each two cameras. This information is then used by the *Sensor Fusion* module to align the visual data from each camera to the same plane and fuse them to form an extended view. The *Visualization* module displays the calibrated image at the selected viewpoint or the image resulting from the fusion.

- 2) Object Classification takes in a set of features to locate objects in videos/images over time in reference of the extracted features (shapes and appearances) of the target object. This approach implements the object identification with this reduced representation. It identifies the target for the Object Tracking module to locate objects in videos/images. *Pose Estimation* calculates the position and orientation of the object in the real world by aligning it to its last pose in the working scene. Additionally the location of the human is roughly estimated by combining the results of human skeleton tracking and face detection.
- 3) With the known location of both the object and human, the Grasping planning module analyzes the current approaching and grasping conditions, based on the present robotic arm and gripper models. The grasping strategies correspond to possible spatial relationships between the target and the robot, as shown in Fig. 5. The occlusion of the object to be grasped is the most likely cause for failures in recognition as well as grasping. Our solution here is to estimate the current object location from its last known pose and extrapolate the current pose making use of the object model. With the estimated pose of the object, the Grasping planning module calculates the possible grasping points and then executes grasping on the object. If none of the grasping points exist in the current situation, the robot will request the human coworker to assist its grasping by adjusting the way he/she presenting the object.

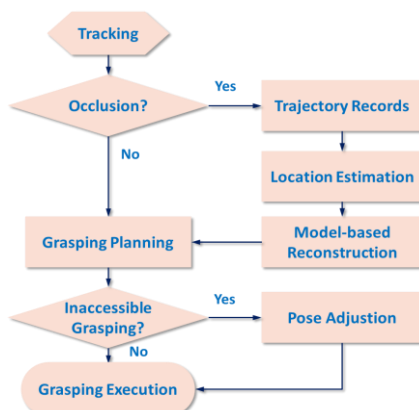


Figure 5. Grasping planning with occlusion.

- 4) At last, the above obtained and processed data are conveyed to the robot controller as *Recognition & Manipulation* inputs to support the operations from the robot on the *3D World*. The visual servoing system assists in making and adjusting the path planning and grasping strategies of robots in real time from *Visual Feedback*.

IV. TOOLS AND METHODS

As mentioned above, the proposed visual servoing system is developed on the basis of several software frameworks (*ROS*, *GrasplIt!*) and image processing libraries (*OpenCV*, *PCL*).

A. Tools

1) ROS

ROS (Robot Operating System) [27] is a software framework for robot software development. It provides standard operating system services such as hardware abstraction, low-level device control implementation of commonly-used functionality, message-passing between processes, and package management. ROS is composed of two main parts: the operating system *ros* as described above and *ros-pkg*, a suite of user contributed packages that implement functionality such as simultaneous localization and mapping, planning, perception, simulation etc.

The *openni_camera* implements a fully-featured ROS camera driver on top of *OpenNI*. It produces point clouds, RGB image messages and associated camera information for calibration, object recognition and alignment. Another package that plays a significant role for our purpose is *tf*, which keeps track of multiple coordinate frames over time. *tf* maintains the relationship between coordinate frames in a tree structure buffered in time, and enables the transform of points, vectors, etc. between any two coordinate frames at any desired point in time.

2) GrasplIt!

GrasplIt! is a simulator that can accommodate arbitrary hand and robot designs. Grasp planning is one of the most widely used tools in *GrasplIt!*. The core of this process is the ability of the system to evaluate many hand postures quickly, and from a functional point of view (i.e. through grasp quality measures).

Automatic grasp planning is a difficult problem because of the huge number of possible hand configurations. Humans simplify the problem by choosing an appropriate prehensile posture appropriate for the object and task to be performed. By modeling an object as a group of shape primitives (spheres, cylinders, cones and boxes) *GrasplIt!* applies user-defined rules to generate a set of grasp starting positions and pregrasp shapes that can then be tested on the object model.

3) OpenCV

OpenCV is an open source computer vision and machine learning software library. OpenCV is built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. The library has a

comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms.

4) PCL

PCL is a large scale, open source project for 2D/3D image and point cloud processing. The PCL framework contains numerous state-of-the-art algorithms including filtering, feature estimation, surface reconstruction, registration, model fitting and segmentation. These algorithms can be used, for example, to filter outliers from noisy data, stitch 3D point clouds together, segment relevant parts of a scene, extract keypoints and compute descriptors to recognize objects in the world based on their geometric appearance, and create surfaces from point clouds and visualize.

B. Methods

1) Visual servoing

There have been two ways of using visual information prevailing in robotic control area [28]. One is the open-loop robot control, which extracts image information and treats control of a robot as two separate tasks where image processing is performed followed by the generation of a control sequence. One way to increase the accuracy of this approach is to introduce a visual feedback loop in the robotic control system, namely visual servoing control. In our human-robot cooperation scenario, the visual servoing system, functioning as shown in Fig. 6, involves acquisition of human pose in addition to object tracking in the traditional systems, in order to carry out the object transfer.

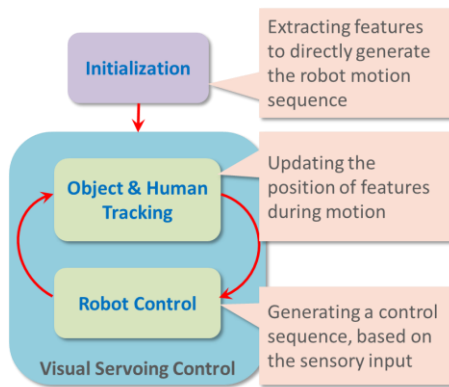


Figure 6. The visual servoing control.

The general ideas behind visual servoing is to derive the relationship between the robot and the sensor spaces from the visual feedback information and to minimize the specified velocity error associated with the robot frame, as shown in Fig. 7. Nearly all of the reported vision systems adopt the dynamic look-and-move approach. It performs the control of the robot in two stages: the vision system provides input to the robot controller; then internal stability of the robot is achieved by conducting motion control commands generated based on joint feedbacks by the controller. Unlike the look-and-move approach, the visual servoing control directly computes the input to the robot joints and thus eliminates the robot controller.

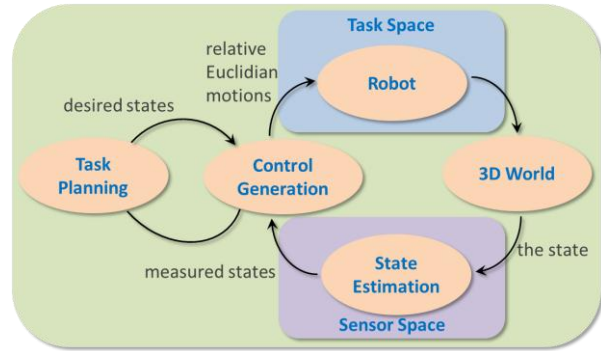


Figure 7. The main processes in vision-based robotic control.

The visual servoing task in our work includes a form of positioning: aligning the gripper with the target object, that is, remaining a constant relationship between the robot gripper and the moving target. In this case, image information is used to measure the error between the current location of the robot and its reference or desired location [28]. Traditionally, image information used to perform a typical visual servoing task is either 2D representation with image plane coordinates, or 3D expression where camera/object model is employed to retrieve pose information with respect to the camera/world/robot coordinate system. So, the robot is controlled either using image information as 2D or 3D. This allows further classifying visual servo systems into position-based and image-based visual servoing systems (PBVS and IBVS, respectively).

In this work, the PBVS approach is applied in the visual servoing system [6, 28]. Features are extracted from the image, and used in conjunction with a geometric model of the target object to determine its pose with respect to the camera. Apparently, PBVS involves no joint feedback information at all, as shown in Fig. 8.

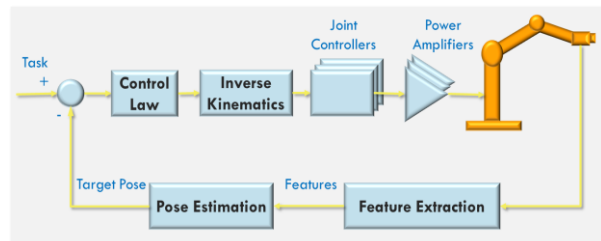


Figure 8. The position-based visual servoing control.

Visual servoing approaches are designed to robustly achieve high precision in object tracking and handling. Therefore, it has great potential in equipping robots with improved autonomy and flexibility in the dynamic working environment, even with humans' participation. One challenge in this application is to provide solutions which are able to overcome position uncertainties. Addressing this tough problem, the system offers dynamic pose information of the target to be handled by the robotic system via the *Object Tracking* and *Pose Estimation* modules.

2) Grasping planning

Grasp planning of a complex object has been thought too computationally expensive to be performed in real-

time. Therefore, the pre-computation, as a preferable substitute to real-time computation, has been always involved in practical applications. Our proposed system adopts the grasping planning approach presented by Sidobre *et al.* [29]. The idea is to compute the grasp affordance off-line (in our case, by *GraspIt!*), for the target object to collect the best variety of possible grasps. This affordance collection will then be used to select the best grasps from, according to a scoring function which indicates good grasp quality with a high score.

Several aspects need to be taken into account to compute a grasp quality score out of our scoring function. Researchers have identified a multitude of properties that an articulated force-closure grasp must possess so that the robot can perform everyday tasks similar to the way that humans perform. The interrelationship between those properties can be delineated into four mutually independent properties [30], [31]: dexterity (the configuration of the robotic grippers), equilibrium (the squeeze imposed on the grasped object), stability (resistance to the external disturbances) and dynamic behavior (the reaction of the grasp to the changing contact status). In our work, a tradeoff is chosen with a score that is a weighted sum of these four measures.

The on-line grasp planning phase takes on when the grasp affordance of the specified object is computed. It begins with the selection of a grasp from the collection. If there were no additional constraints, the highest quality score can be safely counted as an indicator to the optimal grasp choice. However, the robot must consider the limited reachability of its arm and the safety of the human handing over the object. An inverse kinematics computation is required in order to ensure the availability of the grasp. To facilitate the convergence of the robot's and human's motions, the human interprets the robot's motion and assists it with its limited reachability.

V. CONCLUSION AND FUTURE WORK

A. Conclusion

In this paper, we introduce a prototype of a visual servoing system for interactive human-robot object transfer. The system integrates basic vision modules to identify and position the target object. It considers the human coworker's participation and plans for the collaborative task for the robot. Our proposed visual servoing system is characterized by its solutions to partial visual occlusion and inaccessible grasping points in robotic object handling.

Visual occlusion, the major challenge in vision monitoring, is addressed by employing the information from object tracking to generate a history of the trajectory of the target object. In this way, the effect of partial occlusion on the object can be eliminated by reconstructing the object pose and displacement. An extra grasping planning module enables the evaluation of the current grasping affordance, based on the poses of both the object and the human coworker. If the absence of

grasping possibilities is attributed to the robot pose, the system will make the request for the robot to relocate its arm and then re-approach the object. In the case of inaccessible grasping points caused by the human, signals will be sent out as instructions for the human coworker to alter the way he/she presents the object. In the process, real-time visual feedback is provided to relevant modules as visual information update.

B. Future Work

To continue the development of our proposed visual system, more efforts will be made on implementing the vision processing modules in addition to the completed ones. Afterwards, those modules will be integrated into the visual servoing system and the functionalities will be verified through groups of experiments. Moreover, evaluating criteria and approaches will be generated to understand the performances and characteristics of the system. Following that, experimental studies will be conducted towards human-robot interaction with visual signals input to the robotic control unit. A thorough performance evaluation will be carried out and optimizations will be made on the system to achieve the three main goals:

- To propose visual servoing methods to improve the autonomy and situational awareness of the robot partner to increase task-orientation and flexibility in order to work as a team member rather than a tool;
- To enable the robot to realize object recognition with partially obstructed objects;
- To provide visual information for the integrated grasping planning tool to evaluate potential grasps.

ACKNOWLEDGMENT

This work is supported by the German Research Foundation (DFG) within the Cluster of Excellence "Integrative production technology for high-wage countries" at RWTH Aachen University.

REFERENCES

- [1] M. Bratman, "Shared cooperative activity," *The Philosophical Review*, vol. 101, no. 2, pp. 327-341, 1992.
- [2] B. Grosz, "Collaborative systems," *American Association for Artificial Intelligence*, vol. 2, no. 17, pp. 67-85, 1996.
- [3] P. Cohen and H. Levesque, "Persistence, intention, and commitment," in *Intentions in Communication*, Cohen, Morgan and Pollack, eds. MIT Press, 1990, ch.3.
- [4] S. Jeschke, E. Hauck, O. Pfeiffer, and T. Richter, "Robotics as demonstrator discipline for networked virtual and remote laboratories," in *Proc. IEEE Workshop: Advanced Robotics and Its Social Impacts*, Tokyo, 2009.
- [5] P. D. Adsit, "Real-time intelligent control of a vision-servoed fruit-picking robot," Ph. D thesis, University of Florida, 1989.
- [6] P. Corke, "Visual control of robot manipulators-a review," *Visual Servoing*, vol. 7, pp. 1-31, 1993.
- [7] C. Rosen, *et al.*, "Machine intelligence research applied to industrial automation," *Southwest Research Inst. Report*, 1980.
- [8] D. B. Zhang, L. Van Gool, and A. Oosterlinck, "Stochastic predictive control of robot tracking systems with dynamic visual

- feedback," *IEEE International Conference of Robotics and Automation*, 1993, pp. 610-615, 1990.
- [9] P. K. Allen, B. Yoshimi, and A. Timcenko, "Real-time visual servoing," *IEEE International Conference of Robotic and Automation*, 1991, pp. 851-856.
 - [10] N. Houshang, "Control of a robotic manipulator to grasp a moving target using vision," *IEEE International Conference of Robotics and Automation*, 1993, pp. 604-609.
 - [11] N. P. Papanikolopoulos and P. K. Khosla, "Shared and traded telerobotic visual control," *IEEE International Conference of Robotics and Automation*, 1992, pp. 878-885.
 - [12] F. Tendick, J. Voichick, G. Sharp, and L. Stark, "A supervisory telerobotic control system using model-based vision feedback," *IEEE International Conference of Robotics and Automation*, 1991, pp. 2280-2285.
 - [13] C. Breazeal, *et al.*, "Working collaboratively with humanoid robots," in *Proc. 4th IEEE/RAS International Conf. Humanoid Robots*, vol. 1, 2004.
 - [14] P. Corke, *Visual Control of Robots: High-Performance Visual Servoing*, Research Studies Press, 1996, ch. 5.
 - [15] T. Oggier, *et al.*, "An all-solid-state optical range camera for 3d real-time imaging with sub-centimeter depth resolution (swissranger)," vol. 5249, pp. 534-545, 2004.
 - [16] C. Stanger, C. Anglin, W. Harwin, and D. Romilly, "Devices for assisting manipulation: A summary of user task priorities," *IEEE Trans. Rehabilitation Engineering*, vol. 2, no. 4, pp. 256-265, Dec. 1994.
 - [17] R. Bischoff and V. Graefe, "Design principles for dependable robotics assistants," *International Journal of Humanoid Robotics*, vol. 1, no. 1, pp. 95-125, 2005.
 - [18] A. Jain and C. Kemp, "El-e: An assistive mobile manipulator that autonomously fetches objects from flat surfaces," *Autonomous Robots*, vol. 28, no. 1, pp. 45-64, 2010.
 - [19] A. Saxena, J. Driemeyer, and A. Ng, "Robotic grasping of novel objects using vision," *International Journal of Robotics Research*, vol. 27, no. 2, pp. 157-173, 2008.
 - [20] A. Edsinger and C. Kemp, "Human-robot interaction for cooperative manipulation: Handing objects to one another," in *Proc. 16th IEEE International Symposium on Robot Human Interactive Communication*, Jeju Island, Korea, 2007, pp. 1167-1172.
 - [21] N. Sian, K. Yoki, Y. Kawai, and K. Muruyama, "Operating humanoid robots in human environments," in *Proc. the Robotics, Science & Systems Workshop on Manipulation for Human Environments*, Philadelphia, Pennsylvania, Aug. 2006.
 - [22] M. Diftler, R. Ambrose, K. Tyree, S. Goza, and E. Huber, "A mobile autonomous humanoid assistant," in *Proc. 4th IEEE/RAS International Conf. Humanoid Robots*, 2004.
 - [23] PCL.org. [Online]. Available: <http://pointclouds.org/>
 - [24] OpenCV.org. [Online]. Available: <http://opencv.org/>
 - [25] ViSP. [Online]. Available: <http://www.irisa.fr/lagadic/visp/publication.html>
 - [26] GraspIt! [Online]. Available: <http://www.cs.columbia.edu/~cmatei/graspit/>
 - [27] ROS.org. [Online]. Available: <http://wiki.ros.org/>
 - [28] D. Kragic and H. I. Christensen, "Survey on visual servoing for manipulation," *Computational Vision and Active Perception Laboratory*, Fiskartorpsv 15, 2002.
 - [29] D. Sidobre, *et al.*, *Reactive Grasping for Human-Robot Interaction*, 2010.
 - [30] K. B. Shimoga, "Robot grasp synthesis algorithms: A survey," *The International Journal of Robotics Research*, vol. 15, no. 3, pp. 230-266, 1996.
 - [31] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Apr. 2000, pp. 348-353.



Ying Wang (born in China) gained Master of Science in mechatronics engineering from Harbin Institute of Technology in 2009. Since 2012.10 she has been doing her PhD in mechanical engineering in RWTH Aachen University. She worked as a Software Engineer intern in Shenzhen Hans Photoelectric Equipment Company on developing of servo controller for dies bonders (2008-2009). After that, she was Test Engineer on radiation treatment planning software with Elekta Shanghai. She is now a scientific research assistant with the Institute for Management Cybernetics (IfU) in RWTH Aachen University. She works in the interdisciplinary research field in computer vision and HRI.



production, mobile robotics within intralogistic scenarios and human-machine interfaces.



optimization of cooperation, networking processes and management in knowledge-intensive research networks. Additional Dr. Vossen works as lecturer, coach and consultant.



the Institute of Information Technology Services for electrical engineering at the University of Stuttgart (2007-2009), where she was also the director of the Central Information Technology Services at the same time. Some of the main areas of her research are complex IT-systems, robotics and, traffic and mobility and virtual worlds for research alliances. Prof. Dr. rer. nat. Jeschke is vice dean of the Faculty of Mechanical Engineering of the RWTH Aachen University, chairwoman of the board of management of the VDI Aachen, a member and consultant of numerous committees and commissions, alumni of the German National Academic Foundation, IEEE Senior Member and Fellow of the RWTH Aachen University.

Daniel Ewert (born in Germany) studied computer science with a minor in neuro biology at the RWTH Aachen University. He gained his doctorate on "Adaptive scheduling for production in the Factory of the Future" in 2013. His studies focused on artificial intelligence and knowledge-based systems. He is the head of the research group "Engineering Cybernetics" with IfU. Since 2013.8. Current research interests comprise: intelligent planning and control systems for

René Vossen (born in Germany) studied Psychology, Communication Sciences, and Sociology at RWTH Aachen University. Since 2013 he is Managing Director of IfU. Dr. Vossen has been working at the institute cluster IMA/ZLW & IfU as research assistant and research group leader since 2008. In 2012, he completed his doctorate on applied research on the intellectual capital of clusters of excellence. His main areas of research are cybernetic knowledge management, the

Sabina Jeschke (born in Sweden) studied physics, computer science and mathematics at the Berlin University of Technology. After research at the NASA Ames Research Center/ California and the Georgia Institute of Technology/Atlanta, she gained a doctorate on "Mathematics in Virtual Knowledge Environments" in 2004. Following a junior professorship (2005-2007) at the TU Berlin with the construction and direction of its media center, she was head of