25th IEEE International Symposium on
Robot and Human Interactive Communication (RO-MAN)
August 26-31, 2016. Columbia University, NY, USA

Projecting Robot Intentions into Human Environments

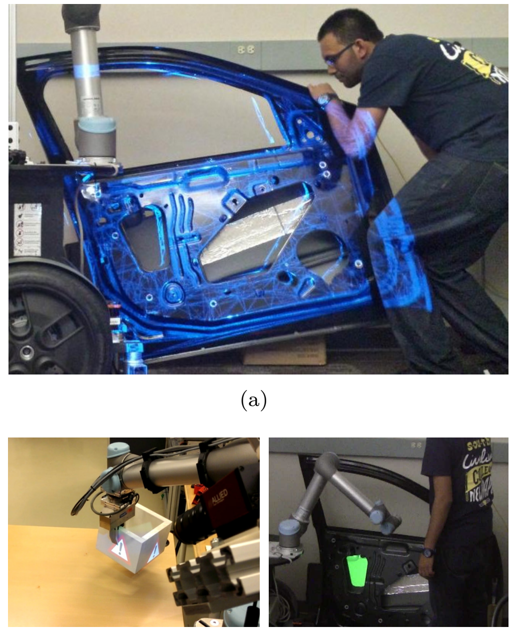
Rasmus S. Andersen[1], Ole Madsen[1], Thomas B. Moeslund[2], and Heni Ben Amor[3]

***Abstract*— Trained human co-workers can often easily predict each other's intentions based on prior experience. When collaborating with a robot co-worker, however, intentions are hard or impossible to infer. This difficulty of mental introspection makes human-robot collaboration challenging and can lead to dangerous misunderstandings. In this paper, we present a novel, *object-aware* projection technique that allows robots to visualize task information and intentions on physical objects in the environment. The approach uses modern object tracking methods in order to display information at specific spatial locations taking into account the pose and shape of surrounding objects. As a result, a human co-worker can be informed in a timely manner about the safety of the workspace, the site of next robot manipulation tasks, and next subtasks to perform. A preliminary usability study compares the approach to collaboration approaches based on monitors and printed text. The study indicates that, on average, the user effectiveness and satisfaction is higher with the projection based approach.**

I. INTRODUCTION

Traditional automated production leverages the speed, precision, and reliability of industrial robots to produce large quantities of identical or similar items. Such industrial robots are well suited for carrying out simple tasks with a high degree of repeatability. For production of smaller production series, however, an increasingly higher degree of customization is required. This pushes demand for flexible manufacturing lines including intelligent robots. Industrial robotics is therefore currently experiencing a major trend moving from robots carrying out simple, repetitive tasks while being separated from human workers by security fences, towards collaborative robots that carry out increasingly complex tasks [1]. Collaborative robots do not need to be physically separated from human co-workers, but can instead work in a shared workspace. Contrary to traditional industrial robots, it is therefore possible for collaborative robots to directly interact with human co-workers on joint physical tasks.

In this paper, we propose to facilitate interaction between industrial collaborative robots and human workers by projecting information about the current task directly into the workspace. This information can include the state and intentions of the robot, instructions to the human, and information about the current state of the task. A novel object-aware projection mapping system is developed and combined with existing object tracking technology, which allows for precise 6-DOF object pose detection from a standard RGB-camera. The tracker enables the display of information on objects before, during, and after they are manipulated by either the human or the robot. Two examples are shown in Figure 1. In (a), a car door is being tracked during a collaborative assembly task[1]. A wireframe model of the car door is projected onto the door itself, while it is manipulated by a human. The projected door provides feedback about the perception quality and accuracy of the robot. In (b), a sign is projected onto an assembled box warning humans that a robot intends to manipulate this particular object.



(a) (b) (c)

Fig. 1. Interaction with tracked objects while information is being projected onto the objects in real-time. In (a), a wireframe is being projected onto a car door while it is being moved by a human. In (b), a sign warns a user that a robot will interact and is interacting with the particular object. In (c), the robot highlights a segment of the object on which it will perform a manipulation.

[1]Rasmus S. Andersen and Ole Madsen are with the Department of Mechanical and Manufacturing Engineering, Aalborg University, Denmark `{rsa,om}@m-tech.aau.dk`

[2]Thomas B. Moeslund is with the Department of Architecture, Design and Media Technology, Aalborg University, Denmark `tbm@create.aau.dk`

[3]Heni Ben Amor is with the Interactive Robotics Lab, Arizona State University, Arizona, United States `hbenamor@asu.edu`

[1]A video is available at `https://youtu.be/FT94aH98bew`

Finally, in (c) the robot identifies the sub-component of the car on which it will perform the next manipulation, thereby clearly indicating to the human its intent. Both the shape of the projection, as well as the color of the projection are used as cognitive cues to communicate information.

The main contribution of this work is an *object-aware* human-robot interaction approach intended for collaborative robots which uses projections to provide relevant information to human co-workers directly in the common workspace. By continuously tracking task-related objects, we can use information about their shape and pose in order to provide accurate visual cues at precise spatial locations to a human collaboration partner. This allows for rapid and explicit communication of the robot's mental states. The approach is evaluated against common interaction approaches in a comparative usability study.

A secondary contribution of this paper is a robust implementation of simultaneous object detection, tracking and projection mapping. In practice, it is challenging to visually track an object while at the same time changing its visual appearance through augmented imagery. This creates a closed-loop tracking-and-projection system, which could result in drifting and unbounded errors. We identify here a reduction of the problem and propose a working solution which in practice provides stable results.

The paper is organized as follows: In Section II, related work within human-robot interaction is reviewed. In Section III, potential use cases for the intention projection system are presented. In Section IV, our approach is described, and finally in Section V, experiments are presented which evaluate both contributions of the paper.

II. RELATED WORK

The idea of using projection to understand the state and intentions of a robot is not new in itself. One of the first attempts was made by Ishii et al. in 2009, who used a projection interface to give commands to a mobile robot [2]. A user draws circles around objects using an IMU pointing device. The drawings are visualized by a projector and when a user accepts the task, the mobile robot moves to the marked object and picks it up. More recent work has explored how to visualize complex robot information in human environments. Several research groups have designed workspaces with a number of projectors mounted in the ceiling covering a large floor area [3], [4]. For instance, the MAR-CPS system presented in [3] combines a ceiling-to-floor projection system with a motion capture system that is capable of tracking flying drones as well as driving robots. This is used for different research areas, including surveillance coverage, path planning, and to identify robot states and error messages in general. Common to these setups, however, is that they display additional information by projecting onto flat surfaces in the environment, e.g., the floor. The surface becomes a replacement for the flat display screen. Hence, the projection does not take into account the 3D nature of the environment and does not include information about the shape and pose of object. Projections onto specific parts of an object are therefore infeasible.

In a similar vein, Leutert et al. propose an interaction system, SAR, which employs projections on flat surfaces to facilitate interactions with an industrial robots [5]. They combine images from an external projector and a flange-mounted projector to visualize various data, including instructions and dynamic movement data. They propose to give instructions to the robot using an input device that is tracked in 3D and then visualize the instructions in real time using projections. The movement data includes floor projections of the intended future pose of the robot. In addition to the above mentioned limitations, their work does not include any evaluation of the usability of the interaction system.

In a recent paper, Chadalavada et al. propose to use a projector mounted on a mobile robot to visualize a small fraction of its intended path a few meters ahead [6]. Test persons were asked to walk towards and pass the robot with and without projection enabled, and the average user rating increased by 53% with projection, if the robot moved in a straight line and by 65% if the robot took a sudden turn. Even though the use case for this study is different from what is suggested in the current paper, it does suggest that direct visual feedback in a common space in general can be advantageous.

Besides approaches that use an explicit approach for communicating robot intentions, e.g. projection, recently there has also been growing interest in implicit approaches. Dragan and colleagues [7]–[9] investigated how robot motion can be planned so as to clearly communicate intentions to a human partner. They argue that legible robot motion allows human interaction partners to infer future goals of the robot. A similar approach is also followed in [10]–[12]. In general, this line of research treats robot motion as a communication channel through which information can be conveyed. Yet, the amount and type of information that can be transmitted through motion is limited. In contrast to that, the explicit projection technique presented in our paper makes it possible to project a much wider variety of information which are useful for actual industrial human-robot collaboration tasks.

III. USE CASES FOR INTENTION PROJECTION

Understanding the internal state and intentions of a robot is relevant to many situations involving modern, collaborative robots. This includes situations where humans and robots have to enter each other's workspace, as well as situations where workspaces are shared between robots and humans for longer durations. Typical tasks at a manufacturing plant involve machine tending,

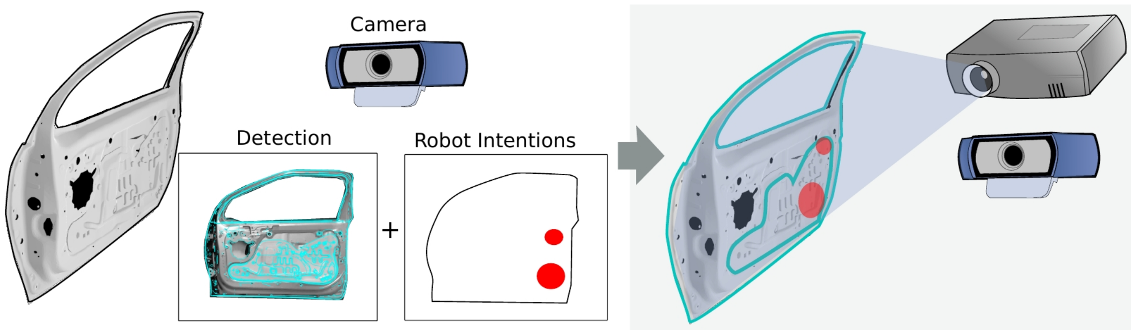


Fig. 2. Overview of the system: A physical object (car door) is first detected using model-based tracking. Information about robot intentions, e.g., next actions, is displayed by rendering a 2D texture on top of the object model. Finally, the rendered image is projected into the real world.

assembly, inspection, logistic tasks, and service tasks. One workspace can be used for many tasks, and if robots and humans have responsibility for different tasks, coordination is necessary. In some cases, humans and robots may even have to cooperate on the same tasks; for instance an assembly task involving heavy parts.

Fig. 3. Example domain: car doors are being transported on a conveyor belt. Both human and robot are required to engage as co-workers on the doors.

A use-case from a car factory is presented as an example throughout the current paper. Car doors are transported hanging from a conveyor belt as shown in Figure 3. Human workers as well as robots must work on the doors while they are being transported. The doors are not fixed, hence their pose must be constantly tracked.

Figure 4(a)-(d) shows an interaction in a laboratory setup. In (a), a human worker moves the car door while the tracked pose is visualized as a wireframe projected onto the door. The display of the wireframe provides feedback about the robot vision system. Failure modes of the vision system can be detected whenever the wireframe does not fit the underlying physical model.

Intention projection can also be used to project the next manipulation location of a robot, or next grasping locations. In Figure 4(b)-(d), for instance, the robot is interacting with different parts of the door, and these parts are marked by the projector in advance. This makes it possible for the human to take the necessary precautions by avoiding those areas. Color coding of the projections can further be used in order to encode current states of the robot, e.g., high stiffness vs. low stiffness modes. In training scenarios, intention projection can be used by a robot to communicate next steps or tasks to a human partner. Projecting such information on top of the required object reduces ambiguities and miscommunication. The human interaction partner is not required to divide attention between an external information display (screen) and the spatial location of the interaction. As a result, attention, fluency and safety can be improved during the collaborative task.

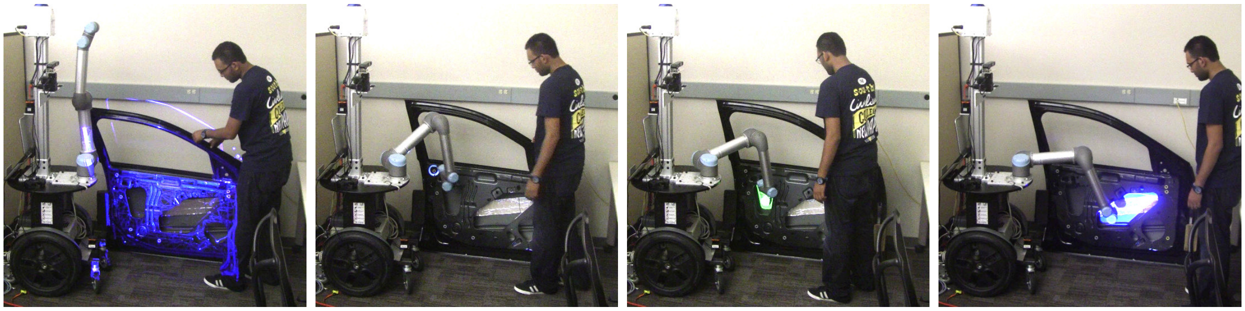
IV. INTENTION PROJECTION FRAMEWORK

The intention projection system uses the environment as a canvas in order to display critical information. To this end, two steps are needed, namely (a) detection and tracking of objects using a standard RGB-camera, and (b) projection of important information into the scene. Figure 2 gives an overview of this setup. The system consists of several subcomponents; object tracking, pose estimation, and rendering. Each of these is described in the following subsections.

A. Object Tracking

Objects can be tracked using different technologies; two the most popular being passive vision and active structured light depth sensing. For the current task, passive vision has the advantage when compared to structured light depth sensing that it is not dependent of the object surface' ability to reflect the emitted light. Additionally, its performance is more robust to direct and indirect sunlight. A problem for passive vision is that projected light in the visible spectrum can disturb the tracking. A way to handle this is to synchronize the shutter systems of the projector and the camera so that the projector is off whenever the camera captures an image.

In the current system, passive vision based on off-the-shelf components is preferred instead of a highly specialized and expensive synchronized projector-camera setup. When much graphics is projected, the interference is instead reduced by projecting and capturing in different ends of the visible light spectrum. The object tracking is based on the Iterative Re-weighted Least Squares (IRLS)-tracker by Choi and Christensen [13].



(a) (b) (c) (d)

Fig. 4. Figure (a)-(d) is from a laboratory setup. In (a), the car door is manipulated with simultaneous real-time visualization of the tracking. In (b)-(d), a robot's intention to interact with particular parts of the car door is projected. This allows a human to take precautions.

The tracker relies on a CAD model of the object to be tracked. Sharp and dull edge features are extracted offline from the model. Depending on a hypothesized pose, a set of salient edges is created by combining all non-occluded sharp edges with dull edges that are part of the object's silhouette. At run-time, edges are extracted frame-by-frame using the Canny edge detector [14]. Points on the model edges from frame *t-1* are matched to detected edges in frame t by searching along a line orthogonal to the direction of the model edge. This results in an error vector which is minimized iteratively by modifying the pose of the model. An exhaustive description of the IRLS-tracker is beyond the scope of this paper, and the reader is instead referred to [13].

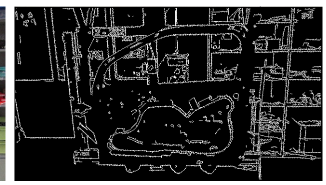
B. SURF Based Tracker Initialization

The IRLS tracking algorithm requires initialization with a pose of the object to track. Also, in the case of tracking failure, it needs to be re-initialized. In previous works by Choi and Christensen, a SURF based initialization is employed with convincing results [13], [15]. A number of images of the object is captured while the object is simultaneously tracked. In each of the images, SURF features are estimated and mapped to the object. When the SURF initialization is subsequently used, SURF features are detected in the present image and mapped to the stored features. In [13], where the pose estimation is used in conjunction with the IRLS tracker, RANSAC is employed to filter SURF matches.

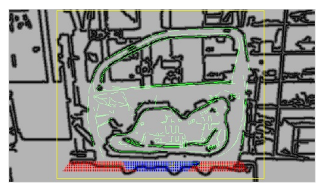
C. Projection-Robust Tracker Initialization

As will be shown through experiments in Section V-A, the SURF based approach performs poorly when graphics are projected onto a significant part of the object. Even relatively dimmed light significantly obstructs the local features that such a method relies on. Experimental results on this are presented in Section V-A. For the system presented in this work, a template-based approach has therefore been employed which relies on edge features similar to the IRLS tracker. The method is illustrated in Figure 5.

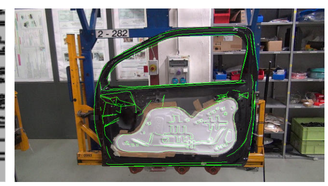
(a) Input frame



(b) Canny edges



(c) ROI and pose hypothesis testing



(d) Final pose

Fig. 5. Edge based initialization.

In the edge-based initialization, a large number of edge maps is first generated offline from the object CAD model for all legal, hypothetical poses. It is assumed that objects are initially placed facing up in a certain height; typically on a supporting surface. The search space can therefore be reduced by only searching through two spatial dimensions and one rotational dimension, instead of the full 6-DOF space.

At runtime, edges are detected using the Canny edge detector as seen in Figure 5(a) and (b). A distance transform using the Euclidean norm is applied to the edge map to facilitate robust matching of the model edge maps. A threshold is subsequently applied to the distance map to avoid extreme values far from any edges. The resulting "softened" edge map is shown as a black/white background image in Figure 5(c).

Recent movement in the scene is used to limit the region-of-interest (RoI), as indicated in Figure 5(c) by the yellow rectangle. The search area has been defined in advance, and hypothetical object positions are displayed in the figure as red and blue dots. To increase computation speed, model edge maps are disregarded if they cannot fit inside the RoI. In the figure, red dots indicate positions where no model edge maps were found to be legal. The remaining model edge maps are matched to the distance transform by computing a sum of all edge locations. The best pose is shown in Figure 5(d).

D. Rendering System

The task of the rendering system is to generate a visual representation of intent and to project it into the environment in a geometrically correct manner as shown in Figure 1. The projector is calibrated intrinsically and extrinsically with respect to the camera (refer to the overview in Figure 2). The resulting transformations are, in turn, used to set projection and the modelview matrix, respectively. Other transformations are added to the modelview matrix, including detected object poses and relative positions in the environment. All objects and visual entities are represented in a scenegraph for rendering. The rendering process in effect pre-warps the generated graphics, thereby producing an accurate projection when the image is distorted by the shape of the physical object. Our current system runs in real-time at frame rates of about 20–30Hz, including object tracking and rendering.

V. EXPERIMENTAL RESULTS

The proposed projection system is evaluated both technically (quantitative) and through a usability study (qualitative). The technical evaluation estimates the robustness of the proposed pose estimation algorithm used for initializing the tracker, while the usability study compares the usability of the proposed intention projection to traditional interfaces.

A. Robustness of Initialization

Both methods for initializing the pose estimation mentioned in Section IV-A, one based on SURF features and one based on edge features, have been tested and compared. The evaluation was performed by repeatedly estimating the pose of the car door in the setup shown in Figure 6.

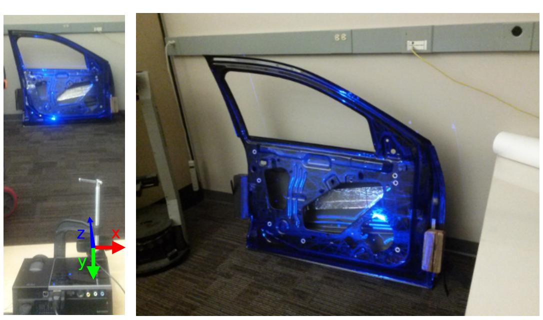


Fig. 6. Setup for the pose estimation test. The camera and projector is shown in the bottom of the left image.

The purpose of the test is to evaluate how the performance of the methods is affected by changing light conditions as well as graphics projected onto the object. Each method has been tested under three different conditions:

1) **Full light:** Normal indoor light conditions. This condition has also been used to train the SURF based method.
2) **Low light:** Lower indoor light.
3) **Projected wireframe:** Lower indoor light with a wireframe being projected onto the object. A wireframe is used because this fills most of the surface of the object. It can therefore be considered as a *worst case* projection.

The effect of the wireframe is minimized by projecting only blue light and using only the red color channel for pose estimation. The success rates and time consumptions have been estimated by performing at least 100 successful pose estimations in each scenario, and the results are shown in Table I.

TABLE I
SUCCESS RATE AND TIME CONSUMPTION FOR POSE ESTIMATION METHODS

Method	Setting	Success rate	Avg duration
SURF	Full light	49.8%	494.7 ms
SURF	Low light	17.6%	487.2 ms
SURF	Proj. wireframe	9.9%	554.1 ms
Edge	Full light	(100%)	15.3 ms
Edge	Low light	(100%)	13.9 ms
Edge	Proj. wireframe	(100%)	16.9 ms

The success criterion for the SURF method is based on the RANSAC algorithm; if too few features are matched, estimation is considered a failure and the pose estimation is restarted. The edge-based method does not employ a similar self-validation method, and all estimations are therefore used. Poor pose estimations of the edge-based method will therefore reduce the mean accuracy of the edge based method, while the SURF based method is able to remove and disregard some poor matches.

It is clear from the success rate that the SURF based approach is heavily affected by both changing light conditions and projections, even when different colors are used for projection and pose estimation. At the same time, the SURF based method is more than an order of magnitude slower than the edge based method. It should be noted, however, that both methods in this test use single threaded CPU implementations and that their computation time can be reduced using parallel GPU implementations.

The accuracy of the successful pose estimations for all scenarios should ideally be determined by comparing the results against a ground truth pose. Ground truth is unfortunately not available in our case. Instead, the steady state pose from the IRLS tracker under optimal light conditions is used. This will unavoidably affect the estimated errors; possibly to the advantage of the edge-based method, since IRLS is itself also edge-based. The SURF-based method is, however, based on surf features, which are also found during steady state IRLS tracking.

The relative performance between the scenarios should therefore only be slightly affected. Results are shown in Figure 7.

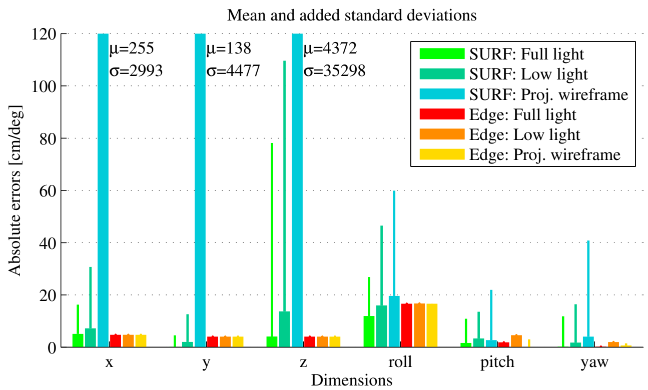
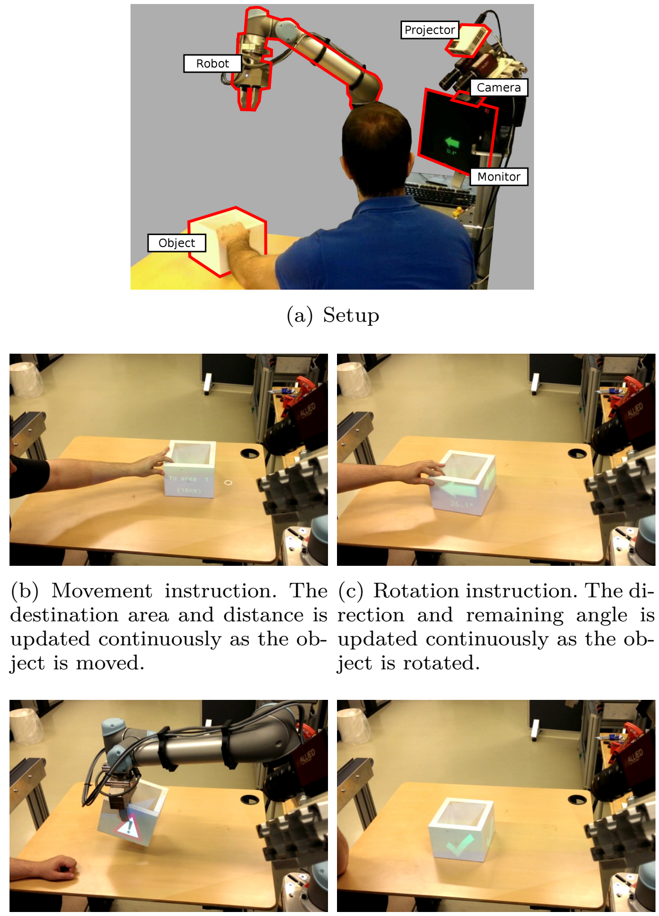


Fig. 7. Comparison of the accuracy of the pose estimation methods. The coordinate frame is oriented as shown in Figure 6. The mean absolute error of each method when compared to the steady state pose detected by the IRLS tracker is shown as wide bars. The thinner lines are the standard deviation added to the means.

SURF-based methods perform poorly when a wireframe image is projected. This is the case even though only the 9.9% of the poses that were evaluated as *successful* are included in these results. A manual sorting through the data indicates that some estimates were actually successful; however other estimations were very far off. The edge-based method performs, on the other hand, almost identical under different conditions. The standard deviation is significantly lower for the edge-based method, partly because the algorithm of this method is deterministic, whereas the SURF based method is non-deterministic as a result of using RANSAC.



(a) Setup

(b) Movement instruction. The destination area and distance is updated continuously as the object is moved.

(c) Rotation instruction. The direction and remaining angle is updated continuously as the object is rotated.

(d) Robot interaction warning

(e) Manipulation complete

Fig. 8. Setup and provided information for the usability study. The task was to either rotate or move the white box based on instructions provided by one of three methods. Figure (b)-(e) shows the projected information. Sometimes a subtask is carried out by the robot, and in this case the test person is notified by a warning triangle as shown in (d).

B. *User Study on Usability*

A usability study has been carried out to evaluate the usability of intention projection compared to traditional approaches. A total of 14 test persons with diverse backgrounds participated. Each test participant was asked to complete three tasks in collaboration with a robot. Each task consisted of 15 subtasks which included rotating an object and moving it between three marked areas. The following subtasks were used:

- Move object to area X (human)
- Rotate object Y° (human)
- Move to area X (robot)
- Rotate Y° (robot)

The test setup is shown in Figure 8(a). For each of the tasks, a different interface for collaboration was used:

1) **Projected information:** With the projection approach, information is given for each subtask as shown in Figure 8(b)-(d). Information is projected directly onto all sides of the object facing the projector. The information is updated in real-time as the object is moved and rotated. For human-move instructions, the destination area is additionally marked with a white circle as seen in Figure 8(b). Whenever a subtask is successfully completed, check marks are shown on the object as in Figure 8(e).
2) **Monitor display:** In the monitor display based approach, the same information is given; the only difference being that it is shown on a monitor attached to the robot, as seen in Figure 8(a). The destination areas are marked with numbers on the table.
3) **Text description:** In the text based approach, the entire task is presented as a numbered list and given to the test person before the start of the experiment.The destination areas are marked with numbers on the table, and illustrations on the text sheet show how the box should be oriented. No validation is performed w.r.t. either position or orientation with this method.

The usability of each approach is evaluated according to the ISO standard 9241-11 (1998) [16], which defines usability as a combination of three factors:

- **Effectiveness:** *"The accuracy and completeness with which users achieve specified goals."*
- **Efficiency:** *"The spent resources in relation to the accuracy and completeness with which users achieve specified goals."*
- **Satisfaction:** *"The freedom from discomfort and positive attitudes towards the use of the product."*

Results for each of the evaluation criteria are presented below.

1) Effectiveness: The effectiveness is measured objectively as the number of times in which the participant required help and/or failed in completing a subtask. In general, not much help was needed since all subtasks were relatively simple. However, especially the text based approach caused some problems. Three test persons asked for assistance in determining if they had completed one or more of their subtasks correctly due to the lack of visual feedback. One subject mistakenly rotated the object around one of its corners instead of its center (as was instructed). A fourth test person was unable to keep track of the current subtask and therefore skipped a human subtask and was not aware of a robot subtask; causing a near-collision with the robot. The collision was only prevented through an emergency stop by the experimenter. A total of six significant questions/errors were thus counted for this approach.

In the monitor based approach, several test persons had difficulties relating an arrow on the monitor to a rotational direction of the objects. Most test persons eventually figured out the relation, but three test persons had to be assisted.

In the projection based approach, no significant questions or errors were present. One test person asked for confirmation before each human subtask but did so for all interfaces and was always correct. The effectiveness of the three approaches is compared in Table II.

TABLE II

EFFECTIVENESS MEASURED AS THE NUMBER OF QUESTIONS AND ERRORS

Method	Questions	Errors	Total pr. person
Projector	0	0	0.00
Monitor	3	0	0.21
Text	3	3	0.43

2) Efficiency: The efficiency is measured objectively as the time spent to complete a task. The time consumption was almost identical for all collaboration approaches as shown in the last column in Figure 9. One reason for this is that the majority of the total time was taken up by the robot's movements and not by the actions of the participants.

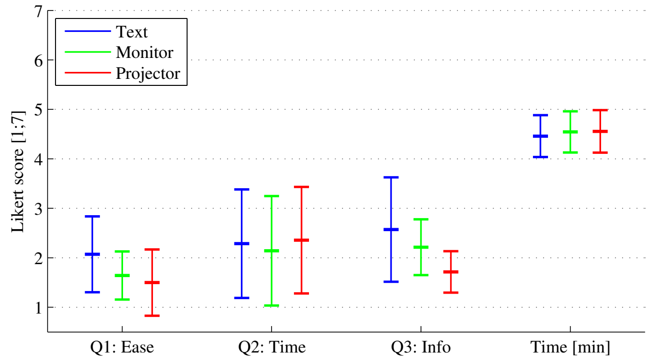


Fig. 9. Results from the usability study concerning satisfaction (Q1-Q3) and efficiency (time). The mean scores 95% confidence intervals are shown. For Q1-Q3, 1 means *strongly agree* while 7 means *strongly disagree*.

3) Satisfaction: The satisfaction of the participants was measured through questionnaires based on Lewis' After-Scenario Questionnaire (ASQ) [17]. For each collaboration approach, they were asked to indicate on a 1-7 Likert scale the extent to which they agreed with the following statements:

Q1) Overall, I am satisfied with the **ease** of completing the tasks in this scenario.

Q2) Overall, I am satisfied with the **amount of time** it took to complete the tasks in this scenario.

Q3) Overall, I am satisfied with the **support information** (projections, display on monitor, text) when completing the tasks.

The results are shown in Figure 9. The scores for question 2, satisfaction with the amount of time, are almost identical for each of the approaches as could be expected with the close to identical time consumptions. The results for the remaining satisfaction related questions, the ease and the support information, shows a clear tendency that the text interface is considered worst whereas the projector interface is considered best.

The average satisfaction scores across all three questions are listed in Table III. On a normalized scale, the user satisfaction was on average 2.3 percentage points better for projection when compared to monitor and 7.5 percentage points better when compared to text. Relatively, the monitor and text based methods scored respectively 16.3% and 52.3% worse than projection.

TABLE III

AVERAGE SATISFACTION SCORES

Method	Likert [1;7]	Normalized [0;100]
Projector	1.86	14.33
Monitor	2.00	16.67
Text	2.31	21.83

The test participants were given the possibility to provide additional qualitative feedback on each approach. Two issues were noted by several participants. For the text based approach, they had an overview on the current state of the overall task. It would be an advantage to also have this for the other approaches. Also for the text based approach, the participants knew what the robot was supposed to do. For the other approaches, it was only shown *when* the robot intended to interact and with *which* item. A richer visualization of the robot's current intentions as well as likely future would be advantageous.

VI. CONCLUSIONS

Collaborative robotics is a new and rapidly expanding field with many challenges and opportunities. In this paper, we propose to improve human-robot interaction with collaborative robots by projecting task information as well as the state and intentions of the robot into human environments. The environments that serve as canvas can include both workspaces that are shared between robots and humans, as well as objects that are tracked in real-time. The primary contribution of this paper is a object-aware, projection approach for human-robot interaction. As a secondary contribution, we identify a reduction of the computer vision problem which allows us to propose a working solution for simultaneous tracking and projection. In our experiments, we show that the approach robustly tracks both simple objects, e.g., cuboids, as well as highly complex objects, e.g., car doors.

The usability of the projection-based interaction approach is evaluated in a comparative user study against two common interaction approaches; one displaying information on a monitor and one with information as printed text. Evaluation is performed according to ISO 9241-11 (1998) [16] as a combination of efficiency, effectiveness, and satisfaction. The efficiency, measured as the time consumption for carrying out tasks, is similar for all methods. The effectiveness, measured as the number of questions and errors from the test participants, is good for all approaches but slightly better for the projection based approach. The user satisfaction is on average best for the projection based approach. The test persons were particularly satisfied with the ease and supporting information of this approach. Usability is better for the projection based approach than for the other approaches. While not conclusive, these results does indicate a potential of projection based approaches in improving the interaction quality in human-robot interaction.

As future research, we intend to conduct additional experiments with a larger pool of participants in order to further investigate this issue and draw more statistically significant conclusions. A large scale comparative usability study with a more involved collaboration scenario could shed further light on the strengths and weaknesses of projection based interaction. Also, based on user feedback, even richer information could be provided as projections. This should include more precise information on what the robot intends to do, as well as indications on what the following actions of the robot are likely going to be.

ACKNOWLEDGMENT

The authors would like to thank Prof. Henrik I. Christensen and the Institute for Robotics & Intelligent Machines at the Georgia Institute of Technology for facilitating this joint work and for providing laboratory space for development and initial tests.

REFERENCES

[1] EUROP, "Robotics visions to 2020 and beyond - the strategic research agenda for robotics in europe (SRA)," EUROP, Tech. Rep., 2009.

[2] K. Ishii, S. Zhao, M. Inami, T. Igarashi, and M. Imai, "Designing laser gesture interface for robot control," in *Human-Computer Interaction–INTERACT 2009*. Springer, 2009, pp. 479–492.

[3] S. Omidshafiei, A.-A. Agha-Mohammadi, Y. F. Chen, N. K. Ure, J. P. How, J. Vian, and R. Surati, "Mar-cps: Measurable augmented reality for prototyping cyber-physical systems," 2015.

[4] F. Ghiringhelli, J. Guzzi, G. Di Caro, V. Caglioti, L. M. Gambardella, A. Giusti *et al.*, "Interactive augmented reality for understanding and analyzing multi-robot systems," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 1195–1201.

[5] F. Leutert, C. Herrmann, and K. Schilling, "A spatial augmented reality system for intuitive display of robotic data," in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2013, pp. 179–180.

[6] R. T. Chadalavada, H. Andreasson, R. Krug, and A. J. Lilienthal, "That's on my mind! robot to human intention communication through on-board projection on shared floor space," 2015, presented at European Conference on Mobile Robots (ECMR).

[7] A. Dragan and S. Srinivasa, "Generating legible motion," in *Robotics: Science and Systems*, June 2013.

[8] A. Dragan, S. Bauman, J. Forlizzi, and S. Srinivasa, "Effects of robot motion on human-robot collaboration," in *Human-Robot Interaction*, March 2015.

[9] A. Dragan, R. Holladay , and S. Srinivasa, "Deceptive robot motion: Synthesis, analysis and experiments," *Autonomous Robots*, July 2015.

[10] L. Takayama, D. Dooley, and W. Ju, "Expressing thought: Improving robot readability with animation principles," in *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on*, March 2011, pp. 69–76.

[11] J. Mainprice, E. A. Sisbot, T. Siméon, and R. Alami, "Planning safe and legible hand-over motions for human-robot interaction," *IARP workshop on technical challenges for dependable robots in human environments*, vol. 2, no. 6, p. 7, 2010.

[12] F. Stulp, J. Grizou, B. Busch, and M. Lopes, "Facilitating Intention Prediction for Humans by Optimizing Robot Motions," in *International Conference on Intelligent Robots and Systems (IROS)*, 2015.

[13] C. Choi, H. Christensen *et al.*, "Real-time 3d model-based tracking using edge and keypoint features for robotic manipulation," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4048–4055.

[14] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 679–698, 1986.

[15] C. Choi and H. I. Christensen, "Robust 3d visual tracking using particle filtering on the special euclidean group: A combined approach of keypoint and edge features," *The International Journal of Robotics Research*, vol. 31, no. 4, pp. 498–519, 2012.

[16] W. ISO, "9241-11. ergonomic requirements for office work with visual display terminals (vdts)," *The international organization for standardization*, 1998.

[17] J. R. Lewis, "Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the asq," *ACM SIGCHI Bulletin*, vol. 23, no. 1, pp. 78–81, 1991.