Transparent Robot Behavior Using Augmented Reality in Close Human-Robot Interaction

Gabriele Bolano[1], Christian Juelg[1], Arne Roennau[1], Ruediger Dillmann[1]

***Abstract*— Most robots consistently repeat their motion without changes in a precise and consistent manner. But nowadays there are also robots able to dynamically change their motion and plan according to the people and environment that surround them. Furthermore, they are able to interact with humans and cooperate with them. With no information about the robot targets and intentions, the user feels uncomfortable even with a safe robot. In close human-robot collaboration, it is very important to make the user able to understand the robot intentions in a quick and intuitive way. In this work we have developed a system to use augmented reality to project directly into the workspace useful information. The robot intuitively shows its planned motion and task state. The AR module interacts with a vision system in order to display the changes in the workspace in a dynamic way. The representation of information about possible collisions and changes of plan allows the human to have a more comfortable and efficient interaction with the robot. The system is evaluated in different setups.**

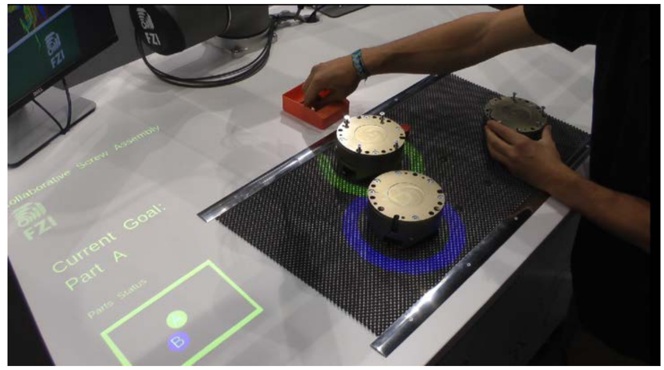


Fig. 1: Augmented information on robot behavior and detected parts visualized with light projection.

I. Introduction

In the early stages of robotics, robots had to operate at safety distance from the human operators or enclosed by fences. During the last years they have gained more and more safety capabilities, in order to sense the environment and dynamically react to changes in their surroundings. Thanks to the research and the achievements in this area, nowadays robots are able to safely work close to humans. The use of 3D sensors and other devices allows them to observe and predict the user's motions and adapt their behavior based on this information.

In order to achieve a good collaboration it is really important that the members of the team know and understand the intention of each other. Especially, in close Human-Robot Collaboration (HRC), the interaction between the participants is a major point. Since they have to work close to each other, the communication of their intentions needs to be as fast and intuitive as possible. To achieve this, several works have studied how the robot can understand the human intentions. Different implicit cues were investigated in order to make it aware of the human behavior and able to foresee the user's next motion.

Although much research has been conducted in this area, the communication of the robot's intentions is an issue that has not yet been thoroughly investigated. This is mainly because they usually cannot provide implicit cues on their intention like those humans generally use during their interaction, such as gaze, facial expressions or gestures. Nevertheless, as highlighted in previous works, it is also really important that the user has awareness and a clear understanding of the robot intentions. Even if the robot is able to dynamically change its motion in order to avoid collisions with the worker, this latter can still have a feeling of anxiety and discomfort. Without any information about the planned robot motion, he is also not able to decide in an intelligent way how to avoid repeated disruption of the robot's current plan. Because of this, the system could be forced to replan repeatedly, inhibiting the progress of the robot's tasks.

Our previous work showed that the representation of the current robot's plan helps the user to achieve a better comfort in the interaction [1]. This was done through the use of acoustic and visual feedback, provided to the human by audio speakers and monitors. However, this solution has some drawbacks and is still not ideal. The use of screens to display the robot's intentions causes the human worker to diverge the attention from his task to these external devices. In addition, the information was shown in a virtual representation of the environment, requiring the user to link it with the real world. This is at the expense of the intuitiveness of the system. Indeed, in close collaboration it is important that the user is able to immediately understand the robot intentions without being distracted from his task.

Due to the possibility of representing virtual information in the real world, augmented reality systems are becoming popular in many applications and also in robotics. Using a video projector to display internal information directly on the worktable, the user is able to visualize and understand immediately the state of the task and the robot's intentions, without the need to look at external devices such as monitors.

The use of AR can help the human worker to quickly

[1]The authors are with FZI Research Center for Information Technology, Haid-und-Neu-Str. 10-14, 76131 Karlsruhe, Germany {bolano, juelg, roennau, dillmann}@fzi.de

understand if the robot has to change its motion, making him aware of its current goal. This is useful in order to reduce the risk of blocking the new robot plan. It can also help the user to detect the status of the parts by highlighting the ones that need human attention.

The structure of this paper is as follows. In Section II, the related work on human-robot interaction and augmented reality is presented. In Section III we describe the system we developed to allows an intuitive representation of robot information using projector-based augmented reality. In Section IV, the results are reported and evaluated. Finally, we provide conclusions and future perspectives in Section V.

II. Related Work

Human-Robot Interaction (HRI) is a topic that has gained a lot of research interest in the last years. Most of the work is focused on making the robots more intelligent and able to take autonomous decision in highly dynamic environments [2], [3]. To achieve this goal, a lot of work has been done in the field of object and human recognition. In order to understand and anticipate the user behavior, research has focused on the prediction of the human intentions [4]. Mainprice et al. propose a system that generates the prediction of human workspace occupancy [5]. This is done by computing the swept volume [6] of learned human motion trajectories. Other approaches track the user's head and gaze in order to locate his focus of attention and predict his intentions [7], [8]. The work from Nomura et al. investigates the relationship between people's emotions and their interaction with a robot [9]. A psychological scale for anxiety was used to predict the human behavior towards the robot. Coupete et al. studied a method to evaluate the acceptability of an operator to work with a robot and how to make the collaboration more natural [10]. The results showed that the use of virtual reality is a helpful tool to evaluate subjective notions on acceptability. A gesture recognition system was used to enable a more efficient and smooth collaboration. This was achieved selecting a set of gestures in order to help the robot to be synchronized with the human worker.

However, in a collaborative task, it is important that both actors have a clear idea of what the other one is doing [11]. Humans continuously use implicit cues to understand the intentions of other people when they work in a common workspace. Eye gaze, gestures, speech and other natural cues allow them to have an efficient and effortless cooperation. Robots usually don't provide this information, especially in the case of manipulators that don't have a human-like appearance. This can also lead to a feeling of anxiety for the user, due to the unpredictable robot motion. The work from Cha et al. explored the use of nonverbal communication signals that a robot could use in a human-robot collaboration [12]. The study focused on multi-modal light and sound signals to request help during a collaborative task.

In the research from Takayama et al., several techniques to improve robot readability are presented [13]. The focus of the study is the use of animation principles in order to illustrate forethought and reaction. The results showed that with this additional information people are more confident in their interpretations of the robot behavior. Furthermore the robot appears more appealing and approachable.

Our previous work has explored how to represent robot information to workers in a manufacturing human-robot shared workspace scenario [1]. Acoustic signals were used to draw the attention of the worker every time that the robot had to change its motion because of unexpected obstacles. The information about the new goal and planned trajectory was displayed on screens and represented in a virtual environment. The results showed that the additional information increased the comfort of the interaction, helping the users to reduce their feeling of anxiety [14], [9]. The drawback of these approaches is that they cause the workers to divert their attention to other devices, making them easily loose the focus on their task. The problem arises also because the human needs time to match the information on displays with the real world.

For this reason, a promising method to represent the robot information is the use of AR tools. This way the information can overlap directly with the real world objects, making the robot's plan understandable in a more intuitive way. This allows to easily display information and virtual markers exactly on the desired position. An early research from Birkfellner et al. explored the use of a head-mounted AR device for visualization in biomedical applications [15]. Other works exploit AR tools to support operators in an industrial workplace [16]. The AR is deployed to visualize the assembly process, production updates and video and text instructions. Fang et al. present the use of AR for an easy and intuitive way to program robots [17]. A virtual robot was used in a real environment in order to program and simulate the execution of the robot trajectory planning process. The user used a device to interact with the virtual model during the programming task. Li et al. proposes a system that combines an AR interface with programming by demonstration to facilitate the robot path planning [18]. The human was enabled to move the probe tip of a 5-DOF robot arm around a virtual object. The position of the tool was tracked in the 3D space by a depth camera. In the work from Green et al. an AR interaction is used to remotely operate a simulated mobile robot [19]. The system provided to the user spatial awareness in order to achieve the feeling of working in a real collaborative environment.

In literature we can find works that studied the representation of information for a mobile platform, in order to show the direction in which it will move in the near future [20], [21]. In these works a projector was mounted on top of a vehicle to cast different shapes on the floor. A similar research presents a study regarding the communication of intentions using a robotic wheelchair [22]. The results showed that the persons passing by the vehicle had a smoother motion with the use of the additional information. Leutert et al. proposed a system to project information directly into the user's workspace. The study was made with a fixed and a mobile projector mounted directly on the robot [23].

The motivation of this work is based on the lack of research on the representation of planned motion for a highly replan-capable robot. We believe that a more intuitive way to show this information can help the human to have a more comfortable interaction during close collaboration. Since the robot is able to change its motion based on the user's actions, the human has to be able to understand if the robot changes its target, as well as the workspace that it will occupy in the near future. In this work we contribute to this goal by representing the robot motion directly in the workspace. The use of a projector allowed us to display the information at the relevant position in the real world. In this way the user is able to perceive feedback from the robot and predict its motion without losing the focus on his task. This has been done in such a way to allow an intuitive understanding of this information.

III. Approach

The scenario considered in this work is a shared assembly workspace. The human and the robot have to share the same worktable, working close to each other. The use of the GPU-Voxels library enables the robot to predict collisions and replan its goal in real time, based on live sensor data from the environment. This information is used to perform collisions checking with the swept volume of the robot's trajectory. The library makes use of highly parallelized algorithms on the GPU in order to minimize the reaction time of the robot.

The aim of this work is to have a better understanding of the robot's behavior, making the user able to understand its planned trajectory and goal. An important information to show is the collision detection information. Sometimes changes of direction in the robot motion can happen, but the human worker can understand them wrongly as reaction to his movement. Without this information, the human could move to try to step away from the robot motion, sometimes instead actually moving into the robot trajectory and making the robot stop uselessly. Giving the operators the possibility to see potential collisions, they can understand why the robot stops and could change their motion in order to accommodate the robot's planned path.

In order to display the additional information in the workspace, we mounted a projector on top of the worktable. The overall setup used is reported in Fig. 2. The information that is shown includes the status of the parts and the current target of the robot. The users are able to see immediately the information, because it is visualized in the position occupied by the objects in the real world. This way they can easily understand the robot's intentions and the available parts on the worktable without the need to divert their attention to external devices.

In our scenario, the parts position and rotation is detected automatically by a camera system positioned on top of the worktable. In Fig. 3 the system architecture diagram is reported.

The parts detected by the camera system are highlighted with colored circle markers around them. They can be placed arbitrarily within the specified area and the projected markers are immediately updated accordingly to the detected position.

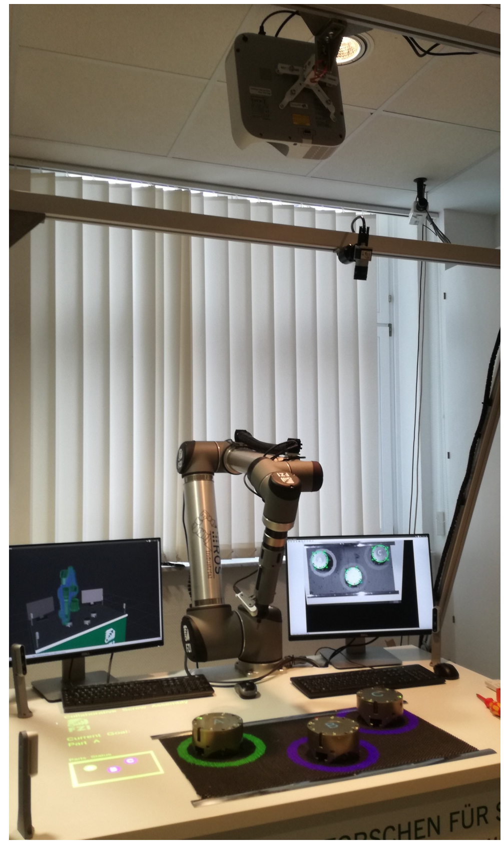


Fig. 2: The overall setup used in this work. A projector mounted on top of the worktable is used to dynamically cast information about the parts and robot planned motion.

This way the user can see the ones that the robot has recognized in the workspace and understand immediately if some of them are not detected or if their position is not updated after a manual change. This feedback is important to make the human aware of the actual status of the vision system. In this way he can detect instantly if the robot has issues in the localization of the parts placed into the

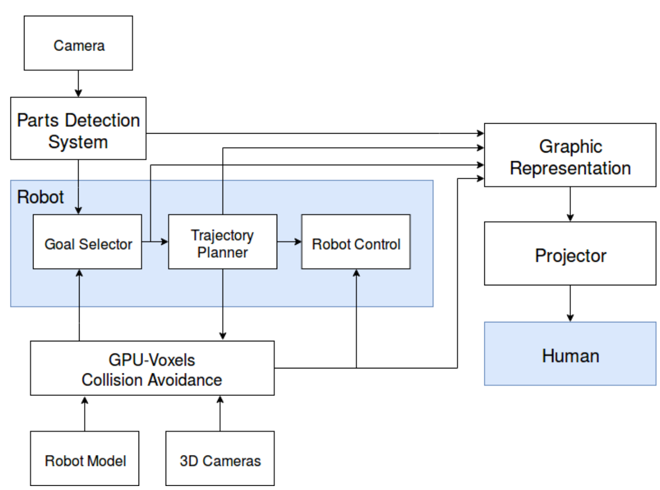


Fig. 3: Diagram of the system architecture. The communication between all the components used is represented.

Fig. 4: Text information on the side of the table allow the human to understand the status of the part and the robot goal even if there are obstruction in the workspace.

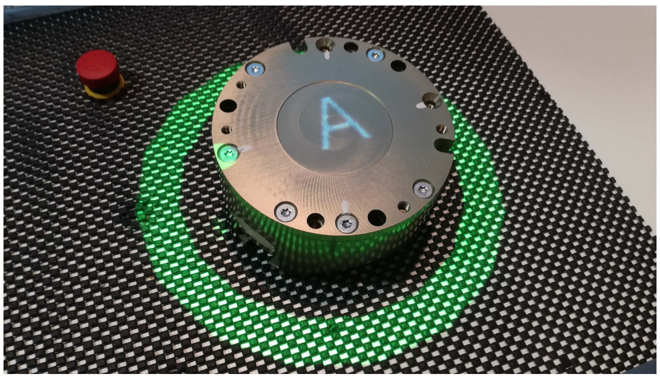


Fig. 5: The current goal of the robot is highlighted with a green marker around the part. The target position of the robot tool is also pointed with a green circle.

shared workspace. Once a part is available to be worked, a circular purple marker is displayed around it. The current status of them is also represented in one small area on the side of the table. In this way the user can see the needed information even if the workspace is completely occupied by the robot and in case of possible projector occlusions. Text information is also displayed to make the human understand the robot actual goal and how this is represented. Fig. 4 shows this information, in the case of three detected parts on the worktable. Furthermore, they are identified with an identification letter that is assigned dynamically and projected on their center. The robot is also able to detect defects in the parts, which need then to be inspected by the human operator. To visualize this request of intervention, the ones involved are then highlighted with a blinking marker. The current target of the robot is represented with a green circle around the related part. The target position of the robot tool is also marked in green as shown in Fig. 5.

We added also additional information to enable the human to detect easily if the robot stops because of a possible collision. In fact, when the human or another obstacle obstruct the planned robot motion, the AR system casts markers to represent this situation. The previous goal marker is then colored with red and additional text on the side explains the current state. Fig. 6 shows the situation in which the robot

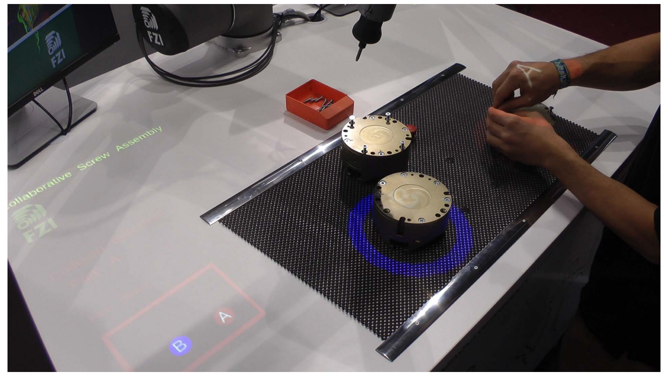


Fig. 6: A red marker on the planned goal of the robot and text information on the side of the table informs the user about the prediction of a possible collision with the highlighted part.

stops in order to avoid a possible collision with the human operator. In this way, the user can have a clear idea when the robot is changing its motion and he can intuitively see the information that allows him to understand the robot intentions and movements, with a consequential improvement of the ergonomics. This leads also to an increase of efficiency. When the human workers know the robot motion and goal, they can decide which part to work on in a more intelligent way, avoiding continuous replanning of the robot's trajectory.

For example Fig. 7 shows a situation in which the user

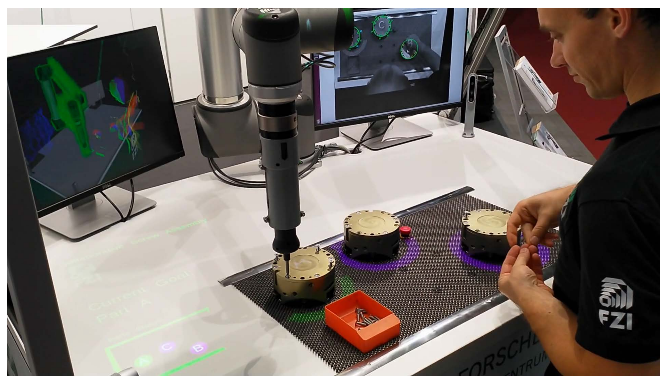
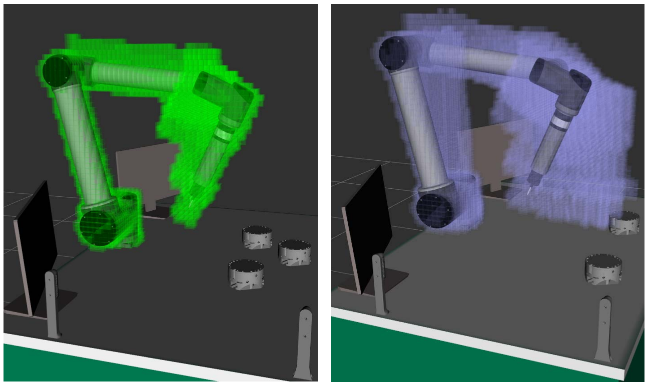


Fig. 7: The user inspects the workspace and clearly notices the current target of the robot.

Fig. 8: The robot moves towards the user, but he can see that the target of the robot is still on the same part. In this way he can continue to work without worrying about the robot motion.



(a) Robot collision model. (b) Trajectory swept volume.

Fig. 9: In (a) the discretized robot collision model is represented in green. This is used to render the trajectories swept volume (b) which is adopted for the collision checks against the live environment. The voxel representation is overlapped with the 3D mesh of the robot.

can work on two parts while the robot works on a third one. Once the human starts to work on one of them, he can easily see that this one is not selected by the robot.

Even if the robot moves suddenly towards his direction, the use of the AR information helps him to understand that the robot's goal it is still located on the same part. Fig. 8 shows this situation in which the robot moves towards the human. The user is able to easily see the AR information, which tells him that the robot is still working on the same part.

Another interesting information which helps the user to understand the robot's plan is the swept volume of the planned trajectory. This is the volume that the entire robot will occupy in its motion towards the current target and it is the one used by the collision checking system. The visualization of this information makes it clear where the human could cause the robot to stop because of a possible collision. This data is provided by the GPU-Voxels library that is used to compute the swept volume for each new trajectory. It is computed by adding the voxel representation of the robot model for each way-point of a trajectory. Fig. 9 shows in green the voxel representation of the robot collision model used for the computation of the swept volume of the robot's trajectories. The blue voxels represent the volume that the robot will occupy in the execution of the current path. During the execution of the current trajectory, the rendered swept volume of the remaining way-points is checked against the live environment camera data for possible collisions.

However, this volume information is based on 3D voxels which cannot be directly represented with the projector setup used in this work. For this reason we have implemented a 2D representation of this volume. The 2D projection of the voxels is rendered onto the table. This allows the user to understand which area of the worktable will be collision free in the next future. This is a useful information if combined with the goal markers. The human can understand the target of the robot and have an idea of its planned motion. In Fig. 10 and Fig. 11 we can see in blue the projection of the swept

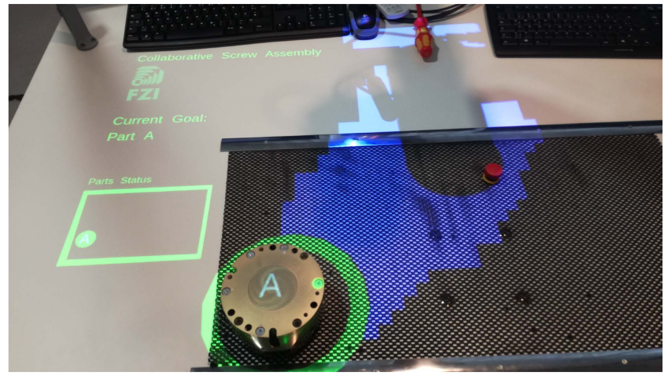


Fig. 10: Projection of the swept volume of the planned robot trajectory. This information enables the user to understand which area of the worktable the robot is planning to move to and the collision free zones.

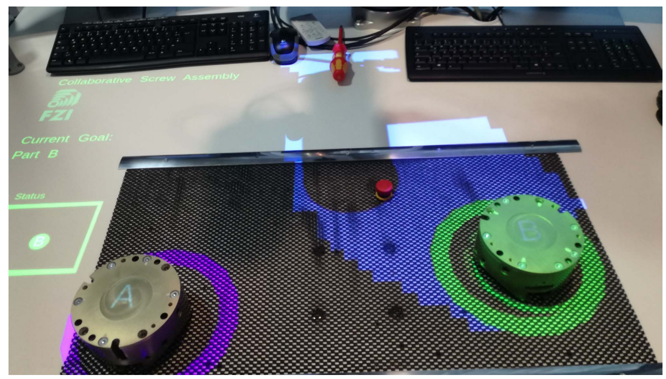


Fig. 11: The robot changes its target and the swept volume visualization is updated accordingly. This way the user is aware of the new collision free areas.

volumes for two different robot targets. Looking directly on the worktable the users can understand in which area of the worktable they can move freely without obstructing the robot's motion.

Another information that we added to the AR system is the end-effector path. This represents the future positions of the robot TCP in the execution of the remaining trajectory.

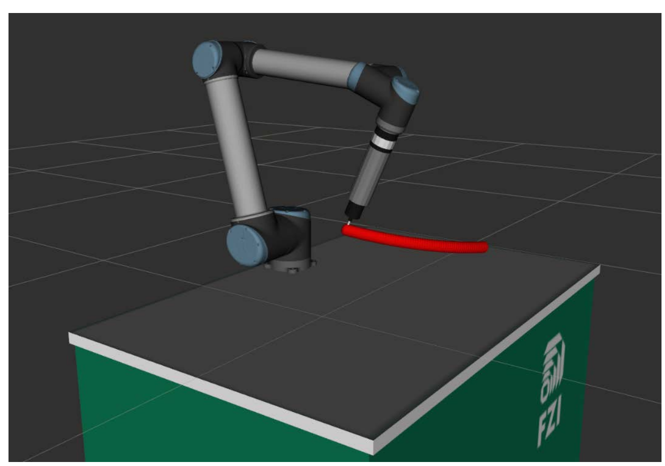


Fig. 12: Representation of the robot end-effector trajectory in the virtual environment. This information help to understand the path of the robot and its goal.

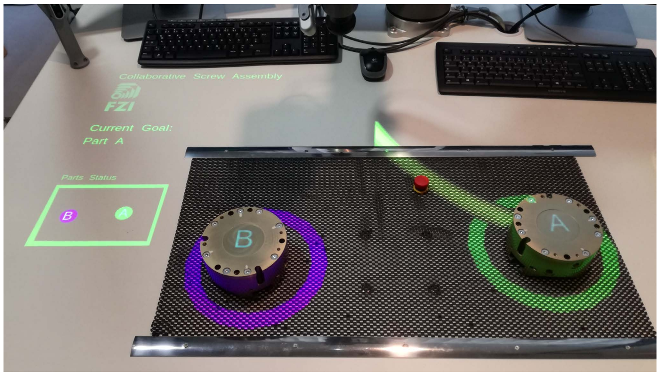


Fig. 13: Projection of the planned end-effector trajectory. The user can thus recognize the path of the TCP towards the current robot goal.

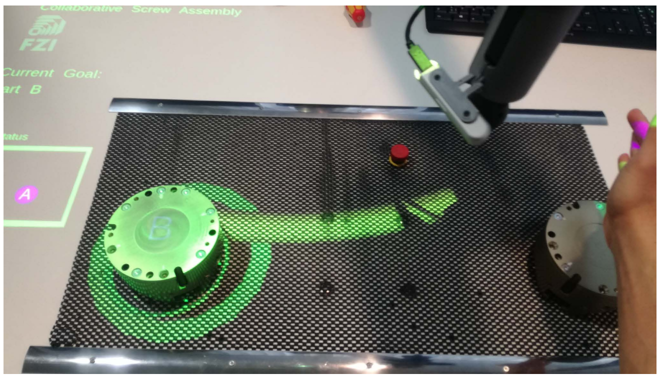


Fig. 14: A possible collision is detected and the robot replans its motion to another goal. The end-effector trajectory is updated immediately to show the user the new robot motion.

The projection of these points can be easily rendered on the worktable, giving the user a clear idea of the end-effector path that the robot will execute. It also provides useful information to recognize which part the robot end-effector is aiming for.

The way-points of the robot trajectory are interpolated in order to have a smooth representation of the motion as the one executed from the robot. These positions, represented in joint space, are then converted into Cartesian space through the computation of the inverse kinematics of each point. Fig. 12 shows the 3D representation of the TCP trajectory in the virtual visualization of the workspace. While the robot is moving, the segment of end-effector path already executed is updated and removed from the visualization. In this way the user can visualize only the remaining trajectory that the robot still has to execute.

In Fig. 13 we can see the representation of the end-effector path on the worktable. Fig. 14 shows the change of trajectory after the detection of a possible collision with the human. Because of that the robot changes its target and the path to new goal is updated accordingly.

IV. Evaluation

The system was presented at the Motek 2018 fair in Stuttgart, Germany. Over four days many visitors tested our projector based AR system. The feedback collected showed us that the additional information related to the robot target and motion were intuitive and easily understandable. The use of a video projector to dynamically display the information related to the entire system made them able to figure out the meaning of the projected markers by simply interacting with the parts and the robot.

The representation of the robot goal was very helpful in recognizing the workpiece that the robot was aiming for. This enabled the users to have more comfort and confidence in the interaction and decreased the feeling of anxiety caused by the lack of information from the robot. The information related to the collision avoidance system allowed the users to perceive when the robot had to change its motion in order to avoid a possible collision. This helped in making them aware of their influence on the robot motion and plan.

The visualization of the robot swept volume on the worktable turned out to be useful for understanding quickly the area in which the robot was not planning to move during the execution of its current motion. Thus the user was able to easily detect collision free zones in the workspace. The TCP trajectory representation turned out to be more suitable for the interpretation of the end-effector path, which is very relevant for some types of tools.

V. Conclusions

The AR system has improved the human-robot interaction in a shared workspace with a fast replanning capable robot. The use of a projector to display the robot motion information directly on the worktable helped the user to understand the robot's plan in a faster and more intuitive way. Using this system, the human workers feel more comfortable and the acceptance is improved.

Future works will explore the use of wearable devices for augmented reality, for example glasses like the HoloLens. In this way the 3D robot information could be represented in the 3D space without the need of projections onto surfaces in the workspace. The swept volume representation in particular could be displayed to give exact information on where the user would be colliding with the current robot trajectory. Further research will investigate how the human can react to this information in an immediate and intuitive way. Using gestures and speech, the user could interactively agree with the new planned motion or force the robot to find another path to the desired goal.

Acknowledgment

This research was funded in part by the European Union's Horizon 2020 program under grant agreement No 680734.

References

[1] G. Bolano, A. Roennau, and R. Dillmann, "Transparent Robot Behavior by Adding Intuitive Visual and Acoustic Feedback to Motion Replanning," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2018, pp. 1075–1080.

[2] A. Hermann, F. Mauch, K. Fischnaller, S. Klemm, A. Roennau, and R. Dillmann, "Anticipate your surroundings: Predictive collision detection between dynamic obstacles and planned robot trajectories on the GPU," in *2015 European Conference on Mobile Robots, ECMR 2015 - Proceedings*, 2015.

[3] C. Juelg, A. Hermann, A. Roennau, and R. Dillmann, "Efficient, Collaborative Screw Assembly in a Shared Workspace," in *Intelligent Autonomous Systems 15*. Springer International Publishing, 2018, pp. 837–848.

[4] A. Turnwald, D. Althoff, D. Wollherr, and M. Buss, "Understanding Human Avoidance Behavior: Interaction-Aware Decision Making Based on Game Theory," in *International Journal of Social Robotics*, vol. 8, no. 2. Springer Netherlands, 2016, pp. 331–351.

[5] J. Mainprice and D. Berenson, "Human-robot collaborative manipulation planning using early prediction of human motion," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 299–306.

[6] W. P. Wang, G. Electric, and K. K. Wang, "Geometric Modeling for Swept Volume of Moving Solids," vol. M, 1986.

[7] M. M. Moniri, F. A. E. Valcarcel, D. Merkel, and D. Sonntag, "Human gaze and focus-of-attention in dual reality human-robot collaboration," in *Proceedings - 12th International Conference on Intelligent Environments, IE 2016*, 2016, pp. 238–241.

[8] O. Palinko, F. Rea, G. Sandini, and A. Sciutti, "A Robot Reading Human Gaze : Why Eye Tracking Is Better Than Head Tracking for Human - Robot Collaboration," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 5048–5054.

[9] T. Nomura, T. Suzuki, T. Kanda, and K. Kato, "Measurement of Anxiety toward Robots," 2006, pp. 372–377.

[10] E. Coupeté, V. Weistroffer, O. Hugues, F. Moutarde, S. Manitsaris, and P. Fuchs, "New Challenges for Human-Robot Collaboration in an Industrial Context: Acceptability and Natural Collaboration," in *Fifth workshop "towards a Framework for Joint Action", IEEE RO-MAN 2016*, 2016, pp. 1–4.

[11] B. Mutlu, A. Terrell, and C.-m. Huang, "Coordination Mechanisms in Human-Robot Collaboration," in *International Conference on Human-Robot Interaction - Workshop on Collaborative Manipulation*, 2013, pp. 1–6.

[12] E. Cha, M. Matarić, and T. Fong, "Nonverbal signaling for non-humanoid robots during human-robot collaboration," in *ACM/IEEE International Conference on Human-Robot Interaction*, vol. 2016-April, 2016, pp. 601–602.

[13] L. Takayama, M. Park, D. Dooley, and W. Ju, "Expressing Thought: Improving Robot Readability with Animation Principles," in *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2011, pp. 69–76.

[14] D. Kulic and E. Croft, "Anxiety Detection during Human-Robot Interaction," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005.

[15] W. Birkfellner, M. Figl, K. Huber, F. Watzinger, F. Wanschitz, J. Hummel, R. Hanel, W. Greimel, P. Homolka, R. Ewers, and H. Bergmann, "A head-mounted operating binocular for augmented reality visualization in medicine - Design and initial evaluation," in *IEEE Transactions on Medical Imaging*, vol. 21, no. 8. IEEE, 2002, pp. 991–997.

[16] G. Michalos, P. Karagiannis, S. Makris, Ö. Tokçalar, and G. Chryssolouris, "Augmented Reality (AR) Applications for Supporting Human-robot Interactive Cooperation," in *Procedia CIRP*, vol. 41. Elsevier B.V., 2016, pp. 370–375.

[17] H. Fang, S. K. Ong, and A. Y.-C. Nee, "Robot Programming Using Augmented Reality," in *2009 International Conference on CyberWorlds*, 2009, pp. 13–20.

[18] P.-c. Li and C.-h. Chu, "Augmented Reality based Robot Path Planning for Programming by Demonstration Augmented Reality based Robot Path Planning for Programming by Demonstration," no. December, 2016, pp. 1–5.

[19] S. A. Green, J. G. Chase, X. Chen, and M. Billinghurst, "Evaluating the Augmented Reality Human-Robot Collaboration System," 2008, pp. 2–4.

[20] E. Bunz, R. Chadalavada, H. Andreasson, R. Krug, M. Schindler, and A. J. Lilienthal, "Spatial Augmented Reality and Eye Tracking for Evaluating Human Robot Interaction," 2016.

[21] 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," in *2015 European Conference on Mobile Robots (ECMR)*, no. 1. IEEE, 2015, pp. 1–6.

[22] A. Watanabe, T. Ikeda, Y. Morales, K. Shinozawa, T. Miyashita, and N. Hagita, "Communicating Robotic Navigational Intentions," 2015, pp. 5763–5769.

[23] F. Leutert, C. Herrmann, and K. Schilling, "A Spatial Augmented Reality System for Intuitive Display of Robotic Data," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2013, pp. 179–180.