2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

A Projection-based Sensor System for Safe Physical Human-Robot Collaboration

Christian Vogel, Christoph Walter and Norbert Elkmann

Abstract— **This paper presents the application of a novel projection-based safety system for ensuring hard safety in human-robot collaboration. We adapted the proposed sensor system to incorporate the joint positions and velocities of a collaborative robot, thus offering the opportunity to establish minimal and well-shaped safety spaces around the robot at any time. In this contribution we explain in detail main challenges and their solutions for generating and monitoring such safety spaces. Furthermore, we build up a future collaborative workplace to demonstrate the practicability of the system under operational conditions.**

I. Introduction

In the area of industrial manufacturing and production robots are still enclosed by static systems ensuring the safety of humans. Fences, light barriers or pressure sensitive mats guarantee safety but simultaneously prevent a coexistence and collaboration of human and robot in close proximity. But this is a main prerequisite for future flexible work cells that include automation as well as manual tasks. So, in research and academia the development of novel safety systems, which enable humans to interact with robots while guaranteeing safety, is a main topic.

In general we can differentiate between approaches that prevent humans from potential collisions with the robot and approaches that deal with different reaction strategies at collision time. Here, especially developments concerning intrinsically safe robots like the DLR KUKA LWR or touch-sensitive sensors [1] [2] used as an artificial skin that encloses parts of the robot are well known. Other systems that are applied to the robot can measure the distance to unknown objects [3][4]. While these systems are applicable for small or mid-sized industrial robots, collision avoidance by optical sensor systems is more general. Multi-camera setups in various configurations and combinations, i.e. multi time-of-flight cameras [6], time-of-flight and stereo cameras [7], multi cameras [5], are used to estimate the minimal distance between a human and a robot, thus leading to a defined reaction of the robot to avoid potential collisions.

In [8] we presented our novel projection-based sensor system and discussed its usage as a safety system in human-robot collaboration scenarios. We argued that our system has main advantages over other current research activities because of its intrinsic safety, dynamic adaption of safety spaces and reduced influence of environmental conditions. Beside the aspect of "Hard-Safety", the system also concerns additional requirements regarding the availability by its "Soft-Safety" features like the visibility of safety space boundaries and the visualization of intended robot movements. Furthermore, the system is capable of supporting the user by additional functionalities like information display and augmented reality for situation awareness, interaction, as well as 3D scene capturing [9].

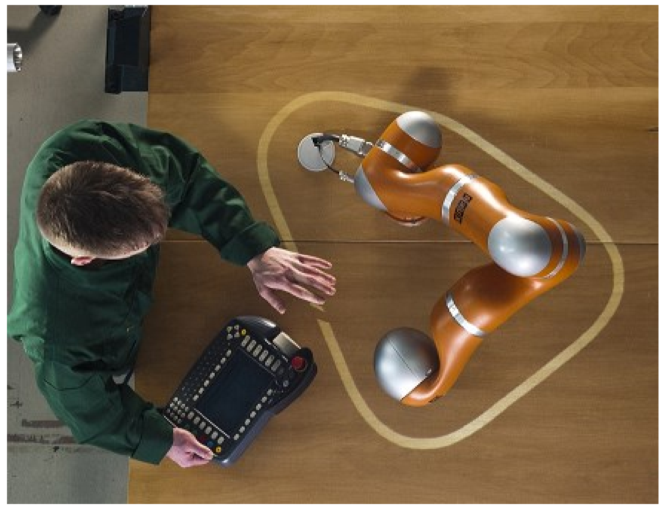

Figure 1: Dynamically established safety space (white line) enclosing an industrial robot.

In this contribution we present a method for monitoring a collaborative human-robot workplace and guaranteeing the safety of humans by using a novel projection-based safety system. We describe our approach of dynamically establishing a safety space around a robot by using its current joint positions and velocities. So, at any time the robot is enclosed by a minimal safety space and enables the user to utilize as much workspace as possible, as seen in Figure 1. For demonstrating usability and testing the system under operating conditions we built up a prototypical collaborative workplace, which consists of the proposed sensor system, additional electronics and workbenches equipped with an industrial lightweight robot.

As this contribution aims at explaining the enhancement of our proposed projection-based sensor system presented in [8], the knowledge of contents of this paper is mandatory.

In the following sections we firstly introduce some prerequisites like workplace setup and internal world representation. After that, we explain and discuss in section III the process of safety space generation and collision detection in section IV. Finally, we present recent insights to the evaluation of the safety system concerning its detection capabilities and limitations.

C. Vogel, C. Walter and N. Elkmann are with the Fraunhofer Institute for Factory Operation and Automation IFF, Magdeburg, Germany <first Name>.<last Name>@iff.fraunhofer.de

II. PREREQUISITES

As we need some prerequisites for explaining the algorithms of safety space generation (section III) and collision detection (section IV) we firstly introduce in this section particular fundamentals. In detail, we describe the prototypical collaborative workplace by specifying the used hardware and their adjustments. Furthermore, we introduce our model of internal world representation that is an elementary part of safety space generation and collision detection. We further give an overview of the entire process flow.

A. Workplace Setup

For testing and demonstrating usability and practicability of the proposed safety system we build up a prototypical collaborative human-robot workplace. This also allows the validation and evaluation of the developed algorithms under operational conditions.

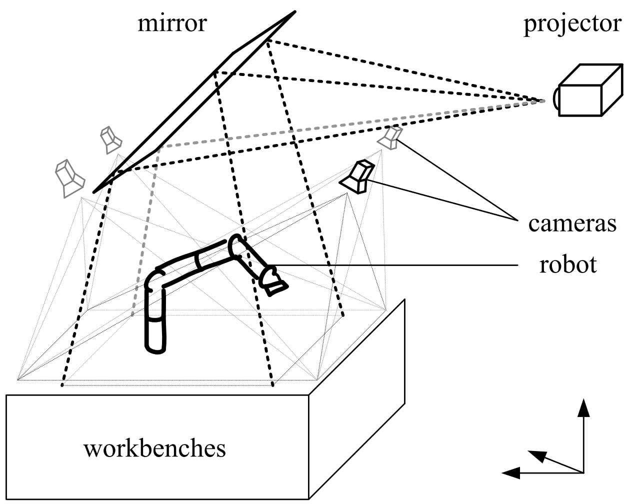

Figure 2: Schematic overview of the system setup consisting of sensor system, robot, mirror and workbenches.

The setup consists of a system carrier that is used to mount the sensor system and additional hardware and electronics. In detail, this includes the four monitoring cameras, which are installed in the four corners of the carrier at a height of 2.50 m. Every camera has a resolution of 640*512 pixels and acquires images with a frame rate of about 50 fps. For trigger purposes and synchronization we connected the cameras to the projector by additional electronics.

While the original optic of the monochrome DLP-projector (resolution: 800*600 pixels) was not able to establish the required large volumes of the safety spaces, we had to replace the wide angle lens by a lens with longer focal length, which leads to an increasing distance between projector and projection plane. We overcome the issue of a very high positioned projector by using a 45 degrees angled mirror on the top of the carrier. So, the emitted light of the projector, which was horizontally mounted on an arm outside of the system carrier, is reflected by the mirror and establishes the safety spaces on the workbenches, i.e. the projection plane. We further equipped the two workbenches that are placed under the carrier with an industrial lightweight robot (KUKA LWR 4+). As this robot is developed for future safe physical human-robot interaction, additional safety systems might not be necessary. However, we used this robot because of its availability, size and design but we also mention that any other industrial robot is just as well possible. A schematic overview of the system setup is depicted in Figure 2.

The four cameras and the projector are calibrated intrinsically and extrinsically relative to a common coordinate system positioned at the surface of the workbenches. Additionally, we measured and further refined by calibration the position and orientation of the robot.

B. Internal World Representation

The internal world representation is an abstract description of the environment with respect to a common coordinate system. This means, that we have to model all relevant facts, which are at least the poses of the cameras and projector as well as the shape of the projection plane.

As we want to establish a safety space that encloses the robot with respect to its current joint positions and velocities, we need an additional internal representation of the robot with appropriate shape, size and kinematics. We designed this internal robot model by using simple 3-dimensional primitives like cylinders, spheres and cubes. The benefit of these primitives is the fact that they are suitable for applying fast perspective transformations, how it is needed by our algorithms concerning safety space generation and collision detection. The specific internal robot model for the used industrial robot can be seen in Figure 3, visualized as red and wired primitives.

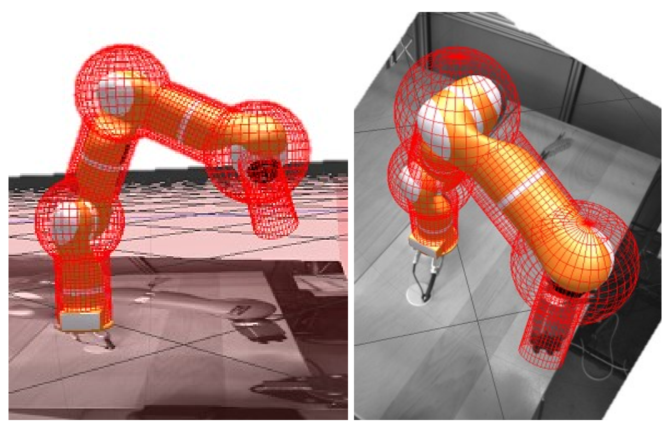

Figure 3: Internal robot model consisting of simple primitives like spheres, cylinders and cubes (red, wired).

As the size of the potentially generated safety space even depends on both the used tool and optional workpiece on it, we also have to consider them in the internal robot model. Instead of defining a certain primitive for every tool and workpiece it may be useful to identify a worst-case primitive that comprises all possible tools or workpieces. However, all primitives can dynamically be added to or removed from the internal representation at operation, thus influencing the resulting established safety space.

The internal world representation also offers the possibilities to consider additional objects in the world, which may influence the safety space generation process. These objects can be modeled as static primitives or can even be added dynamically to the internal world representation. However, for unknown objects we firstly have to identify their shape and size for approximating them by a primitive, which is recently an open task.

C. Process Flow

In the following section, we give an overview of the entire process flow, as depicted in Figure 4. We start by obtaining the current joint angles and velocities from the robot controller. These data are used for updating the internal world representation i.e. the internal robot model that forms the basis on the one hand for generating the projector image, and on the other hand to compute the virtual reference images, also called *expected-state mask* for every camera. Here, we also incorporate the intrinsics/ extrinsics of projector and camera as well as the safety space configuration that is based on the internal world representation. While the projector establishes the safety space by emitting the aforementioned projector image into the environment, it triggers the cameras for alternate imaging a light and a dark image, which are further used to compute a difference image and to determine the appropriate *current-state mask*. On the basis of the two masks a negative (no violation) or positive (violation) safety signal is generated concerning the resulting match or mismatch of both.

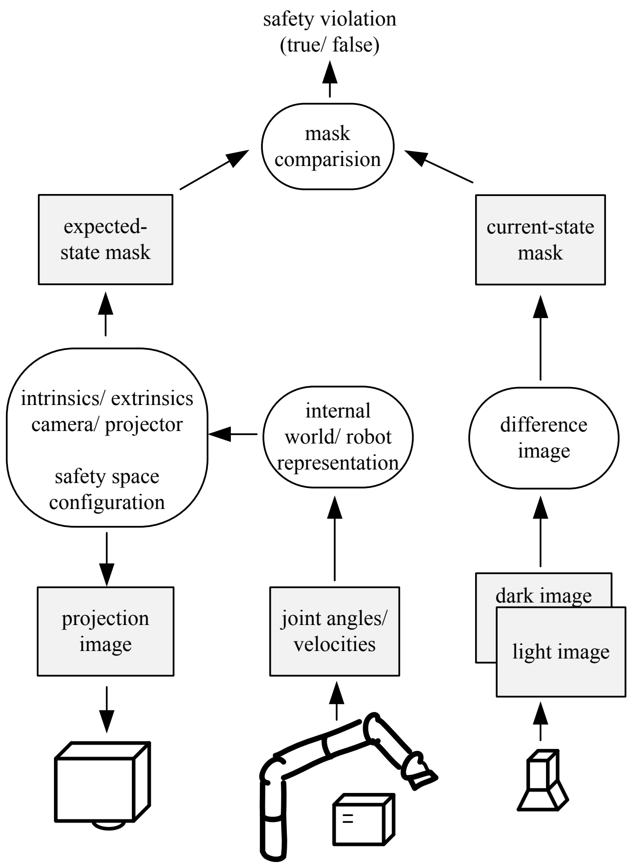

Figure 4: Schematic overview of the entire process flow.

III. Generation of Safety Spaces

In this section we explain in detail the single steps for generating safety spaces. As our sensor system is capable of establishing safety spaces by variably composing different shapes like dots, lines or planes, it also offers the possibility to generate a safety space shape that encloses the robot minimally. Obviously, we therefore need knowledge about the robot's size, shape, kinematics and the current joint positions. Therefore, we introduced the internal world representation that incorporates a model of the robot with an appropriate design (see section II.B).

A. Update the Internal World Representation

Since the generation of the safety space is based on the internal model we firstly have to update it. Therefore we adapt the positions and angles of the 3-dimensional primitives of the internal robot model by the current data of the robot controller, finally representing the current state of the real robot.

At this point it is also of utmost importance to consider the currently used tool and potential workpiece on it. While in most cases the tool (e.g. gripper) is static, its size, shape and position at the robot can be approximated by another primitive at the internal robot model. Concerning the workpiece, this may be different. Here, the currently used workpiece may be dependent on the single process steps and may differ in size, shape and position at the tool. To overcome this issue we identified the possible workpieces, approximated them by suitable primitives and added them dynamically to the internal robot model regarding the current process step. While the entire work process of the robot was implemented at the robot controller, the signal for adding or removing a workpiece to the internal robot model as well as the appropriate type of workpiece are provided by the robot controller.

Beside the adaptation of the internal robot model, at this point it is also possible to manage further objects, which should have an influence on the safety space generation process. At the end of this step we got an updated internal world representation that forms the basis for the safety space generation process.

B. Computation of Projector Image

The 3-dimensional internal robot model states an abstract representation of the real robot at certain times. As we sized the shapes of the primitives greater than the real sizes of the single parts of the robot, we can consider the internal robot model as a closure of the real robot. More precisely, it can be assumed as a "safety hull". So, by simply perspective transforming these 3-dimensional primitives to the 2-dimensional image plane of the projector, we got an intermediate image that contains the boundaries of all primitives represented as white lines on a black background. Therefore we need the updated internal model as well as the intrinsics and extrinsics of the projector. The resulting projector image that represents the final safety space was firstly determined by computing the overall contour of these boundaries. At this point we identified two issues. The first one concerns the primitives of the internal robot model building the shape of the safety space. As the radii of the cylinders are always less than the radii of the adjacent spheres, they never influence the resulting shape of the safety space. For this reason we simplified the internal robot model by dropping these primitives.

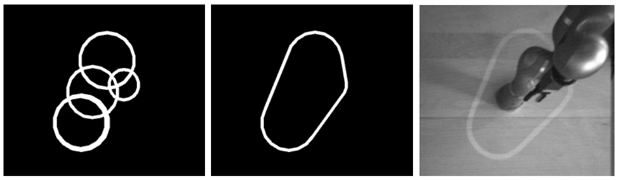

Figure 5: Zoomed view of intermediate (left), final projector image (mid) and camera image (right) after perspective transforming primitives of the 3D internal robot model to 2D projector image plane.

The second point concerns the visibility of the entire safety space to the monitoring cameras. Here, every single part of the safety space has to be seen by at least one camera. As this requirement could not be met under all circumstances, because of the occlusion of particular parts of the safety space by the robot itself, we implemented a second approach that computes the convex hull of the contours in the projector image. The resulting safety space is indeed in most cases bigger than the safety space of the first approach but the visibility of the projected lines to at least one camera is guaranteed. In Figure 5 we depicted exemplarily an intermediate projector image (left) computed on the basis of the simplified robot model as well as the corresponding final projector image (mid) based on the computation of the convex hull.

Actually, we generate safety spaces that are based on the robot's current state i.e. joint angles at certain time. As we also wanted to incorporate the velocities of the robot joints, we extended the generation of safety spaces by adapting the computation of the projector image to a two-stage process. At first we transform the updated internal robot model representing the current state of the robot to the 2D image plane as we described above. At the second step we compute the velocities of every robot joint and utilize them for approximating the joint's position after a predefined time. This time was determined by incorporating the reaction time and additional braking time. The estimated joint positions are used to update the internal robot model whose primitives are additionally transformed to the 2D image plane of the projector image. The final projector image is further determined by computing the convex hull of all these primitive boundaries, as depicted in Figure 6.

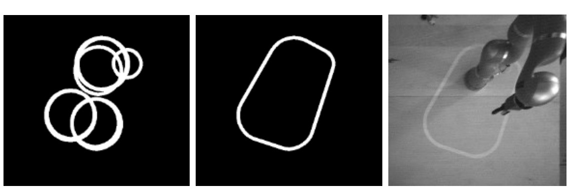

Figure 6: Zoomed view of intermediate (left), final projector image (mid) and camera image (right) by incorporating joint velocities.

C. Safety Space configuration

The projector image defines the safety space from the perspective of the projector. At this point there is no information about the shape, position or size of the safety space in world. But this is necessary for computing the virtual reference images i.e. the expected state mask for every camera (see section IV.C).

So, we compute the 3-dimensional representation of the safety space by using the intrinsics and extrinsics of the projector as well as the internal world representation. In detail, we determine for every pixel in the projector image that belongs to the safety space the corresponding light ray by using the intrinsics and extrinsics of the projector and intersect them with the projection plane defined by the internal world representation. The total of all these intersection points defines the safety space in world and will be stored as a list of 3D coordinates in the appropriate safety space configuration. Here, we also define the method of interpreting these coordinates to compound the safety space, e.g. as single dots or line strip, and specify additional information like line width.

IV. Collision Detection

As we described the generation of safety spaces in section III, we now explain the single steps to detect violations of them. A safety space violation occurs, if an object disrupts the emitted light rays of the projector, which represent the safety space. This disruption is detected by the surrounding monitoring cameras.

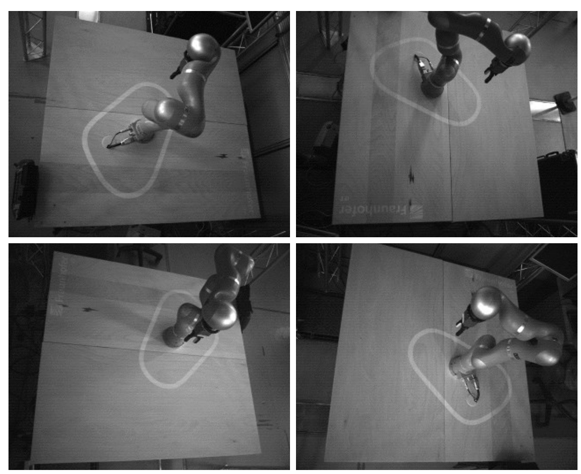

Figure 7: Images of the four monitoring cameras.

The cameras were mounted and adjusted in a way that they are capable of monitoring almost the whole area of the workplace but especially the single parts of the workplace directly in front of the cameras. By this, we can ensure that the composition of all camera images lead to an area-wide monitoring of the entire workplace at any time, although some parts of them are occluded by the robot, as can be seen in Figure 7. Here, we depicted the camera images of all four monitoring cameras. As the area-wide monitoring of the entire workplace is essential concerning safety, the overlapping of single observation parts by several cameras additionally increases the safety capabilities of the system.

A. Image Preprocessing

On the basis of the setup (see section II.A) and the appropriate perspective of the cameras (see Figure 7), it is obvious that projected lines close to the camera appear in the camera image with a wider line width than projected lines far away. As we further process the images by several morphological operations for extracting these lines, the varying line widths lead to weak results.

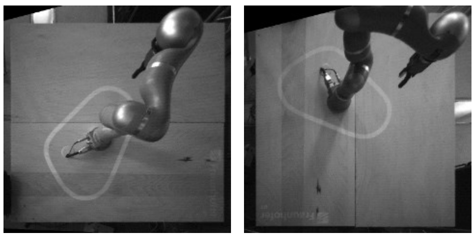

Figure 8: Resampled camera images from perspective of a virtual camera positioned in the center above the workplace.

Therefore, we firstly resample the camera images from perspective of a virtual camera positioned in the center above the workplace. Here, we defined a resolution of 800*800 pixels for this virtual camera, because this is a suitable trade-off between the detection capabilities (spatial resolution) and processing time of the collision detection process. The resampled camera images are further used for the processing of collision detection. Exemplarily, we depict two images in Figure 8.

B. Extraction of current-state mask

The first step for detecting a disruption of the safety-relevant emitted light-rays concerns the extraction of them in the camera image. Here, we compute pixel by pixel the absolute differences of consecutive camera images. While the image acquisition is synchronized with the light emission of the projector, consecutive images differentiate remarkably at pixel positions representing safety spaces. This difference image forms the basis for applying some additional morphological operations and final thresholding. Any pixel position in the resulting binary current-state mask that represents a safety space has a value of 1, the others 0. Figure 9 shows the extraction process for one camera.

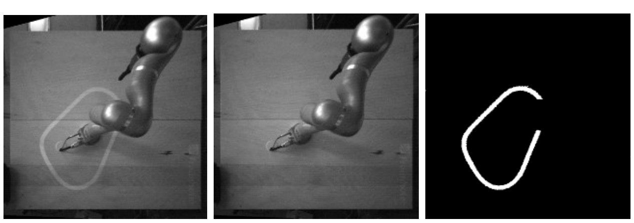

Figure 9: Computation of current-state mask (right) by differentiation of light (left) and dark (mid) image.

Actually, the cameras are triggered by a frequency of about 50 Hz for alternate image acquisition of a light (image with emitted light rays) and dark image (image without emitted light rays). The short time lag between imaging of these two images reduces the influence of changing light conditions to the extraction process.

C. Generation of expected-state mask

The current-state mask represents the pixel positions of the current safety space in the camera image. On the basis of this, it is not possible to decide if there exists a disrupted light ray i.e. violation of the safety space, or not. So, we need a reference camera image that determines the pixel positions of the safety-relevant light rays in an undisrupted state. This binary virtual reference image is computed on the basis of the intrinsics and extrinsics of the corresponding camera and projector and is further called expected-state mask. At all pixel positions in this virtual camera image a light ray of the projected safety space is expected, the pixel gets a value of 1, the others 0.

For determining the expected-state mask we take the current safety space configuration (see section III.C) into account and perspective transform the 3-dimensional coordinates of the safety space to 2D image plane of the camera by using the corresponding intrinsics and extrinsics. As the safety space configuration specifies just the parameters and properties of the safety space, at this point the resulting expected-state mask does not consider the geometry of objects in the workplace like the robot. This means, under certain circumstances, the expected-state mask defines pixel positions of light rays the camera is not able to see because of the occlusion by the robot. This depends on the position of the camera and the current kinematics of the robot. In Figure 10 we depicted a resampled camera image (left), the corresponding current-state mask (mid) and determined expected-state mask (right). It can be seen, that at some pixel positions light rays are expected, which are actually not visible for the camera.

Figure 10: Zoomed view of a resampled camera image (left), appropriate current-state mask (mid) and erroneous determined expected-state mask (right).

For determining the correct expected-state mask we also have to consider the internal robot model, which is used to eliminate such pixel positions of expected light rays in the expected-state mask that are not visible for the camera. Similar to the generation of safety spaces in section III, we transform the updated primitives of the model to the 2D image plane of the camera and compute a hull around the boundaries of these shapes. This hull represents the area in the camera image, at which the camera observes the real robot. Now, we eliminate all pixels representing expected light rays that are positioned in this area. The resulting final expected-state mask consists only of such pixel positions of expected light rays that are visible for the camera.

In general, for determining the expected-state mask it is a major task to consider the geometry of all objects positioned in a safety space.

D. Determine safety violation

For detecting a safety violation it is necessary to compare the expected-state mask against the current-state mask (see Figure 11). At every pixel position in the expected-state mask that contains a value of 1, the corresponding pixel in the current-state mask has to coincide. If there is a mismatch at a single position, a positive safety violation is signaled.

Figure 11: Zoomed view of expected-state mask (left), current-state mask (mid), resulting match of both (right).

V. Evaluation Results

In this section we want to present some remarkable results and findings of the evaluation of the proposed safety system according to its system behavior as well as detection capabilities and limitations. For this, we elaborated various test criteria based on the Technical Report 61496-4 and its subparts -2 and -3 that give us first insights to a future safety certification of the sensor system and its feasibility. All computation was done by our host system consisting of two Intel Xeon E5645 2.40 GHz processors and 12 GB memory.

One of the important characteristics of safety systems is the response time. Here, we consider the time from the moment an object intrudes the safety space to the moment the system generates the violation signal. Due to a camera frame-rate of about 50 fps, additional image processing and collision detection we identified a worst case response time of less than 80 ms. While incorporating the robot's joint velocities at safety space generation, we also have to consider this response time as well as the robot's latency and breaking distance, to establish sufficient safety spaces.

We further analyzed the safety system regarding its detection capabilities. Here, we firstly identified the minimal size of objects that are robustly detected by using cylinder-shaped test pieces with various diameters. However, as the system just detects the disruption of the safety space at the projection plane (i.e. the shadow of the object), this detection capability depends not only on the size of the object, but also on the height of the object above the projection plane. So, the task is to determine the smallest possible disruption width at the projection plane. On basis of our current system setup and parameterization of virtual cameras as well as image processing, the system is capable of robustly detecting disruptions of 15 mm. Due to the fact the system detects the disruption of the safety space at the projection plane and not the object itself, object properties like colors or textures do not influence the detection capabilities of the system. In contrast to this, full transparency and high reflectance (e.g. mirror) of objects can lead to erroneous violation signals. As these are shortcomings of camera-based systems in general, we recently have to avoid them in operation.

Moreover, we analyzed the system regarding changing environmental conditions. Firstly, we affected the system by strong vibrations and bumps. The reason for the resulting positive violation signal is the fact that projector and camera become decalibrated which results in a mismatch of the corresponding current-state mask and expected-state mask. Another point are changing light conditions and varying illumination intensities. The system is able to handle these influences until a certain limitation that is actually defined by several camera properties (exposure, gain, brightness) and parameterization of image processing. Here, also high-dynamic range modes have to be examined. But generally, if the light conditions are too poor the resulting weak current-state mask mismatches the expected-state mask and leads to a positive violation signal. In summary, the system remains safe at changing environmental conditions but the availability is decreased.

VI. Conclusion

In this paper we presented our method for dynamically generating and establishing safety spaces around a robot by using a projection-based safety system. We explained in detail the process of safety space generation by incorporating the robot's current joint positions and velocities. We further introduced an internal world representation that forms the basis for safety space generation and collision detection. The detection capabilities of the system were evaluated by several test criteria that give first insights to a future safety certification. Beside these first positive results, the next step is to analyze the system's behavior under real industrial environmental conditions.

Acknowledgment

This work was supported by the EU Project EC FP7-ICT-231143 ECHORD.

References

[1] P. Heiligensetzer, H. Wörn, "Safeguarded human-robot interaction in production"; In: Proc. of ISR, 2002

[2] Fritzsche M.; Elkmann N.; Grützner J.; Saenz J.: „A Tactile Sensor with Cushioning Elements for Enhanced Safety in Human-Robot Interaction", In: Robotics and Application (RA), 2010 15th IASTED International Conference on, Nov. 2010

[3] B. Ostermann, M. Huelke, A. Kahl, "Freed from fences - Safeguarding industrial robots with ultrasound"; To be published in: Proc. of 6th Working on Safety Conference, 2012

[4] Feddema J.T., Novak J.L.: "Whole Arm Obstacle Avoidance for Teleoperated Robots", In IEEE Robotics and Automation Proceedings, pp.3303 – 3309, 1994.

[5] Henrich D.; Gecks T.; "MULTI-CAMERA COLLISION DETECTION BETWEEN KNOWN AND UNKNWON OBJECTS". In: 2nd ACM/IEEE International Conference on Distributed Smart Cameras, ICDSC, Stanford, USA, 2008

[6] Fischer M.; Henrich D.: 3D Collision Detection for Industrial Robots and Unknown Obstacles using Multiple Depth Images", In: German Workshop on Robotics – GWR; Braunschweig, Germany, 2009

[7] Walter C.; Vogel C.; Elkmann N.: „A Stationary Sensor-System supporting Manipulators at Safe Human-Robot interaction"; In: Proceedings for the joint conference of ISR 2010 und ROBOTIK 2010, Munich, Germany, 2010

[8] Vogel C., Poggendorf M., Walter C., Elkmann N.: "Towards Safe Physical Human-Robot Collaboration: A Projection-based Safety System "; In: Proc. of IROS, 2011, San Francisco, USA, 2011

[9] Vogel C., Walter C., Elkmann N.: „Exploring the possibilities of supporting robot-assisted work places using projection-based sensor system"; In: Proc. of ROSE 2012, Magdeburg, Germany, 2012