

Human-Robot Interaction and Collaboration (HRI-C) Utilizing Top-View RGB-D Camera System

Tariq Tashtoush^{1*}, Luis Garcia², Gerardo Landa³
School of Engineering
Texas A&M International University
Laredo, TX, 78041 USA

Fernando Amor⁴, Agustin Nicolas Laborde⁵,
Damian Oliva⁶, Felix Safar⁷
Department of Science and Technology
National University of Quilmes
Quilmes, Argentina

Abstract—In this study, a smart and affordable system that utilizes an RGB-D camera to measure the exact position of an operator with respect to an adjacent robotic manipulator was developed. This developed technology was implemented in a simulated human operation in an automated manufacturing robot to achieve two goals; enhancing the safety measures around the robot by adding an affordable smart system for human detection and robot control and developing a system that will allow the between the human-robot collaboration to finish a predefined task. The system utilized an Xbox Kinect V2 sensor/camera and Scorbobot ER-V Plus to model and mimics the selected applications. To achieve these goals, a geometric model for the Scorbobot and Xbox Kinect V2 was developed, a robotics joint calibration was applied, an algorithm of background segmentation was utilized to detect the operator and a dynamic binary mask for the robot was implemented, and the efficiency of both systems based on the response time and localization error was analyzed. The first application of the Add-on Safety Device aims to monitor the working-space and control the robot to avoid any collisions when an operator enters or gets closer. This application will reduced and remove physical barriers around the robots, expand the physical work area, reduce the proximity limitations, and enhance the human-robots interaction (HRI) in an industrial environment while sustaining a low cost. The system was able to respond to human intrusion to prevent any collision within 500 ms on average, and it was found that the system's bottleneck was PC and robot inter-communication speed. The second application was developing a successful collaborative scenario between a robot and a human operator, where a robot will deposit an object on the operator's hand, mimicking a real-life human-robot collaboration (HRC) tasks. The system was able to detect the operator's hand and it's location then command the robot to place an object on the hand, the system was able to place the object within a mean error of 2.4 cm, and the limitation of this system was the internal variables and data transmitting speed between the robot controller and main computer. These results are encouraging and ongoing work aims to experiment with different operations and implement gesture detection in real-time collaboration tasks while keeping the human operator safe and predicting their behavior.

Keywords—Robotics manipulator; robot end-effector; computer vision; human-robot interaction (HRI); human-robot collaboration (HRC); robotics safety; scorbobot; Kinect; RGB camera; industrial system modeling; manufacturing systems design

I. INTRODUCTION

The demands and trends of the current market require enhanced manufacturing systems with reduced delivery times, mass production, and product customization, which impose a greater need for system flexibility and adaptability. Collaboration between humans and robots is considered a promising

technique to increase productivity and decrease the cost of production by combining both the robot's fast repetition and high production capabilities, and a human operator's ability to judge, react and plan. Collaborative robots (Co-bots) represent an evolution that can resolve a few challenges presented in the manufacturing and assembly environments. Co-bots allow physical interaction with humans within the work-space. Matheson and his team [1] described different ways a robot and an operator can work together, (1) *Co-existence*: the operator and robot are in the same work-space, but no interaction, (2) *Synchronized*: the operator and robots work within the same work-space, but at different times, (3) *Cooperation*: the operator and robots work together in the same work-space but have independent tasks, (4) *Collaboration*: the operator and robots work together to complete an assigned task. In a collaboration environment, it is important to note that any action will have immediate consequences for the other entity.

According to the International Standard ISO 10218 (1 and 2), and more extensively in Technical Specification ISO/TS 15066:2016, [2–5] four classes of safety requirements for collaborative robots are required:

- *Supervised stop*: The movement of the robot is stopped before an operator enters the collaborative work-space to interact with the robot and complete the desired task.
- *Manual guide*: The operator uses a manually operated device located on or near the robot's end-effector to transmit movement commands to the robot's system.
- *Monitoring speed and separation*: The robot and operator can move within the collaborative work-space simultaneously. The reduction of risk is achieved by always maintaining a distant separation between the operator and robot.
- *Power and force limitation*: Where the system must be designed to adequately reduce the risk for an operator by not exceeding the threshold as defined by the risk assessment.

Additionally, it is important to note that collaborative methods can be adopted even when using traditional robots. However, this requires the use of several and expensive safety devices such as laser sensors or visual systems. For these reasons, the team started to work on evaluating and developing affordable and accurate sensory systems that can measure the distance between the operator and the robot. This study utilizes

a lowcost RGBD camera to measure the position of an operator with respect to the robotic manipulator. While this configuration of specific measurement was utilized to track human beings [6], to our knowledge and based on the conducted literature review, it was not previously studied in the context of human-robot interaction and collaboration. Some researcher [7–15] analyzed the literature review and found that most of the RGB-D use was meant for human identification and tracking, human activity recognition, human behavior analysis for shopping and security purposes, intelligent health care systems, detecting defects in produce and animal recognition, also a data-based had been developed to summarize all these uses and algorithms. It was proven that the top-view RGB-D cameras can be utilized successfully in several applications where behaviors and interactions can be analyzed and they are very attractive due to their affordability and the sufficient information extracted from the provided pictures or live feed.

The paper is organized as follows: Section II is a literature review about robotics and their application in the industrial system, robotics safety regulations and standards, and collaboration and interaction between human and robots, Section III is a description of the robot and sensory system developed in this research, Section IV describes the followed methodology including the geometric model of the robot-sensor system, the process of calibration, and the detection of the operator, while in Section V, we evaluate the two methods of interaction between human and robot and reporting our findings and Section VI concludes the paper and describes the future plan.

II. LITERATURE REVIEW

The world has come to a point of many technological innovations where the presence and use of robotics are growing. Robots had been presented in manufacturing, hospitals, personal-use robots, service robots, etc. These robots aid the productivity of several tasks depending on their surrounding environment. In general, robotics could be used in many different settings where their intended purpose is to aid on a specific goal, complete a set of tasks that is difficult/tedious for a human to achieve, or simply make processes faster. Expedite services in systems such as Industrial/manufacturing, Health, or personal use, is a great enhancement to all current systems as their efficiency will increase. Therefore, safety standards are essentially required and must be implemented to achieve a safe operation of robotics in certain areas and or near human beings. Traditional robots have been separated from humans in workplaces trying to avoid any risk, injuries, or fatal incidents. This separation was implemented in the form of physical barricades or shut off robots whenever a human is present. However, technological improvements have shown great results where robotics no longer need to be separated and robots can be collaborative by working closely with humans, by developing new safety standards to design collaborative robotics to ensure humans' safety.

The Existence of robots in industrial settings enhances the production to meet the required demands while keeping the cost low. Robotics is considered as a flexible cell within a manufacturing line as they can be programmed to conduct different processes when needed. Safety is of utmost priority when designing robots and placing them in such environments and because of the rapid rise of robotics presence, safety

standards are to be frequently developed and improved to meet the new technology trends.

Few researchers and their teams [3, 16–21] discussed different industrial environments, the safety approaches that should be followed, and some real-life case studies. It was showed that lead designers must develop and evaluate safe, human-centered, ergonomics, and efficient collaborative assembly workstations, where the operator's feedback was provided in regards to occupational health and safety. Additionally, the Human Industrial Robot Collaboration (HIRC) workstation design process was evaluated through computer-based simulations based on the performance and safety characteristics such as Ergonomics, Operation time, operational costs, Maximum contact forces, and maximum energy density, this research illustrated how difficult is to evaluate safety and performance characteristics due to lack of physical workstations.

Parigi-Polverini [18], developed a new safety assessment tool "*Kinetostatic Safety Field*" which identifies *sources of danger* which could be an obstacle, human body part, or another robot link. The main advantage of this tool is the real-time applications and real-time collision avoidance with the use of a reactive control strategy. Another researcher suggested that robotics no longer need to be separated from humans, as robots can enforce safety by proposing a kinematic control strategy and maintain the robots' max level of productivity by reduced when humans are present in a working area.

Incorporating the industrial regulations such as the International Standard ISO 10218, Technical Specification ISO/TS 15066:2016, the American ANSI/RIA R15.06, the European EN 775 ISO 10218, and the national standards Spanish Association of Normalization and Certification, is the main procedure that is followed by manufacturing systems. These standards are outdated and have not been improved in the last five years, therefore some researchers introduced new concepts to cover techniques for estimation and evaluation of injuries focusing on various areas of the human body and the importance of developing new devices to detect impact, and minimize the human-robot impact.

Risk assessment is a crucial tool that must be used to enhance safety for both humans and robot systems. From literature review [22–27] discussed some history of operators and robots and how industrial robots have been evolved, differences between collaborative and non-collaborative robot cell safeguarding, voluntary industry consensus standards, and the risk assessment. Risk assessment should include quantitative head injury index for service robots as mechanical risk and incidents such as robot throws or drops and trapping and crushing are more to happen with such robots. Another proposed method to address safety in the human-Robot collaboration setting is Cooperative Collision Avoidance in dynamic environments [25]. This method computes a collision-free local motion for a short time horizon, which restricts the actuator motion but allows a smooth and safe control. Modeling human behavior and errors is another proposed method [28]. This formal verification methodology was developed to analyze the safety of collaborative robotic applications with a rich non-deterministic formal model of operator behaviors that captures the hazardous situations, which allows safety engineers to refine their designs until all plausible erroneous behaviors are considered and mitigated.

Other researchers [29–31] discussed different aspects of robotics design and their relationship to their safety ranking. Robot design principles should include robustness, fast reaction time, context awareness, energy, and power limitations. These principles will facilitate the following features as speech processing, vision processing, and robot control that also follow guidelines that will allow the robot to recognize speech, gestures, and correlations which eventually learns in the long run while also keeping humans safe. Predicting human behaviors, collision avoidance, collision reduction by data analysis, collisions reduction by design, perceptions affecting design, boundaries, sensors, adaptability to the surrounding environment, path planning, statistical probability, and robotic decision making are some of the safeguards that can be implemented in a high speeds and payload levels industrial settings.

III. SYSTEM DETAILS AND SETUP

The system is developed based on available educational and off-the-shelf components to model real-life robotics tasks, which are explained below.

A. Robotic Manipulator

The robotic manipulator selected for this project was the Scorbot ER-V Plus show in Fig. 1. This robot has five degrees of freedom, the Fig. 2 shows the length of the links and the degree of rotation and operation range determining the work-space of the robot. The direct kinematics of this robot determine the pose of tool $\{T\}$ with respect to the base $\{B\}$ is resolved using equation (2) based on Fig. 3. The base of the robot is at a fixed position on a workbench.

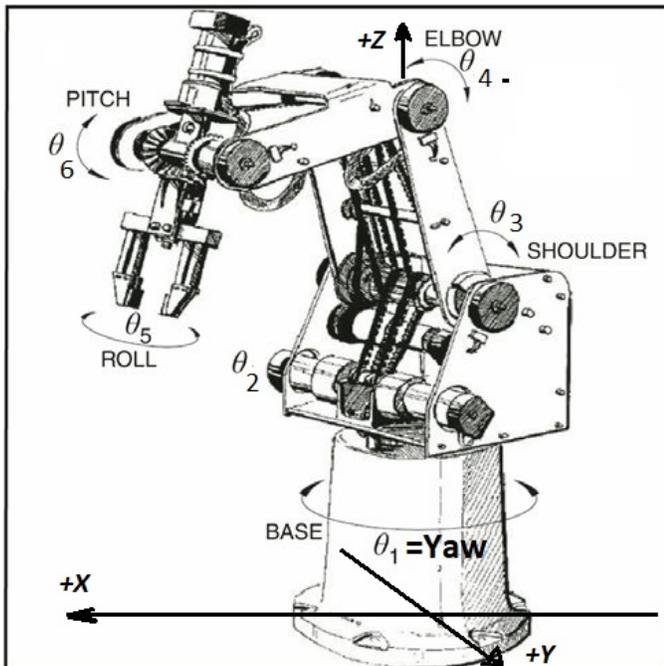
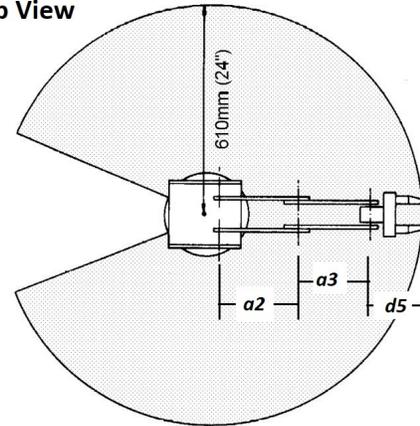


Fig. 1. Scorbot ER-V Plus Robot Details of Angles and Conventions.
Note: the arm occupies a plane coinciding with the z-axis of its base [32].

The robot is controlled using ACL, which is a language that can be used as a multitask robotic programming environment

A: Top View



B: Side View

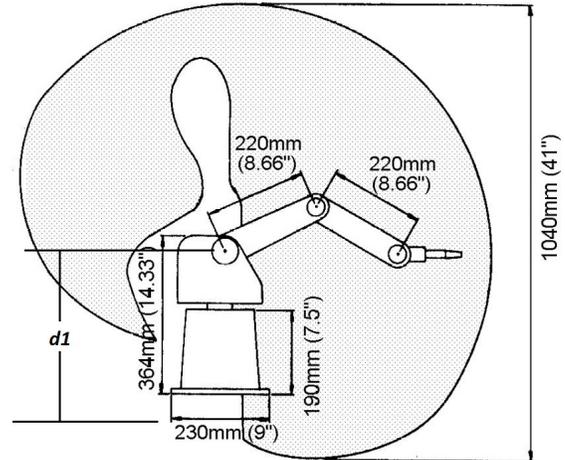


Fig. 2. The Operation Range that Defines the Space and Parameters of the Robot used in the Kinematics [32].

[33, 34]. MATLAB functions were created to establish bidirectional serial communication with the Scorbot controller. Both systems (robot and Computer Vision systems) are running in MATLAB, and give ACL commands which allowed the robot to execute specific tasks, read and load pose data into the controller, and modified the manipulator's movement speed. Fig. 3 also shows the flow of exchanging information between system components.

B. Vision Sensory System

The Kinect V2 sensor (RGB-D sensor) is composed of two cameras, the RGB and an infrared IR camera. The IR camera can be utilized to obtain depth maps, with a field of vision (70° horizontal and 60° vertical). The Kinect camera is capable of running at a rate of (30 fps) at a resolution of (512X424 pixels) and the operational range for the IR camera is between (0.5 m to 4.5 m). The sensor operates based on the time-of-flight principle [35]. The depth data obtained in each pixel corresponds to the Z_i coordinate measured on the optical axis of the IR camera as illustrated in Fig. 4.

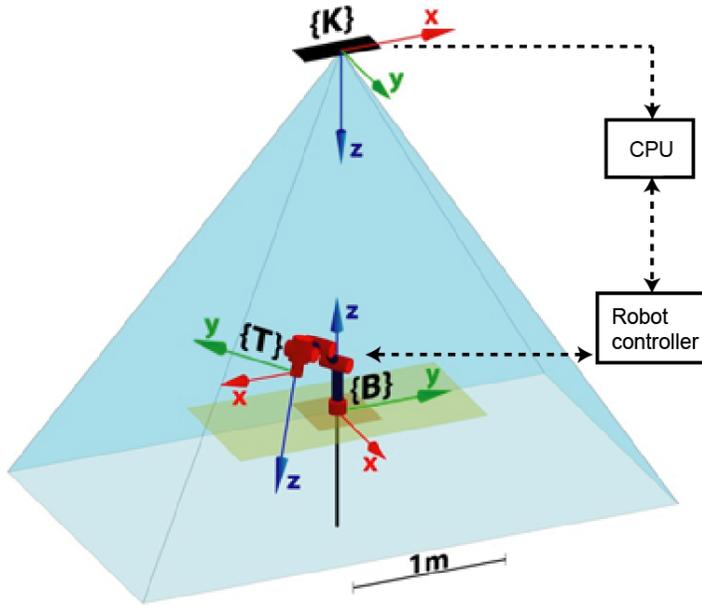


Fig. 3. The Frame Associated with the System Problem Modeling, $\{B\}$ is the base of the Robot, $\{K\}$ Represents the Kinect Sensor, $\{T\}$ Represent the Robot's Tool. The Dotted Line Indicates the Information Flow that is being Exchanged in the Systems.

IV. METHODOLOGY

A. Kinect-Robot Modeling and Calibrating

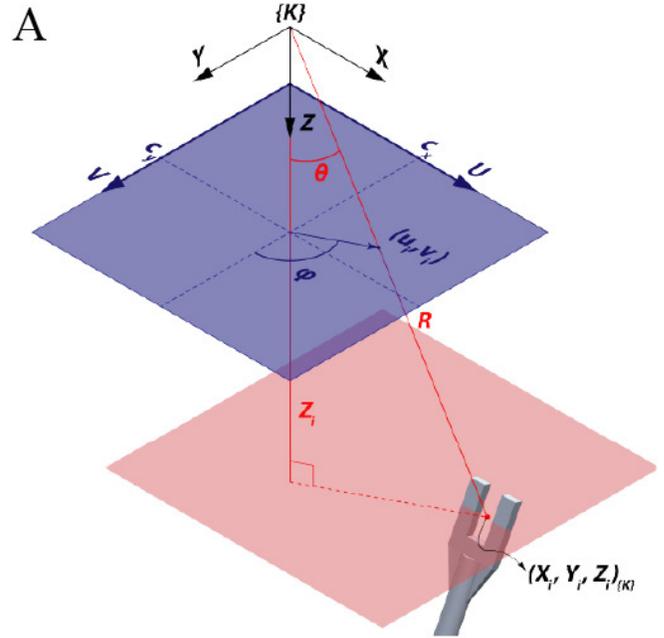
The objective of this work is to provide a system that allows the robot to sense its surroundings and act accordingly. It is necessary to represent the three-dimensional space around the manipulator. There are three important frames [36], the center base of the robot $\{B\}$, the robot's tool $\{T\}$, and the origin of the physical model of the Kinect's depth camera $\{K\}$ as shown in Fig. 3. The robot's task was defined in Cartesian coordinates referred to the base frame $\{B\}$.

For the geometric description of the system, a homogeneous coordinates based on the knowledge gained from [36] was used. The coordinates of a point p with respect to the frame $\{K\}$ is written as ${}^K p = (X_K, Y_K, Z_K, 1)^T$. To calculate the coordinates with respect to frame $\{B\}$ the expression ${}^B p = {}^B_K T \cdot {}^K p$ is used.

The homogeneous matrix T is given by equation (1) where ${}^B_K R$ is a rotation matrix that describes the orientation of the frame $\{K\}$ with respect to the base $\{B\}$, and ${}^B t_K$ corresponds to the coordinates of origin $\{K\}$ in frame $\{B\}$.

$${}^B_K T = \begin{bmatrix} {}^B_K R & {}^B t_K \\ 0_{1 \times 3} & 1 \end{bmatrix} \quad (1)$$

1) *Robot Geometric Model:* In a previous project [37], the direct and inverse kinematics of the Scorbot ER-V Plus were studied. The results presented allowed the calculating of position (${}^B p = (X_B, Y_B, Z_B, 1)^T$) and orientation ($\alpha = yaw, \beta = pitch, \gamma = roll$) of the tool $\{T\}$ in a function of the five rotational angles of the robot ($\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$). Equations



B

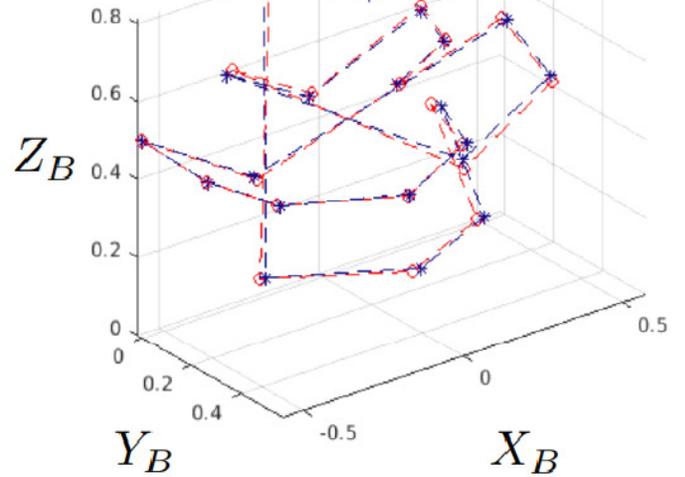


Fig. 4. (A) Geometric Model Utilized to Calibrate the Kinect-Scorbot. Z_i Results in the Intensity of the Pixel (u_i, v_i) . (B) Calibration of Kinect-Scorbot, the Blue Points Indicate the Coordinates of the tool's Position, while the Red Points Indicate the Prediction of the Positions after Calibrating the Parameters f and ${}^B_K T$ of the Model.

(2) and (3) represent the results while Fig. 1 represents the parameters of the robot.

$$\begin{aligned} X_B &= c_{\theta_1} (a_3 c_{\theta_S} + a_2 c_{\theta_2} + d_5 c_{\theta_W}) \\ Y_B &= s_{\theta_1} (a_3 c_{\theta_S} + a_2 c_{\theta_2} + d_5 c_{\theta_W}) \\ Z_B &= d_1 + a_3 s_{\theta_S} + a_2 s_{\theta_2} + d_5 s_{\theta_W} \end{aligned} \quad (2)$$

$$\begin{aligned} \theta_S &= \theta_2 + \theta_3 \\ \theta_W &= \theta_2 + \theta_3 + \theta_4 \end{aligned}$$

$$\begin{aligned}\alpha &= \theta_1 = \arctan(Y_B, X_B) \\ \beta &= -\theta_2 - \theta_3 - \theta_4 = -\theta_W \\ \gamma &= \theta_5\end{aligned}\quad (3)$$

Note that once θ_1 is defined, the robot is contained in a plane that coincides with the Z -axis of the first articulation. This observation is important because it allows the construction of a binary mask that allows the Scorbot detection when moving.

2) *RGB-D geometric model*: The camera model used was a *pinhole* type. The hypothesis is as follow: the origin of the frame $\{K\}$ coincides in XY with the center of the image (c_x, c_y) , and the focal distance f is the same in X as in XY . Spherical coordinates were employed to map the coordinates (u_i, v_i) and data Z_i with coordinates $(X_i, Y_i, Z_i, 1)_{\{K\}}^T$ as shown in equation (4)-(6).

$$\varphi_i = \text{atan}\left(\frac{v_i - c_y}{u_i - c_x}\right) \quad (4)$$

$$\theta_i = \text{atan}\left(\frac{\sqrt{(u_i - c_x)^2 + (v_i - c_y)^2}}{f}\right) \quad (5)$$

$$R_i = \frac{Z_i}{\cos(\theta_i)} \quad (6)$$

Thus, the three-dimensional point p_i has coordinates in frames $\{K\}$ and $\{B\}$ given by (7)-(8).

$${}^K p_i = (R \cos(\theta_i) \cos(\phi_i), R \cos(\theta_i) \sin(\phi_i), Z_i, 1)^T \quad (7)$$

$${}^B p_i = {}^B T \cdot {}^K p_i = ({}^K T)^{-1} \cdot {}^K p_i \quad (8)$$

3) *Geometric calibration of Kinect-Scorbot system*: The geometric calibration of the robot was accomplished in [38]. The only intrinsic parameter that was considered unknown in the Kinect was the focal distance of f . Also, the extrinsic parameter that represents the pose of the camera $\{K\}$ with respect to the robot $\{B\}$ needed to be calibrated.

The experiment had the robot take a wooden cube using its claw in a way that allows the center mass of the block to be aligned with the manipulator's tool frame. The Kinect was placed on the roof of the lab as seen in Fig. 3. Each measurement is represented by the index i , making a total of $N = 22$. As a pattern for the adjustment of the camera model, the coordinates were obtained from the robot driver (${}^B p_{exp,i}$) through serial port communication. Given the parameters $(f, {}^B T)$ coordinates can be predicted ${}^B p_{pred,i}$ and calculate a prediction error, defined by equation (9).

$$\epsilon = \frac{1}{N} \sum_{i=1}^N |{}^B p_{exp,i} - {}^B p_{pred,i}(f, {}^B T)| \quad (9)$$

To simplify the optimization problem, it was assumed that the optical axis of the Kinect camera was perpendicular to the XY plane of the plot $\{B\}$, and the Z axes of the two frames were parallel and opposite to each other. The optimized parameters resulted in the following values: the focal length, and the position of the camera $\{K\}$ with respect to the frame $\{B\}$: ${}^B t_K$.

First, f was adjusted so that ϵ equation (9) is minimal, starting with a ${}^B t_K = (0, 0, 2.40)^T m$. This resulted in a perfect alignment in XY of the camera and the robot in Fig. 3 with a value of Z_i taken from a depth image. With the focal distance optimized, the 22 points are re-projected and a mean error for each axis was computed. These deviations were introduced as corrections in ${}^B t_K$ to reduce the mean re-projection error. With the adjusted transformation, a new focal point was computed. With the mean error being negligible for each axis, the parametric adjustment at that point was finalized.

The optimized focal distance resulted in 362.8 *pixels*. The mean error of re-projection from the 22 coordinates was 1.70 *cm* and a standard deviation of 0.87 *cm* and a peak of 3.66 *cm*. The transformation that maps the frame $\{K\}$ with $\{B\}$ was determined by equation 10.

$${}^B T = \begin{pmatrix} 0 & 1 & 0 & -0,0142 \text{ m} \\ 1 & 0 & 0 & -0,0289 \text{ m} \\ 0 & 0 & -1 & 2,3907 \text{ m} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (10)$$

B. Human Detection using Background-Foreground Technique

For human or foreign object detection in the scene, a Background-Foreground (B-F) technique [39] was used. 100 frames of depth images were captured within 10 *sec* and used to form images of the *background* making sure the scene stayed static.

As previously mentioned, rotating and fixed rectangular binary masks were generated to avoid the detection of the robot's movement by the *foreground*. A captured image was printed on the screen, and the mouse determined the vertices of the two rectangles and a fixed point for one of them to rotate. The non-rotating rectangle was used to hide the base of the robot from the *foreground*. The rotating rectangle did the same with the extension of the maximum possible arm. The fixed point corresponded approximately with the robot axis. The angle of rotation of this mask was computed by reading the status of the encoders of the robot, and applying direct kinematics as in (2) so that it could follow the movement of the plane occupied by the robot Fig. 2.

Human detection scenarios differ slightly for the applications selected, and they are described below.

1) *Collision prevention*: Three areas were determined to be evaluated in the depth images, which represent the severity of the collision. Starting from the robot base, and utilizing the Kinect sensor calibration, two sections were established to determine the *red* and *yellow* areas in the images. The sections were 660 *mm* which is 50 *mm* more than the maximum reach of the robot for the *red zone* and 1150 *mm* for the *yellow zone*. The *green zone* was considered outside the radius of the *yellow zone*. The behavior of the robot was modeled as

a machine of finite states. There were (1) *Green* is a normal speed, (2) *Yellow* is a medium speed, (3) *Red* is minimum/very slow speed, which is illustrated in Fig. 5.

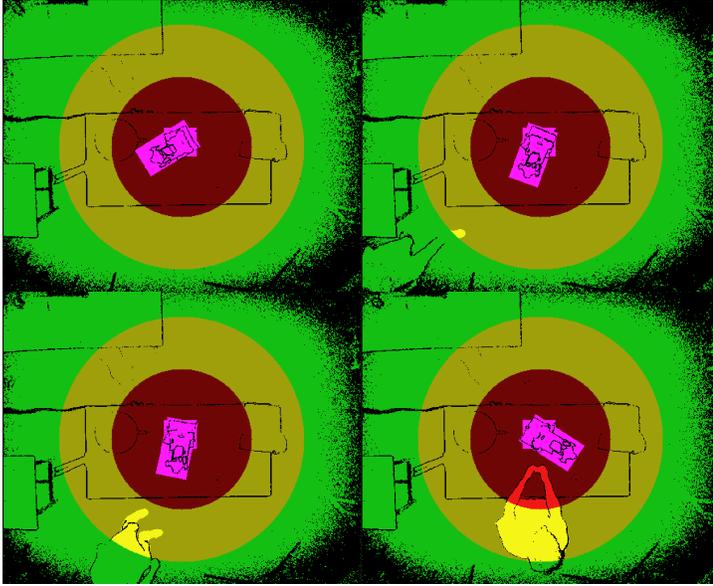


Fig. 5. Collision Prevention Scenario.

At the beginning of each iteration, a depth image will be captured, and the value of the robot's base encoder will be gathered. Then a binary mask will be added to the foreground where the captured image was subtracted from the background then a binary mask was applied to hide the robot. An opening was performed to the resulting image with a 5 pixels radius kernel disk to remove the noise. Finally, a 50 mm depth threshold was used to binarize the image.

The B-F results combined with the areas of interest to determine the behavior of the robot's speed. If the *foreground* binary area within the *red* zone exceeded 100 pixels, the state turns *red*. If the area is not exceeding 100 pixels in the *red* zone but reaching at least 500 pixels in the *yellow* zone, then the state turns *yellow*. If none of the above conditions are met, the state updates to *green*.

2) *Collaborative Scenario*: Each iteration started with the robot taking an object located at a pre-established location and a request that will appear on the user screen to guide the operator to position her/his hand where she/he wanted to receive the object from the robot. Subsequently, a binary mask will be generated for the *foreground*. The *background* image will be subtracted from the captured image and applied a 15 mm depth threshold to make it binary. It was decided to analyze a 200 pixels radius to avoid dealing with peripheral noise. The radius was equivalent to the calibration at 1.32 m at the height of the workbench. The *foreground* was cleaned by imposing an opening using a disk of 4 pixels radius like a kernel. After closing was imposed with a kernel disk with a radius of 3 pixels to remove any imperfections remaining in the *blobs*. The *blobs* with an area smaller than 800 pixels were discarded. A binary mask was generated with the remaining *blobs*. Two zones were separated by heights zones in the resulting *blobs* using Otsu's method [40].

Given the characteristics of the system, the system will be able to identify the head, torso, and arms. The portion of *blobs* that had the torso and arms inside the robot's work area will be isolated. The pixels within the radius of the work-space were filtered from the processed mask. Finally, the coordinate (u_i, v_i) of the pixel corresponding to the center of the hand was found. To find the center of the palm, a skeletonized binary image was obtained [39]. The team looked for the radius with the maximum circumference in pixels that could fit the binary mask. The coordinate of the pixel with the largest radius was preserved and the depth value Z_i assigned was that of greater repetition within the maximum circumference that could fit in the mask with center (u_i, v_i) , and applied to the originally captured image.

The coordinates (u_i, v_i) and Z_i obtained were transformed. First to coordinates $(X_i, Y_i, Z_i, 1)_{\{K\}}^T$ using equations (4)-(6), and then to $(X_i, Y_i, Z_i, 1)_{\{B\}}^T$, by using equation (10). At this position, a height increase, Z_B , of 7 cm was made to prevent collisions with the operator's hand. Then adjusted height was entered automatically by serial communication to an internal variable of the robot controller. This allowed the end effector to deposit the object at the desired position. As a result, the collaborative job will be completed as shown in Fig (6).

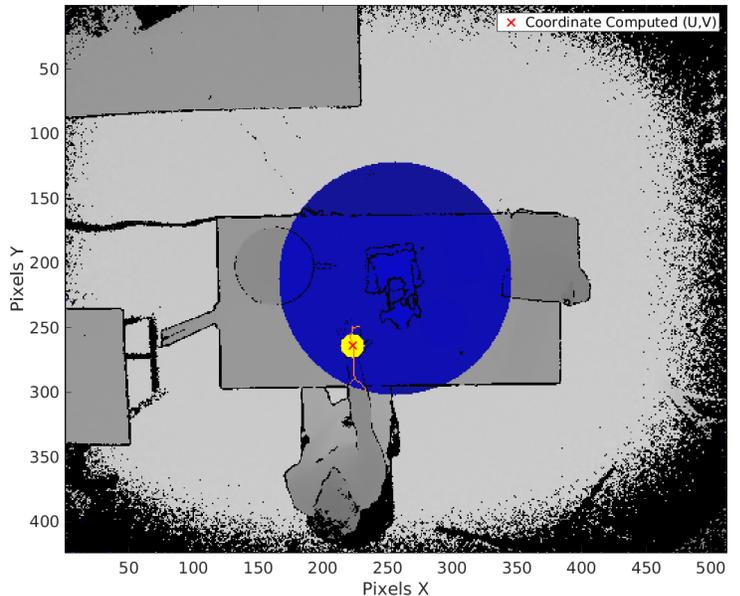


Fig. 6. Collaboration with an Operator Scenario.

V. RESULTS

This work is meant to develop an affordable prototype that can be added to industrial robots to increase robot safety, decrease the barriers between human operators and robots, and facilitate a collaboration system between them. The designed system addressed these goals as follow:

A. Collision Prevention System

The detection of an operator in the pre-established zones is exemplified in Fig. 5. To test the operation of the system, 10 tests were made in areas of interest, where a human operator will introduce his/her hand into the robot surroundings. The

system was used to detect the operator when first entering the yellow zone (Operator's leg) where the system forced the robot to move at half of its original operation speed. Then the operator introduced his/her hands within the red zone to force the robot to slow significantly to almost not moving. These actions were captured by the camera and highlighted by the associated colors shown in Fig. 5, which displays the areas of interest corresponds to the detection of the pixels as part of the foreground and the outline of the ScorBot is shown inside the binary mask.

In all test cases, the system behaved correctly as intended, by identify the existence of the human operator and change the robot speed according to the distance between the human and the robot. The robot response time to change the end-effector speed was recorded and the mean system update time was 0.45 s with a standard deviation of 0.30 s, which is a significantly fast response.

Modify the speed of the robot was accomplished through an ACL command called 'CLRBUF', which was introduced as an instant stop to the robot followed by an immediately a new movement speed was set, and a new trajectory was generated from the current pose until the next corresponding task resuming the job. the team implemented other methods to change the speed by changing the task priorities on the robot or send speed change commands during a test but all failed since these commands could only be utilized after completing the previous tasks.

B. Collaborative Scenario

Collaboration between the robot and human operator was simulated by having the automated system detect the operator's hand and estimate the spatial coordinate of the center of the hand then command the robot will move to pick up an object from a predefined location then place it on the operator's hand, this is illustrated in Fig. 6. The blue region represents the Scorbot work-space, the orange lines show the skeletonization of the operator's arm while the yellow area shows the mask's maximum circumference where the robot should place the object on.

Experiments with 20 different hand positions within the robot's work-space were conducted, the system gave satisfactory results, where the job was done correctly, and placement coordinates mean error was 2.4 cm.

VI. CONCLUSION

This work showed that an overhead low-cost RGB-D camera can measure the position of an operator with respect to a robotic manipulator, and thus improve human-robot interaction safety and increase the collaboration opportunities through 3D sensing of the robot surrounding environment. This proposed system will allow manufacturing and industrial companies to update their existing robotics and automation system by adding an affordable add-on safety and collaboration device without influencing their manufacturing lines with a lower cost of investment.

In the collision prevention scenario, the captured video analysis proved that the reaction times of the system was 500ms and the system's bottleneck was the PC and robot

inter-communication which required relatively longer times and added pauses and checkpoints to make sure it is reliable.

In the collaborative scenario, detecting the operator's hand and have the robot placing an object was achieved, and similar to the other scenario, the internal variables, and data transmitting speed between the robot controller and the main computer was the main factor to defined the speed of the system.

The team is working on a few improvements to the proposed system including enhancing the B-F algorithm internal variables and date, exploring the application of dynamics methods that can assimilate changes in the scene on slower times scales. Also, an RGB camera system development is being conducted to detect a particular color or clothing as an activator for robot tasks. Additionally, more sophisticated moving object classification techniques such as convolution neural networks will be explored.

ACKNOWLEDGMENT

This project was funded and supported by "100,000 Strong in the Americas Innovation Fund – Exxon Mobil", and the U.S. Department of Education - Minority Science and Engineering Improvement (MSEIP) Program (Grant # P120A170102), National University of Quilmes, and Texas A&M International University. With their support, this project was a successful collaboration between the two institutes.

REFERENCES

- [1] Matheson, E., Minto, R., Zampieri, E. G., Faccio, M., & Rosati, G. (2019). Human-Robot Collaboration in Manufacturing Applications: A Review. *Robotics*, 8(4), 100.
- [2] Ferraguti, F., Bertuletti, M., Landi, C. T., Bonfe, M., Fantuzzi, C., & Secchi, C. (2020). A control barrier function approach for maximizing performance while fulfilling to iso/ts 15066 regulations. *IEEE Robotics and Automation Letters*, 5(4). <https://doi.org/10.1109/LRA.2020.3010494>
- [3] Robla-Gomez, S., Becerra, V. M., Llata, J. R., Gonzalez-Sarabia, E., Torre-Ferrero, C., & Perez-Oria, J. (2017). Working together: a review on safe human-robot collaboration in industrial environments. *IEEE Access*, 5. <https://doi.org/10.1109/ACCESS.2017.2773127>
- [4] Harper, C., & Virk, G. (2010). Towards the development of international safety standards for human-robot interaction. *International Journal of Social Robotics*, 2(3), 229-234.
- [5] Anandan, T. M. (2013). Safety and control in collaborative robotics. Published on: Aug, 6, 1-4.
- [6] Paolanti, M., Romeo, L., Liciotti, D., Cenci, A., Frontoni, E., & Zingaretti, P. (2018). Person Re-Identification with RGB-D Camera in Top-View Configuration through Multiple Nearest Neighbor Classifiers and Neighborhood Component Features Selection. *Sensors*, 18(10), 3471. MDPI AG. <http://dx.doi.org/10.3390/s18103471>
- [7] Zhang, J., Li, W., Ogunbona, P. O., Wang, P., & Tang, C. (2016). RGB-D-based action recognition datasets: A survey. *Pattern Recognition*, 60, 86-105. <https://doi.org/10.1016/j.patcog.2016.05.019>.
- [8] Coupeté, E., Moutarde, F., & Manitsaris, S. (2015). Gesture recognition using a depth camera for human-robot collaboration on an assembly line. *Procedia Manufacturing*, 3, 518-525. <https://doi.org/10.1016/j.promfg.2015.07.216>.
- [9] Shukla, D., Erkent, Ö., & Piater, J. (2016, August). A multi-view hand gesture RGB-D dataset for human-robot interaction scenarios. In 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN) (pp. 1084-1091). IEEE. <https://doi.org/10.1109/ROMAN.2016.7745243>.
- [10] Lorenzo-Navarro J, Castrillón-Santana M, Hernández-Sosa D. On the Use of Simple Geometric Descriptors Provided by RGB-D Sensors for Re-Identification. *Sensors*. 2013; 13(7):8222-8238. <https://doi.org/10.3390/s130708222>
- [11] Lorenzo-Navarro J., Castrillón-Santana M., Hernández-Sosa D. (2012) An Study on Re-identification in RGB-D Imagery. In: Bravo J., Hervás

- R., Rodríguez M. (eds) Ambient Assisted Living and Home Care. IWAAL 2012. Lecture Notes in Computer Science, vol 7657. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-35395-6_28
- [12] Paolanti, M., Pietrini, R., Mancini, A., Frontoni, E., & Zingaretti, P. (2020). Deep understanding of shopper behaviours and interactions using RGB-D vision. *Machine Vision and Applications*, 31(7), 1-21. <https://doi.org/10.1007/s00138-020-01118-w>
- [13] Liciotti D., Paolanti M., Frontoni E., Zingaretti P. (2017) People Detection and Tracking from an RGB-D Camera in Top-View Configuration: Review of Challenges and Applications. In: Battiato S., Farinella G., Leo M., Gallo G. (eds) *New Trends in Image Analysis and Processing – ICIAP 2017*. ICIAP 2017. Lecture Notes in Computer Science, vol 10590. Springer, Cham. https://doi.org/10.1007/978-3-319-70742-6_20
- [14] Moallem, P., Razmjoo, N., & Ashourian, M. (2013). Computer vision-based potato defect detection using neural networks and support vector machine. *International Journal of Robotics and Automation*, 28(2), 137-145. <https://doi.org/10.2316/Journal.206.2013.2.206-3746>
- [15] Turarova, A., Zhanatkyzy, A., Telisheva, Z., Sabyrov, A., & Sandygulova, A. (2020, March). Child Action Recognition in RGB and RGB-D Data. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 491-492). <https://doi.org/10.1145/3371382.3378391>
- [16] Gualtieri, L., Rauch, E., Vidoni, R., & Matt, D. T. (2020). Safety, ergonomics, and efficiency in human-robot collaborative assembly: design guidelines and requirements. *Procedia Cirp*, 91, 367-372. <https://doi.org/10.1016/j.procir.2020.02.188>
- [17] Ore, F., Vemula, B., Hanson, L., Wiktorsson, M., & Fagerstrom Bjorn. (2019). Simulation methodology for performance and safety evaluation of human-industrial robot collaboration workstation design. *International Journal of Intelligent Robotics and Applications*, 3(3), 269-282. <https://doi.org/10.1007/s41315-019-00097-0>
- [18] Parigi Polverini, M., Zanchettin, A. M., & Rocco, P. (2017). A computationally efficient safety assessment for collaborative robotics applications. *Robotics and Computer Integrated Manufacturing*, 46, 25- 37. <https://doi.org/10.1016/j.rcim.2016.11.002>
- [19] Zanchettin, A. M., Ceriani, N. M., Rocco, P., Hao, D., & Matthias, B. (2016). Safety in human-robot collaborative manufacturing environments: metrics and control. *IEEE Transactions on Automation Science and Engineering*, 13(2). <https://doi.org/10.1109/TASE.2015.2412256>
- [20] Vicentini, F. (2020). Terminology in the safety of collaborative robotics. *Robotics and Computer Integrated Manufacturing*, 63. <https://doi.org/10.1016/j.rcim.2019.101921>
- [21] Vicentini, F., Askarpour, M., Rossi, M. G., & Mandrioli, D. (2020). Safety assessment of collaborative robotics through automated formal verification. *IEEE Transactions on Robotics*, 36(1), 42-61. <https://doi.org/10.1109/TRO.2019.2937471>
- [22] Franklin, C. S., Dominguez, E. G., Fryman, J. D., & Lewandowski, M. L. (2020). Collaborative robotics: a new era of human-robot cooperation in the workplace. *Journal of Safety Research*, 74, 153-160. <https://doi.org/10.1016/j.jsr.2020.06.013>
- [23] Echavarrí Javier, Ceccarelli, M., Carbone, G., Alen Cristina, Muñoz Jose Luis, Díaz Andres, & Muñoz-Guijosa, J. M. (2013). Towards a safety index for assessing head injury potential in service robotics. *Advanced Robotics*, 27(11), 831-844. <https://doi.org/10.1080/01691864.2013.791655>
- [24] Smit-Anseeuw, N., Remy, C. D., & Vasudevan, R. (2019). Walking with confidence: safety regulation for full order biped models. *IEEE Robotics and Automation Letters*, 4(4), 4177-4184. <https://doi.org/10.1109/LRA.2019.2931225>
- [25] Alonso-Mora, J., Beardsley, P., & Siegwart, R. (2018). Cooperative collision avoidance for nonholonomic robots. *IEEE Transactions on Robotics*, 34(2), 404-420. <https://doi.org/10.1109/TRO.2018.2793890>
- [26] Gurriet, T., Mote, M., Singletary, A., Nilsson, P., Feron, E., & Ames, A. D. (2020). A scalable safety-critical control framework for nonlinear systems. *IEEE Access*, 8. <https://doi.org/10.1109/ACCESS.2020.3025248>
- [27] Oyekan, J. O., Hutabarat, W., Tiwari, A., Grech, R., Aung, M. H., Mariani, M. P., & Dupuis, C. (2019). The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans. *Robotics and Computer Integrated Manufacturing*, 55, 41-54. <https://doi.org/10.1016/j.rcim.2018.07.006>
- [28] Askarpour, M., Mandrioli, D., Rossi, M., & Vicentini, F. (2019). Formal model of human erroneous behavior for safety analysis in collaborative robotics. *Robotics and Computer-Integrated Manufacturing*, 57, 465-476. <https://doi.org/10.1016/j.rcim.2019.01.001>
- [29] Giuliani, M., Lenz, C., Muller Thomas, Rickert, M., & Knoll, A. (2010). Design principles for safety in human-robot interaction. *International Journal of Social Robotics*, 2(3), 253-274. <https://doi.org/10.1007/s12369-010-0052-0>
- [30] Raiola, G., Cardenas, C. A., Tadele, T. S., de Vries, T., & Stramigioli, S. (2018). Development of a safety- and energy-aware impedance controller for collaborative robots. *IEEE Robotics and Automation Letters*, 3(2), 1237-1244.
- [31] Hull, T., & WCX™ 17: SAE World Congress Experience Detroit, Michigan, United States 2017-04-04. (2017). Intelligent robotics safeguarding. *SAE International Journal of Engines*, 10(2), 215-221. <https://doi.org/10.4271/2017-01-0293>
- [32] Intelitek Inc. Scrobot-ER V plus: User's Manual. Catalog #100016 Rev. C February 1996.
- [33] Siciliano, B. & Khatib, O. (Eds.). (2016). *Springer handbook of robotics*. Springer.
- [34] Robotec, E. (1994). *ACL-Advanced Control Language Versions 1.43, F1. 44 and ATS-Advanced Terminal Software Version 1.44*.
- [35] Villena Martínez, V. (2015). Comparative analysis of calibration methods for RGB-D sensors and their influence on the recording of multiple views. [36] Craig, J. J. (2005). *Introduction to Robotics*.
- [37] Del Vigo, L., Peyton, R., Bussi, U., & Oliva, D. E. (2018, November). Computer-aided design and simulation of a robotic manipulator for educational purposes. In *2018 Argentine Conference on Automatic Control (AADECA)* (pp. 1-6). IEEE.
- [38] Pessag, F., De Cristoforis, P., Perri, V. & Oliva, D. (2014). Implementación y validación de un modelo cinemático para el manipulador robótico SCORBOT ER-Vplus. *Argentina, CABA. VIII Jornadas Argentinas de Robotica, ISBN-13*. pp. 978-987-1087-19-9.
- [39] Russ, J. C. & Neal, F. B. (2016). *The image processing handbook*, 7th ed. CRC Press. pp. 381-391
- [40] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1), 62-66.