

Research Article

Toward Design of a Drip-Stand Patient Follower Robot

Zewen Wu and Shahram Payandeh 

Network Robotics and Sensing Laboratory, School of Engineering Science, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

Correspondence should be addressed to Shahram Payandeh; payandeh@sfu.ca

Received 16 October 2019; Revised 29 January 2020; Accepted 3 February 2020; Published 9 March 2020

Academic Editor: Gordon R. Pennock

Copyright © 2020 Zewen Wu and Shahram Payandeh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A person following robot is an application of service robotics that primarily focuses on human-robot interaction, for example, in security and health care. This paper explores some of the design and development challenges of a patient follower robot. Our motivation stemmed from common mobility challenges associated with patients holding on and pulling the medical drip stand. Unlike other designs for person following robots, the proposed design objectives need to preserve as much as patient privacy and operational challenges in the hospital environment. We placed a single camera closer to the ground, which can result in a narrower field of view to preserve patient privacy. Through a unique design of artificial markers placed on various hospital clothing, we have shown how the visual tracking algorithm can determine the spatial location of the patient with respect to the robot. The robot control algorithm is implemented in three parts: (a) patient detection; (b) distance estimation; and (c) trajectory controller. For patient detection, the proposed algorithm utilizes two complementary tools for target detection, namely, template matching and colour histogram comparison. We applied a pinhole camera model for the estimation of distance from the robot to the patient. We proposed a novel movement trajectory planner to maintain the dynamic tipping stability of the robot by adjusting the peak acceleration. The paper further demonstrates the practicality of the proposed design through several experimental case studies.

1. Introduction

One of the common types of service robotics is a person following robot, which has found applications in several fields such as security, surveillance, and elderly monitoring [1, 2]. A nurse following robot can efficiently reduce the workload for nurses and health delivery system [3]. The design and development of an autonomous mobile robot to haul and transport hospital supplies and improving hospital efficiency was proposed in [4]. Nurses may spend as much as 30% of their time away from the patient being involved in tasks such as fetching various medications or reports. Using robots to perform those tasks, thereby reducing or eliminating the time, can have a significant benefit for the patient and health delivery system. Similarly, the proposed design of an autonomous drip-stand patient follower robot can offer a more convenient tool for ease of patient mobility and support during their movements in the hospital.

Currently, patients need to manually pull their drip stand to accommodate it along with their motion patterns. The patient usually holds on to the stand on their side and ensures that there is adequate slack in the feeding tubes. Through study of various notions for better design of a patient interaction system, this paper proposes a novel design of a robotic system. The system can autonomously follow the patient from behind and ensures that there exists a proper distance between the drip stand and the patient for a given allowable slack in the infusion tubes. In the proposed design, we included a standard, off-the-shelf RGB camera as a part of the sensor processing unit. We placed the sensor closer to the ground level in order to better protect the patient's privacy by constraining the viewing angle. Additional sensing modality based on depth sensing is utilized only for performance comparison study with the computational model obtained through an RGB camera. Patient's regular hospital clothing is rectified with artificial markers for robust distance estimation, thus enabling easy hospital deployment.

The overall performance requirements for an autonomous drip stand are similar to that of person following robot. We designed the target tracking and motion planning to maintain an allowable distance between the robot and the target. We defined and processed various features in the specified part of the input image, which we obtained by an onboard camera or depth sensor. These features can be used to determine the position of the object with respect to the sensor frame and its associated speed. For example, in [5], a target detector is implemented based on the colour properties of the image. They used a colour histogram and a mean-shift tracker to find the area that is locally similar to the stored template defined at the initial frame. The tracking algorithm proposed in [6] improves the classical mean-shift with spatial-colour feature to define a new similarity measure. The Bhattacharyya coefficients are used to compute the similarity between the previously selected target and potential targets. Another approach to use to track an object is by using an RGB-D sensor in combination with other sensing modalities such as laser and thermal sensing. The RGB-D sensor captures RGB images along with per-pixel depth information in the live data stream in order to assist in scene segmentation and detection [7]. The person tracking system introduced in [8] is based on a laser sensor, a thermal sensor, and a depth sensor combining the leg and vest detections with the heat sensing. Here, the laser sensor can localize moving blobs to extract the position of the legs. Concurrently, RGB-D images are used to track the target and estimate the position of the target during the operational tasks and through vest detection.

We organized the paper as follows. Section 2 presents the basics of the visual tracking methodologies we utilized. Section 3 presents an overview of the enhanced trajectory controller for the drip-stand patient follower robot. Section 4 presents a detailed experimental evaluation and comparison of the camera-based patient follower robot, and finally Section 5 presents some concluding remarks.

2. Visual Tracking

Tracking of the patient is accomplished through processing the video images obtained from the onboard RGB camera. We also explored other sensing modalities, such as depth sensing, and we compared the results with the tracking information obtained through RGB cameras. Primarily, we focused on utilization of images we obtained through these standard RGB cameras. The small size of these cameras and their low costs allowed easy integration with the final design of the drip-stand patient follower robot. The placement of these cameras was also essential in protecting privacy of the patient. We also found that any processing of the video images faces further challenges associated with the presence of uniform colour distribution in the hospital environment. Both the surrounding background and foreground images containing patient wearing hospital clothing had uniform colour distribution. Figure 1 shows examples of typical hospital environments.

2.1. Design for Artificial Marker. For the proposed patient following robot, we selected an artificial marker with known size and colour. We placed this landmark on the hospital's clothing for the patient. Previously, we have followed a similar visual tracking methodology for tracking tools in surgical environments [9]. We designated a colour identification patch for each patient occupying a common area. A drip-stand robot can be assigned only to follow the assigned colour associated with the specific patient. Figure 2 shows an example of image processing results where the system can identify the red colour patch and further determine its width. The pinhole camera model can then use this information for estimating the distance of the marker to the robot. The presence of the yellow marker in the figure demonstrates an example where we tuned the algorithm to only segment the assigned red colour [10].

Given the segmented image of the marker and associated estimation of its width in the projected plane, it is now possible to estimate the distance of the marker to the robot. We computed the distance using the geometry of similar triangles. For example, the geometry of a pinhole camera is shown in Figure 3(a). Through similar triangles shown in Figure 3(b), we can determine the distance along the x_3 axis. The captured image plane is located at the distance f from the origin in the negative direction. The camera aperture is located at (x_1, x_2, x_3) . Let P be a point located on a physical marker attached to the patient with measured distance of $\sqrt{x_1^2 + x_2^2}$. Point Q is on the projection of the segmented image of the marker on the image plane with the measured distance of $\sqrt{y_1^2 + y_2^2}$. The similarity relationship between the actual image and the projected image can be written as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = -\frac{f}{x_3} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad (1)$$

where f is the focal length of the camera, which can be obtained through various calibration methods. For the experimental setup of this paper, the calibrated parameter is determined to be $f = 15.42$. The autofocus function of the camera was turned off to ensure better accuracy in the computation. We further calibrated to associate the physical size to the pixel size. The above relationship is used to compute the magnitude of the distance of the marker in the (x_2, x_3) projection plane. We obtained a similar relationship for the triangles in the (x_1, x_3) plane, which results in determining the coordinate of the representative point P with respect to the camera frame.

In this study, we are considering two primary categories for the placement of the marker on the patient, which are dependent on the type of clothing they might wear. The first category is for the cases where the patient is wearing no special clothing or when they are wearing hospital pants or socks. Figures 4(a)–4(c) depict example images for this category. For all three cases, we used a known sized strip of colour marker. The second category for marker attachment is for the case when the patient is wearing a hospital gown. Unlike the first category, the shape of the attached marker



FIGURE 1: Two typical views from the hospital environment, demonstrating the existence of colour uniformity.

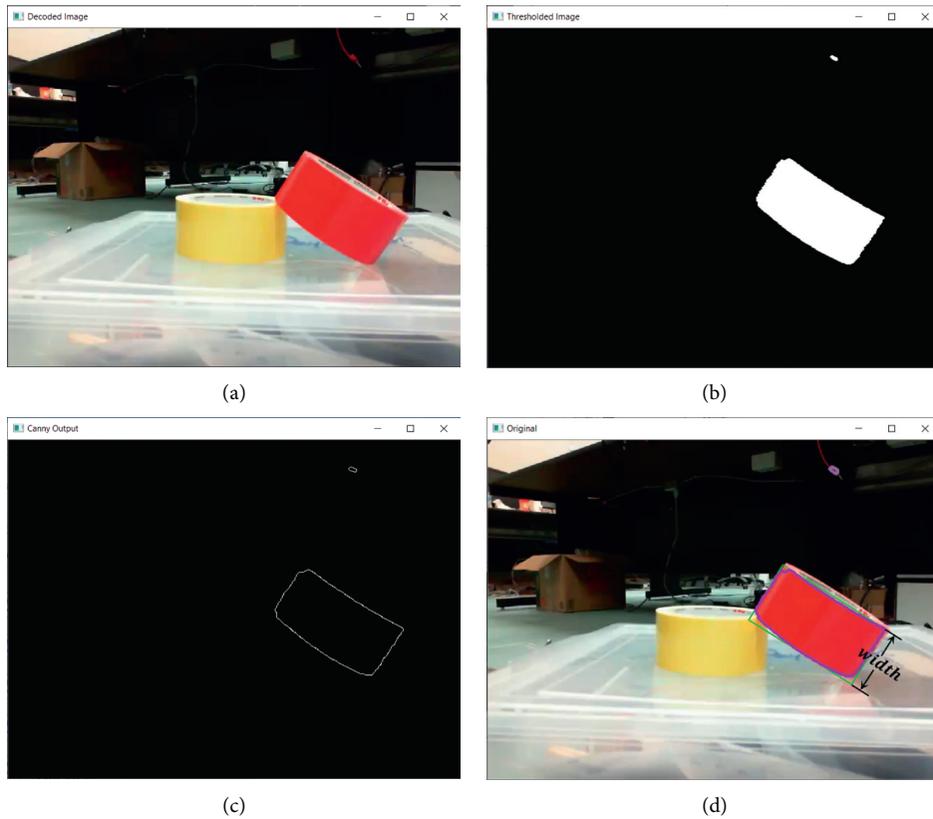


FIGURE 2: Results of image processing for identifying a red colour marker and its width in the projected image. The processing can be adjusted to recognize different colour identifiers for each patient. (a) The input sample image containing two different colours. (b) Segmentation of the red colour and further thresholding. (c) Edge detection of the segmented image. (d) Computation of the width of the projected image which can further be used in the pinhole camera model.

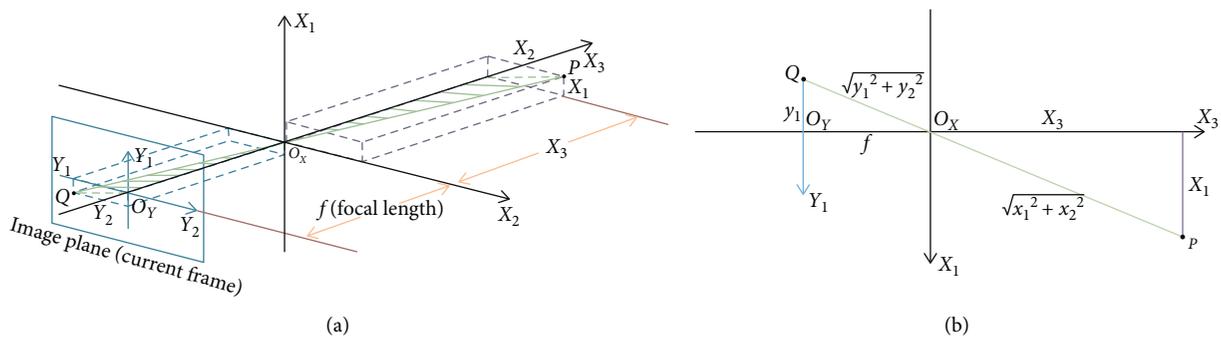


FIGURE 3: (a) Pinhole camera model representation and (b) an associated projection plane showing similar triangles.

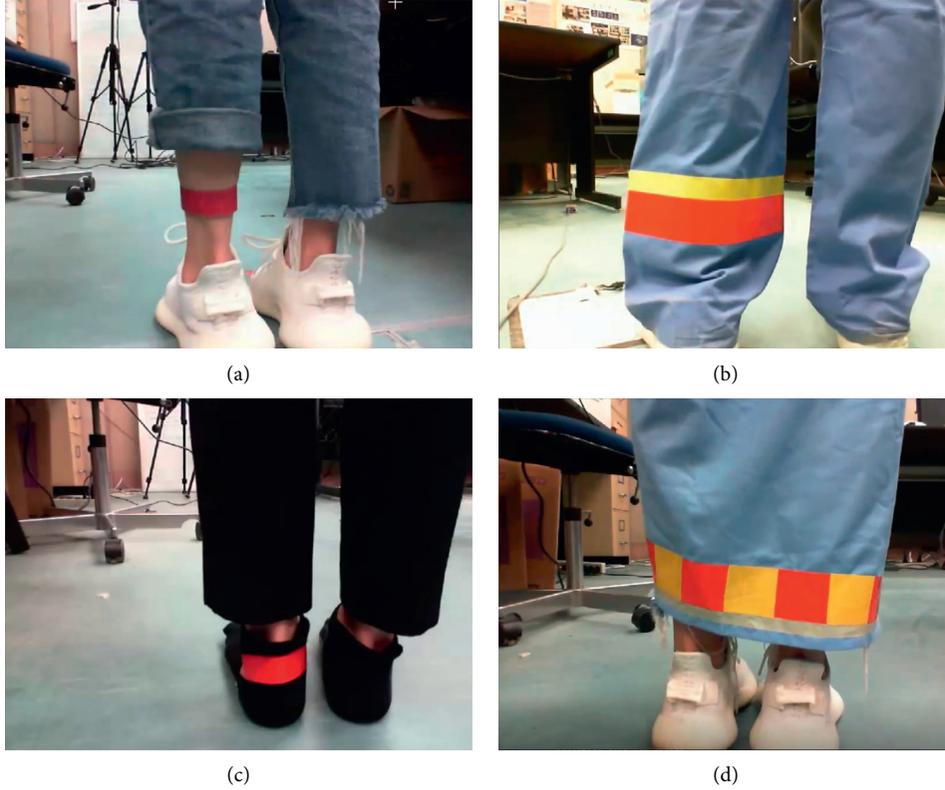


FIGURE 4: Attachment of artificial colour marker for the first and second category of hospital clothing. (a) For the case where the patient does not wear hospital clothing. (b) When wearing hospital pants. (c) When wearing hospital socks. (d) When wearing hospital gown.

can undergo large deformation. Figure 4(d) shows an example of attachment of this type of marker on a patient's gown.

2.2. Initialization for Rigid Target Identification. The marker detection and tracking algorithm explores the utilization of both template matching and colour histogram of the object within the initially defined region of interest (ROI). The ROI corresponds to the initial state of the patient, which the drip stand robot controller needs to use to initiate the tracking phase. For the first class of marker design and its placement (Figures 4(a)–4(c)), the initialization of ROI is defined depending on the size of the first segmented marker. We defined the region of the template through our experiments to be slightly larger than the marker. The larger size alleviates impacts caused by initial position of the legs and their movements. The robot also initiates a self-rotational calibration strategy about its axis to have the initially detected marker closer to the center of the frame. This results in the center of the marker to be within the desired range. The robot then can self-navigate following the position of the patient.

2.2.1. Template Matching. The most direct way to orient the ROI at a current frame is by corresponding the definition of a template defined in the initial frame. There exist numerous approaches for performing template matching [10]. Here, we

have selected the method based on the correlation coefficient, which we have defined as

$$R(x, y) = \sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y')), \quad (2)$$

where $T'(x', y') = T(x', y') - (1/w \cdot h) \cdot \sum_{x'', y''} T(x'', y'')$ and $I'(x + x', y + y') = I(x + x', y + y') - (1/w \cdot h) \cdot \sum_{x'', y''} I(x + x'', y + y'')$.

$R(x, y)$ represents the similarity between the template (T) and the selected patch in the image (I). In general, the similarity measure can also be defined through other approaches such as square difference, normalized square difference, cross-correlation, and its normalized version, which computes a number within the broad allowable range, e.g., $R(x, y) \in [0, +\infty)$. However, it was found to be challenging to determine the reliability of these matching algorithms. On the other hand, the method based on the correlation coefficient offers a similarity measure within a nominal practical range, i.e., $R(x, y) \in [0, 1)$, which can be easily implemented within the context of the decision-making process.

Due to the movements of the patient, we anticipated that the size of the marker could also change. We defined a relationship that maintains the ratio of the marker within the ROI consistent. For example, when the patient is moving away from the robot, the marker gets smaller than the initial captured template. In this case, for a consistently sized template, the matched region within ROI would include

some of the surrounding backgrounds. As a result, even though the matched ROI contains the marker, it is not within the suitable region which can be utilized for further processing. To increase the accuracy, we experimentally determined to reduce the template size by 10% which can result in the similarity matching measure of greater than 0.5. Figure 5 shows the overall flow diagram of the proposed template matching algorithm.

2.2.2. Bhattacharyya Distance. Similar to the template matching algorithm, Bhattacharyya distance evaluates the difference between the two selected image patches. Colour histogram generation would be a time-consuming process for real-time tracking of the target, which may also cover the entire frame (especially when the size of the frame and the template are large). In this work, we have employed mean shift algorithm in conjunction with the Bhattacharyya distance computation. For each colour histogram, the number and range of colour coefficients must be equivalent for a given same size image. The flowchart of Bhattacharyya distance calculation combined with mean shift adjustment is shown in Figure 6(a). The flowchart summarizes operation on a single pixel which is in the region selected by the mean shift algorithm. For each region, the Bhattacharyya distance is computed multiple times (in case 3) in total, referring to target models in different sizes. This step is essential as it is the preparation for the selection of the size of the target model which results in significant similarity between the colour histogram of the target model and the ROI at the current frame.

2.2.3. Decision Making with respect to Template Matching and Histogram Comparison. Template matching and colour histogram comparison are two algorithms for estimating the ROI based on the region defined at the initial state. Both approaches have some advantages and disadvantages, which can be integrated as a part of the decision-making strategies. The proposed algorithm for template matching allows estimation of ROI over the entire frame by analyzing its effectiveness using a similarity measure. Template matching is a pixel-to-pixel comparison, and it has excellent performance for the recognition of rigid targets. Although ROI is meant to be defined on the patch containing the most consistent information associated with the design of the artificial marker, the information within ROI can still be distorted, resulting in a nonrigid ROI. We also implemented a colour histogram comparison. The comparison is the right approach, as it increases the reliability of the template matching. Since the computation of the colour histogram is exponential in time, we also applied a mean shift algorithm to estimate the probability of locating ROI at current frame. We assumed that the displacement of ROI is not large compared to the entire frame. Figure 6(b) summarizes the proposed decision-making process.

2.2.4. Tracking Deformable Marker. The previous section introduced an algorithm for tracking artificial marker on

patients who do not wear hospital gowns. The tracking marker attached to deformable objects such as hospital gowns introduces additional complexities in designing visual tracking algorithm. Compared to the first class of marker placements, here, the target model is more deformable (Figure 4(d)). We could not use such a marker as a valid distance estimation based on pinhole camera model. We show our design for a marker in Figure 4(d) as a strip with a sequence of red grids. This strip allows an increase in the probability of detection of a valid marker, which can be identified and used as a rigid one. Using similar strategies as before, all the red marked grids whose center points appear inside ROI at the current time frame are segmented. The segmented image is further analyzed and searched in order to identify a single grid with a relatively unchanged shape, which can be used for further tracking. In our case, for an unchanged shape, the ratio of the width to height is (5/4.6). Similar to our previous analysis, segmented grids are bonded together by a rotatable bounding box. The height here is the distance between the top-left corner and the bottom-left corner and the length is the distance between the top-left corner and the top-right corner.

Figure 7 shows an example of the detected markers without initialization of the ROI. Figure 7(a) shows the bounding areas around some of the detected markers. Figure 7(b) shows the threshold image of the segmented markers. We marked the estimated contours for these grids on the image. As long as the segmented area is large enough to be recognized as a grid, we identified and marked it by a green boundary. The tracking algorithm then analyzes all identified green boundary regions to compute and rank their similarities to the reference marker. The blue bounding region in the figure is an example of the selected reference marker, which we used for further distance computation.

For some cases with extreme deformability, it becomes challenging to determine the accurate distance between the gown and the camera. For example, the left three grids in Figure 8 are most likely have similar depth information, further away from the camera than that of right four grids. When two grids are overlapping, they are going to be combined and considered as a single grid (yellow bounding box).

When patients do not wear a hospital gown, we could not place markers. Here, the navigation commands for the robot depend on the position of the segmented marker with respect to the robot. For example, at the initial frame, an ROI is selected. We obtained the position of the ROI and the movements of the markers at the current frame, and then we estimated the direction of movement of the patient with respect to the robot. Due to the unique pattern structure of the marker applied on the hospital gown, the navigation commands associated with the first category do not directly apply to the second category mentioned above since the marker consists of several segmented grids. There we aimed to recognize the hospital gown to make a more accurate detection. The complexity arose when the position of the marker is inside the field of view of the camera. Typically, the hospital patient gown is in blue or white colour, which is very close to the overall colour of the hospital environment.

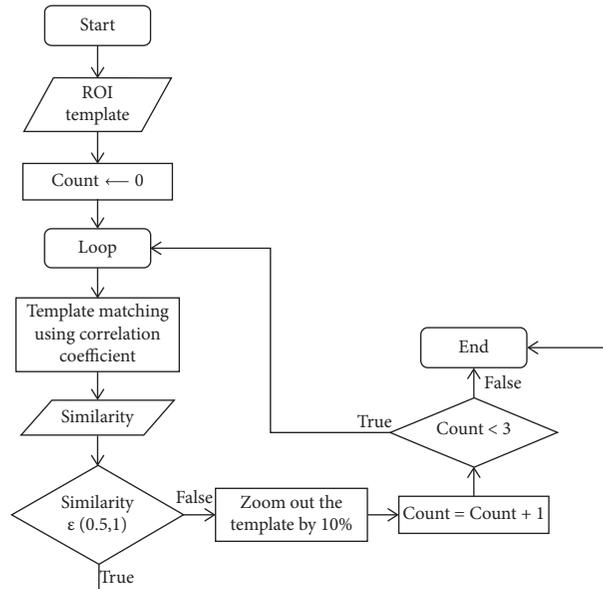


FIGURE 5: Flowchart of the modified template matching algorithm.

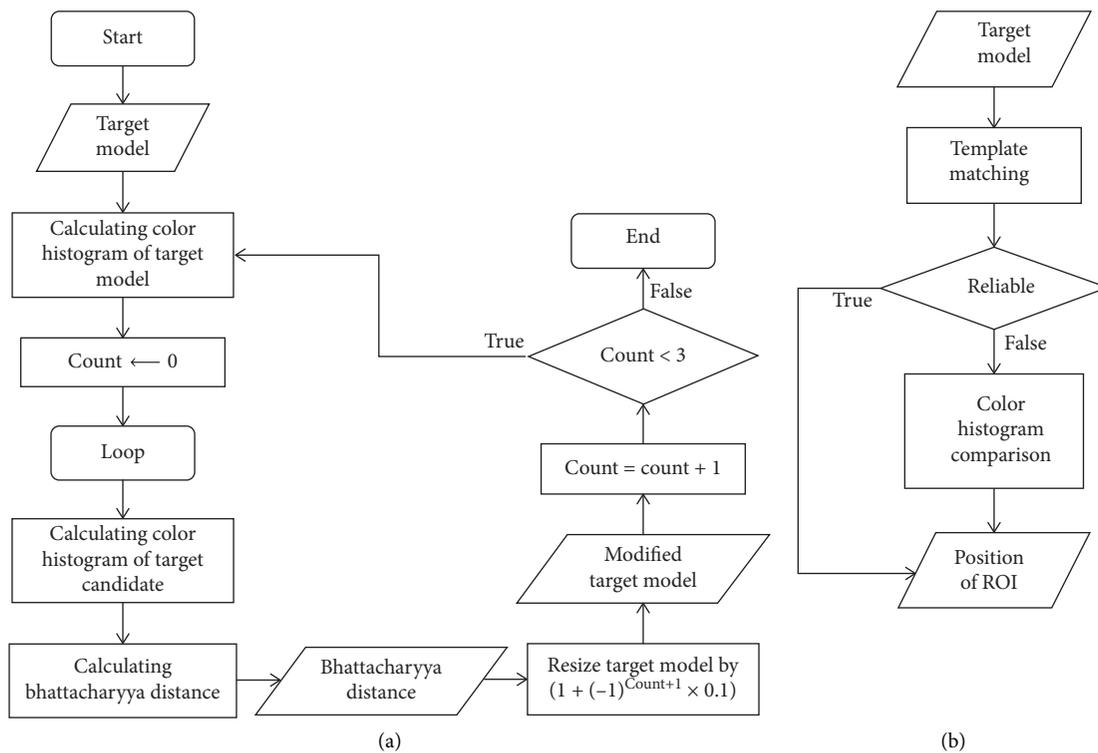


FIGURE 6: (a) Flowchart of histogram comparison algorithm applying Bhattacharyya distance and (b) decision making considering results for template matching and histogram comparison algorithms.

Here, the objective is to segment all grids inside the ROI defined through template matching. We used the average coordinate value of the grid whose centroid has the smallest value along the direction of the frame buffer and the grid whose centroid has the highest value to represent the current location of the patient. This information is further used by the trajectory controller of the robot. For example,

Figure 9(a) shows the overall bounding box of the ROI shown in black which contains all segmented grids. The blue box labeled as Grid 2 is the grid that is assigned to be the reference grid for distance estimation based on the pinhole camera model. The position of the patient is then assigned to be the average of x values of the two grids which we identified as Grid 1 and Grid 3 (Figure 9(b)).

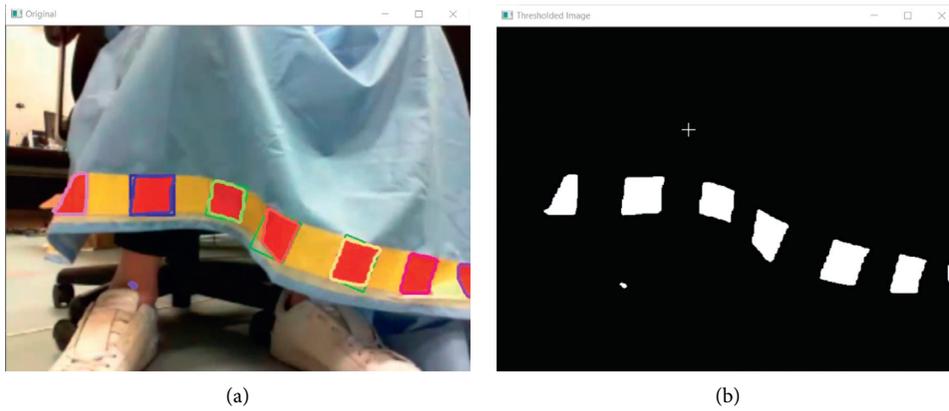


FIGURE 7: Results of the image processing when applying a reference selection algorithm for a deformable marker. (a) Results of grid recognition and the selected grid, which we highlighted with a blue boundary. (b) The result of thresholding.

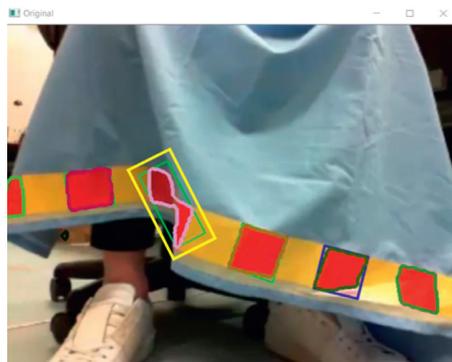


FIGURE 8: The case where two grids overlapped.

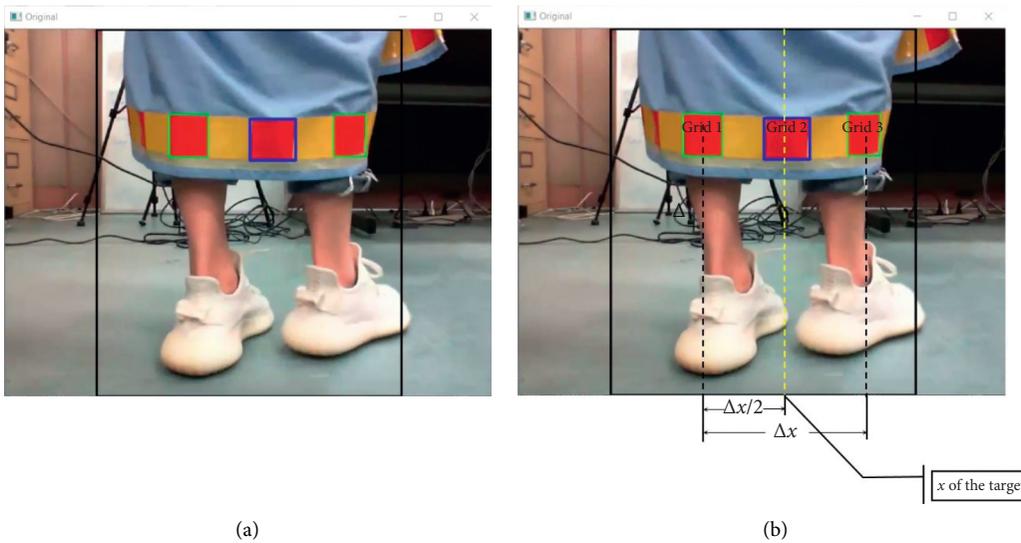


FIGURE 9: Description of horizontal displacement of the marker strip attached to the hospital gown: (a) the initial overall bounding box and (b) labeled grid locations and target definition.

3. Control for Trajectory Following

Previous sections highlighted some of the key components that were utilized as a part of the image processing for patient localization. This section presents the robot navigation algorithm for following a patient given the detected position information of a marker.

Let a representative position information of a marker with respect to the robot be defined as (P_x, P_y, P_z) (for convenience, this coordinate frame is defined in the image plane of the camera). Here, (P_x, P_y) is the estimated position of the segmented marker in the image buffer and its z coordinate is obtained through the pinhole camera model. Figure 10 shows an example of tracking information which we obtained for a patient wearing a foot band.

In the proposed trajectory control of the robot, the image of the marker is maintained about the center of the current frame. In any consecutive frame, the angular deviation θ and change in linear distance to the robot ΔD can be computed (Figure 10(b)). We then mapped these incremental changes to the angular velocity of the wheels.

A trapezoidal velocity profile is one of the practical profiles which most industrial type robotic applications utilize [11–13]. The profile can be divided into three segments corresponding to the acceleration, constant velocity, and deceleration phases. In order to reach to a desired position, initially the robot is accelerating, then remains at a constant speed, and finally decelerates toward the incremental goals. For example, the area under the profile is the total displacement of the robot, which is related to the linear deviation defined by the current position of the marker. One of the key advantages of this profile is that we can tune it for the drip-stand patient follower robot. For example, we can adjust it to have smooth acceleration and deceleration phases for avoiding the tipping action of the drip stand (Figure 11).

We propose a design of a controller for a robot to determine when the robot is to follow linear translational or rotational movements. In general, for a given definition of the trapezoidal profile and the incremental changes in the movements of the patient, the controller is either in acceleration or deceleration phases of the trajectory without reaching the constant velocity phase. To alleviate this, we have proposed a modified trajectory profile for both linear and angular movements of the robot (Figure 12). For example, the translational velocity profile is defined as

$$v = \begin{cases} v_{\max}, & z \geq d + e_1, \\ v_{\min} + \frac{v_{\max} - v_{\min}}{e_1 - e_0} (z - (e_0 + d)), & d + e_0 \leq z < d + e_1, \\ 0, & d - e_0 < z < d + e_0, \\ -v_{\min} + \frac{v_{\max} - v_{\min}}{e_1 - e_0} (z - (d - e_0)), & d - e_1 < z \leq d - e_0, \\ -v_{\max}, & z \leq d - e_1, \end{cases} \quad (3)$$

where d is the desired incremental distance of the patient with respect to the robot (Figure 10(b)). A similar description can be defined for the angular velocity profile as shown in Figure 12(b). For example, in the description of the profile in our study, x is the desired horizontal displacement of the marker in the image frame, and 320 is determined to be the biased value based on the properties of the camera and central pixel location.

4. Experimental Evaluation

We carried out a series of experiments in a laboratory environment to examine the proof of concept proposed in this paper. The environment layout consists of three connected corridors, where the subjects can walk. To simulate a typical environment, we placed standard furniture in the area. We used a single subject at the time, i.e., no additional subject was visible in the field of view. The experimental setup consists of a mobile platform equipped with an RGB camera and a depth sensor. Both the camera and the depth sensors are mounted on the robot closer to the ground level. There is also a camera placed in the experimental environment overlooking and recording the experiments. We recorded an external video to show the overall movements of the robot and the patient in the monitoring area. We used a depth sensor as an auxiliary sensing modality to compare and verify the distance computed through an RGB sensor, in the absence of any ground truth. We connected the robot to a wireless network through a central server. The central server is in the laboratory environment, which minimizes any delay in communication. We transmitted onboard information from the RGB camera and the depth sensors from the robot to the local host computer. We computed the required patient localization, and the desired robot movements based on the trajectory profiles are transmitted back to the robot.

We performed three sets of experiments as we describe in the following section with the subject wearing (a) non-visible hospital clothing; (b) hospital pants; (c) hospital socks; and (d) hospital gowns. Since the results of a patient wearing hospital socks are similar to the first case, we will not show the associated experimental results here. Subjects in the experiments first stepped in front of the robot to initiate the tracking. The patient following action of the robot is accomplished through previously described image processing and trajectory control profiles. The maximum and minimum values defined in Figure 12 are manually modified based on the expected speed of the patient.

4.1. Performance of Drip-Stand Patient Following Robot for the First Category of Hospital Clothing. For the first category, when the patient wears a marker band, hospital pants or hospital socks (as shown in Figures 4(a)–4(c)), similar performances were observed for the drip-stand follower robot. Here, we present a representative result for the case when the patient is wearing no hospital clothing (wearing marker band) and hospital pants.

Figure 13 shows the initialization of the tracking algorithm.

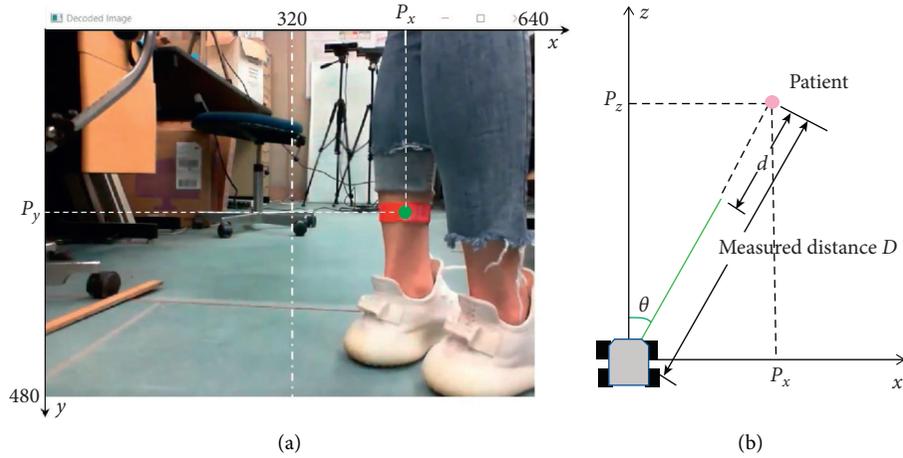


FIGURE 10: Position of the patient with respect to the robot coordinate frame. (a) Position of the detected marker in the projected image (x - y plane). (b) Position of the marker in x - z plane.

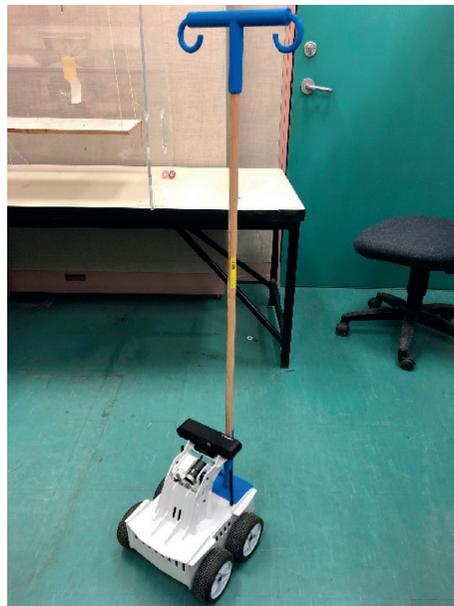


FIGURE 11: The prototype of the design of a robot following a patient's drip stand.

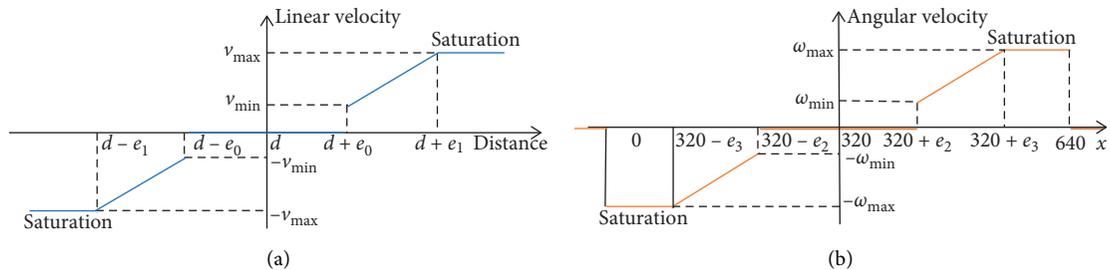


FIGURE 12: A modified velocity profile for the drip-stand patient following robot. (a) A linear translational velocity profile. (b) Angular velocity profile.

The patient wearing a marker band above the ankle stands in front of the robot at a distance (Figure 13(a)). The marker is segmented, and its distance to the robot is computed (Figure 13(b)). Figure 13(c) shows a segmented point cloud of the depth map obtained through a time of flight depth sensor.

We used this auxiliary measure of the distance of the patient to the robot for comparison purposes with the distance obtained through the pinhole camera model.

Figure 14 shows sample frames of the experimental study. Like the previous figure, the figure shows an

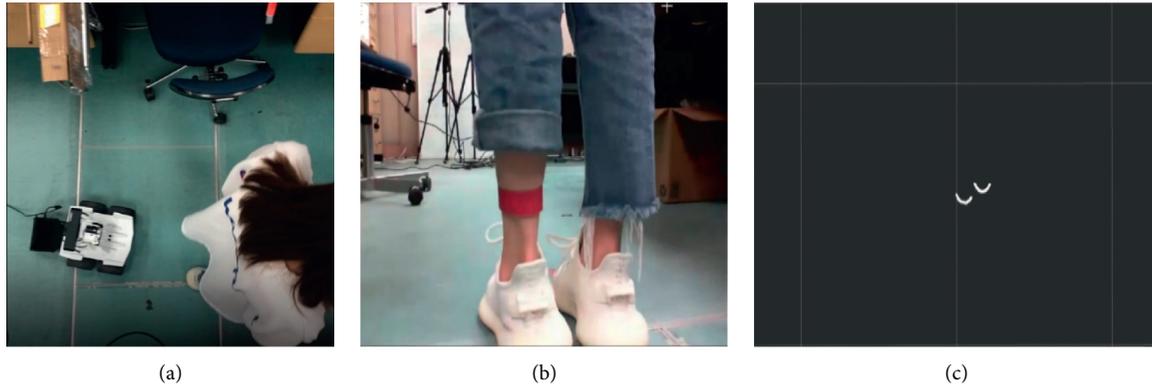


FIGURE 13: Initialization of the patient tracking algorithm. (a) The patient stands in front of the robot at a representative distance. (b) Detection of the marker using RGB camera and computation of initially required distance between the robot and the patient. (c) Point cloud representation of the depth map of the feet within the scanned plane.

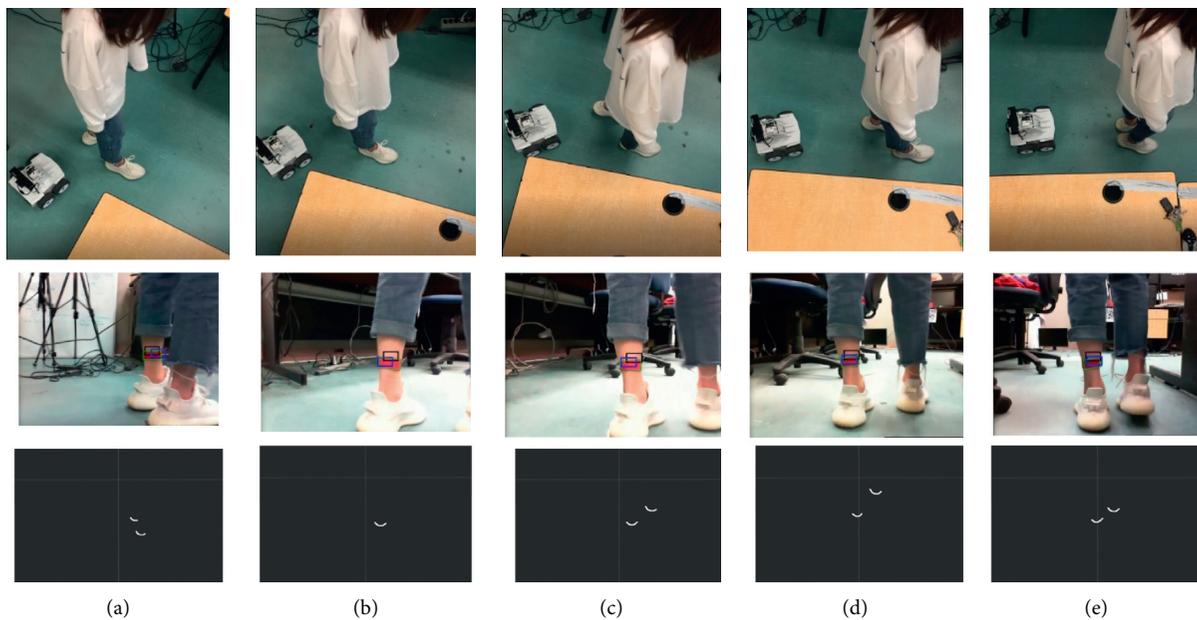


FIGURE 14: Examples of tracking frames of the drip-stand patient follower robot for the case of patient wearing nonhospital clothing (only wearing a marker band). Each sample column shows a view of the patient and the robot, the segmented marker for computation of the actual distance of the patient to the robot, and the results of distance computation using depth sensor co-located with the RGB sensor. (a) Sample frame 1. (b) Sample frame 2. (c) Sample frame 3. (d) Sample frame 4. (e) Sample frame 5.

image of the patient and the robot moving in the test area, the segmented marker, and the associated depth map.

To verify the accuracy of distance estimation algorithm using the pinhole camera model, we have compared the results with the actual distance measurement obtained through a depth sensor. We show these two datasets in the plots of Figure 15. We found that both methods have similar and consistent measurements. For example, if we assume that the data measured by depth sensor are an exact measure, the pinhole camera model can measure the depth information correctly. Most of the points, which represent the measurement error, aggregate in the range $[0, 0.99]$ (Figure 15(b)). From the collected data, we found that compared to pinhole camera model, depth measurement is

more sensitive to minor distance changes when the robot was moving. However, when the robot did not have either linear or angular adjustments (i.e., no changes in illumination condition in the viewing scene), the measured and computed data were more stable around the ideal distance. The maximum difference between these two methods (shown as a red point in the figure) is 2.4 cm.

Figure 15(a) shows the results of distance estimation between the robot and the patient for the case when the person wears marker band and for both pinhole camera model and depth sensor measurements. We set the desired distance between the robot and the patient to be at 40 cm. The maximum error occurred when the robot was compensating for angular displacement. In this experiment, we also verified that we could compensate through adjustment

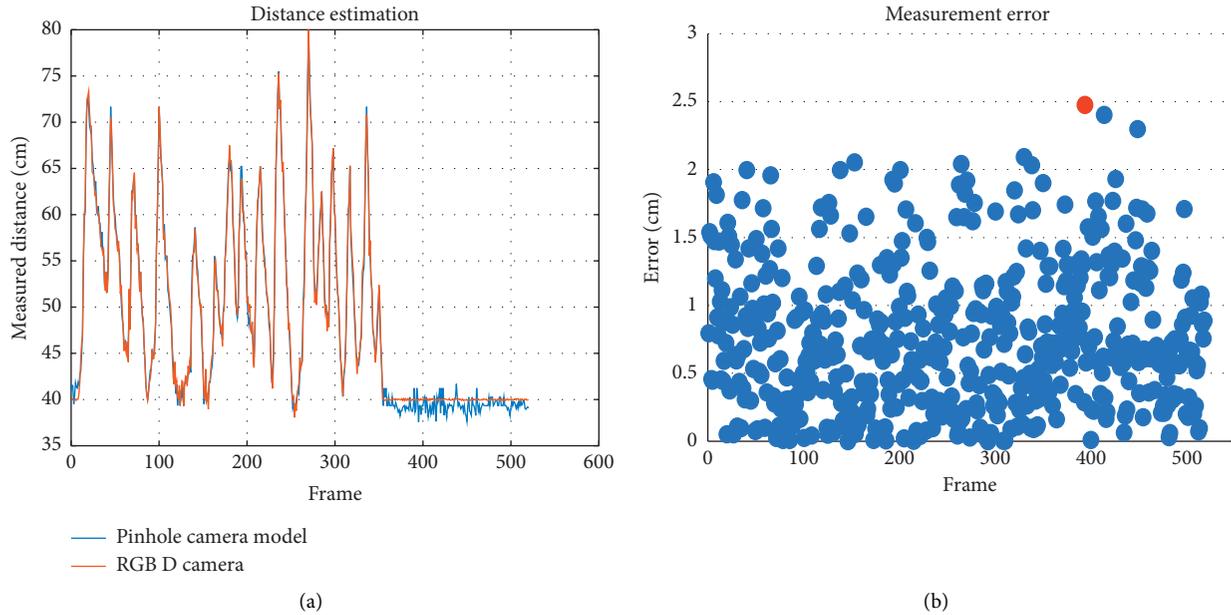


FIGURE 15: Tracking performance of the drip-stand follower robot for the case where the patient is wearing nonhospital clothing and a marker band. (a) Comparison between the pinhole camera model estimation and the depth sensor measurements. (b) Collection of measurement error during the tracking phase during the walking cycle between the pinhole camera model and the depth measurements.

of the robot maximum and minimum velocities defined in the trajectory profiles for when a patient was walking at a higher velocity.

We further evaluated the performance of the robot. We computed a set of trajectories from the movement commands of the robot and the estimated position of the subject. As we can see from the plots of Figure 16, overall, the robot follows the person accurately. However, when the robot received a command to make an angular adjustment, it did not always stay behind the subject. The robot started to rotate at the time when the person starts to turn along its path. The green point in the plot shows the location where such angular adjustments have occurred. The two red points indicate the place along the trajectory where there was an unexpected delay in the image transmissions. The delay was either caused by the speed of the connection or the long image processing time. To be specific, for the instances when the person starts to turn, the image buffer that the program is processing may be one frame behind. Hence, the robot would move forward for a short time (~0.125 seconds). However, we found that we can recover such delay once the processing time is less than 0.125 seconds. We compensated through the next frame adjustment, where the position deviation occurred, starting at the blue point. The robot trajectory deviates from the person trajectory due to the tolerance of the angular deviation.

Figure 17 shows the initialization of the tracking algorithm when the subject is wearing hospital pants. The patient stands in front of the robot at a distance (Figure 17(a)). The marker which is attached to the hospital pant is segmented and the distance of the marker to the robot is computed (Figure 17(b)). Figure 17(c) shows the segmented point cloud of the depth map obtained through a time of flight

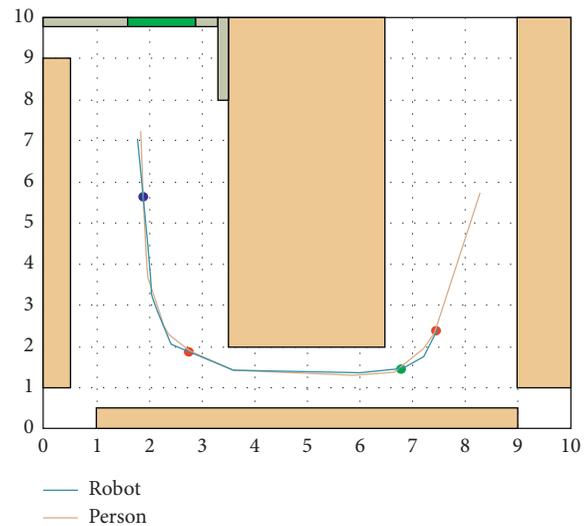


FIGURE 16: Trajectory of patient and the drip-stand robot for the case when the patient is in nonhospital clothing and wearing a marker band.

depth sensor. We used this auxiliary measure to compare with the distance obtained through the pinhole camera model.

Figure 18 shows sample frames of subject moving in the test area. Similar to the previous figure, it shows the image of a patient and the robot moving in the test area, the segmented marker, and the associated depth map.

Figure 19 shows the results of distance estimation between the robot and the person for the case when the person wears hospital pants. We set the desired distance between the robot and the patient to be at 40 cm. The maximum error occurred when the robot was compensating for angular

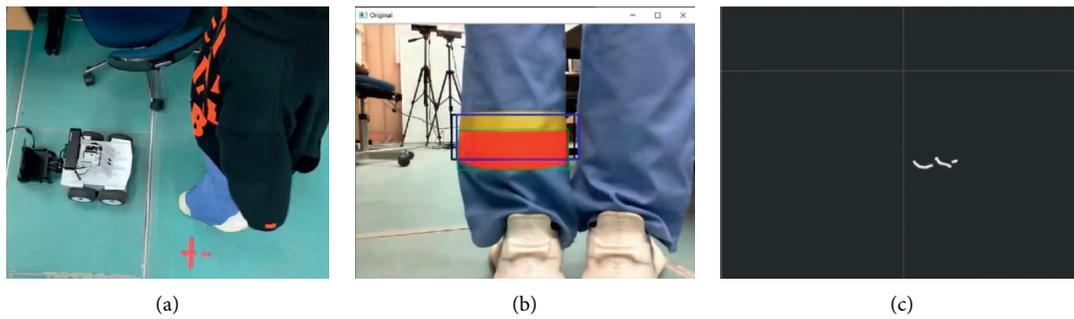


FIGURE 17: Initialization of the patient tracking algorithm. (a) The patient stands in front of the robot at a representative distance. (b) Detection of the marker using RGB camera and computation of the initially required distance between the robot and the patient. (c) Point cloud representation of the depth map of the feet within the scan plane.

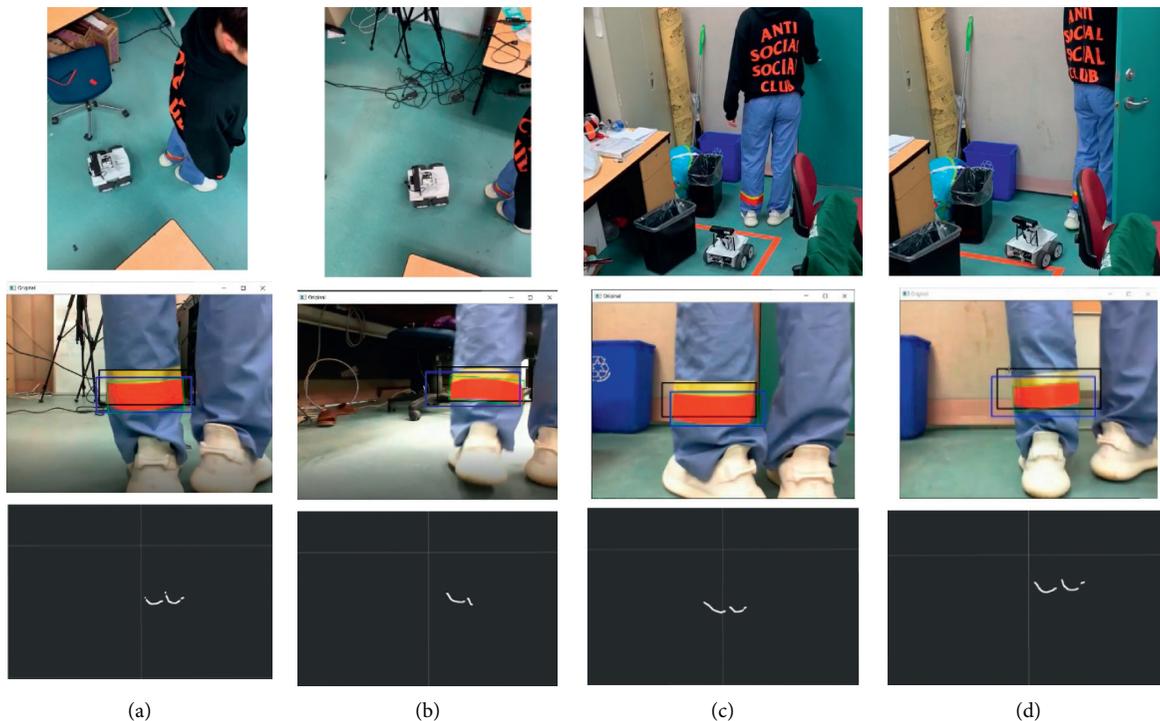


FIGURE 18: Examples of tracking frames of the drip-stand patient follower robot for the case of hospital pants. Each sample column shows a view of the patient and the robot, the segmented marker for computation of the actual distance of the patient to the robot, and the results of distance computation using depth sensor co-located with the RGB sensor. (a) Sample frame 1. (b) Sample frame 2. (c) Sample frame 3. (d) Sample frame 4.

displacement. In this experiment, we also studied how we can compensate for the faster walking velocity of the patient through adjustment of the robot maximum and minimum velocities defined in the trajectory profile. Also, in some cases and due to the gait pattern of the legs, the pants get wrinkled, which can cause an inconsistency between the two marker detection methodologies. In these cases, the detected marker is not close to a rectangular shape. As a result, the detected region would be larger than the actual marker. This situation happened numerous times throughout the entire tracking.

Figure 20 shows the overall movement trajectory from the starting point to the time when the patient exits the experimental room through the door. Here, we can see the

robot performing its tasks satisfactorily. The two red marks indicate the time when there was a delay in the transmission of images. In the same figure, blue marks indicate instances when the angular deviations of the patient from the center part of the image are within the acceptable range.

4.2. Performance of Drip-Stand Patient Following Robot for the Second Category of Hospital Clothing. When the patient stands in front of the robot, the initialization process is triggered. As a part of this initialization, the computed distance between the robot and the patient is also defined (Figure 21). At each frame, we estimated the position of the

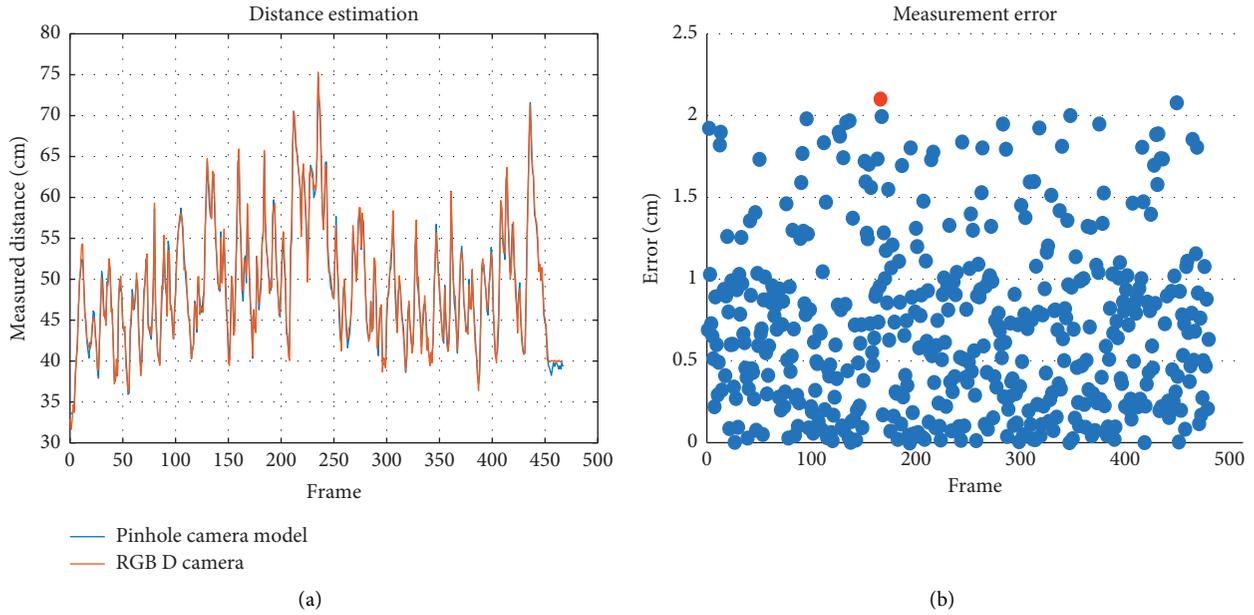


FIGURE 19: Tracking performance of the patient follower drip-stand robot for the case where the subject is wearing hospital pants. (a) Comparison between the pinhole camera model and the depth sensor measurements. (b) Collection of measurement error during the tracking phase between the pinhole camera model and the depth measurements.

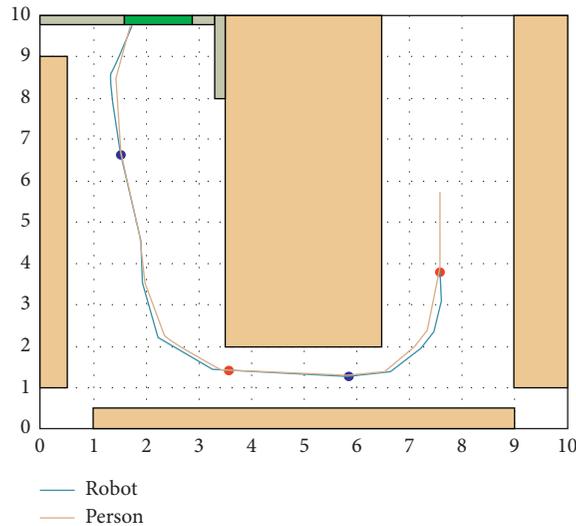


FIGURE 20: Trajectory of patient and the drip-stand patient follower robot for the case when the patient is wearing hospital pants.

patient by referring to the two side grids representing the actual location of the person.

We showed the results of image processing for various sample frames in Figure 22. For hospital gowns, the use of redundant marker grids obtained a better estimate of the robot.

Overall we found the performance of distance measurement to be consistent with other category of markers shown in Figure 4. However, we can expect that the distance error is the largest among all other types of markers. Depth image processing retrieves the average value of distances of pixels on the segmented contour on the gown, which is approximately the position highlighted by the green point shown in Figure 23. RGB image processing only selects a grid

that is not distorted by the effect of wrinkles in the cloth. In this case, the reference grid can either be the closest one to the camera or furthest away. In both cases, the results are not consistent with the one measured by the depth sensor. In Figure 23, this difference between two measurements is identified by Δd , which we referred to as the measurement error of the method based on the pinhole camera model.

Figure 24(a) shows the overall performance results of the drip-stand patient follower robot for the case of the hospital gown. Compared to the other three cases, we observed that the frequency that the distance is less than the desired value is higher, and the maximum measurement error is also the largest. This deviation usually happened while the patient is

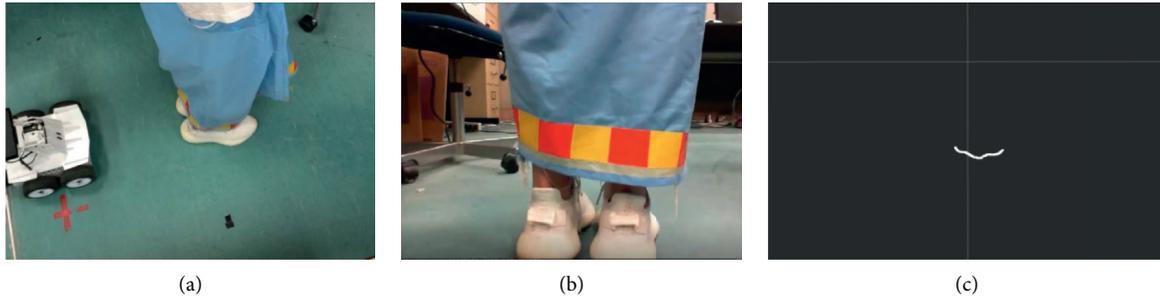


FIGURE 21: Initialization phase of the drip-stand patient follower robot when wearing hospital gown. (a) The patient stands in front of the robot at a representative distance. (b) Detection of the marker using RGB camera and computation of the initially required distance between the robot and the patient. (c) Point cloud representation of the depth map of the gown within the scan plane.

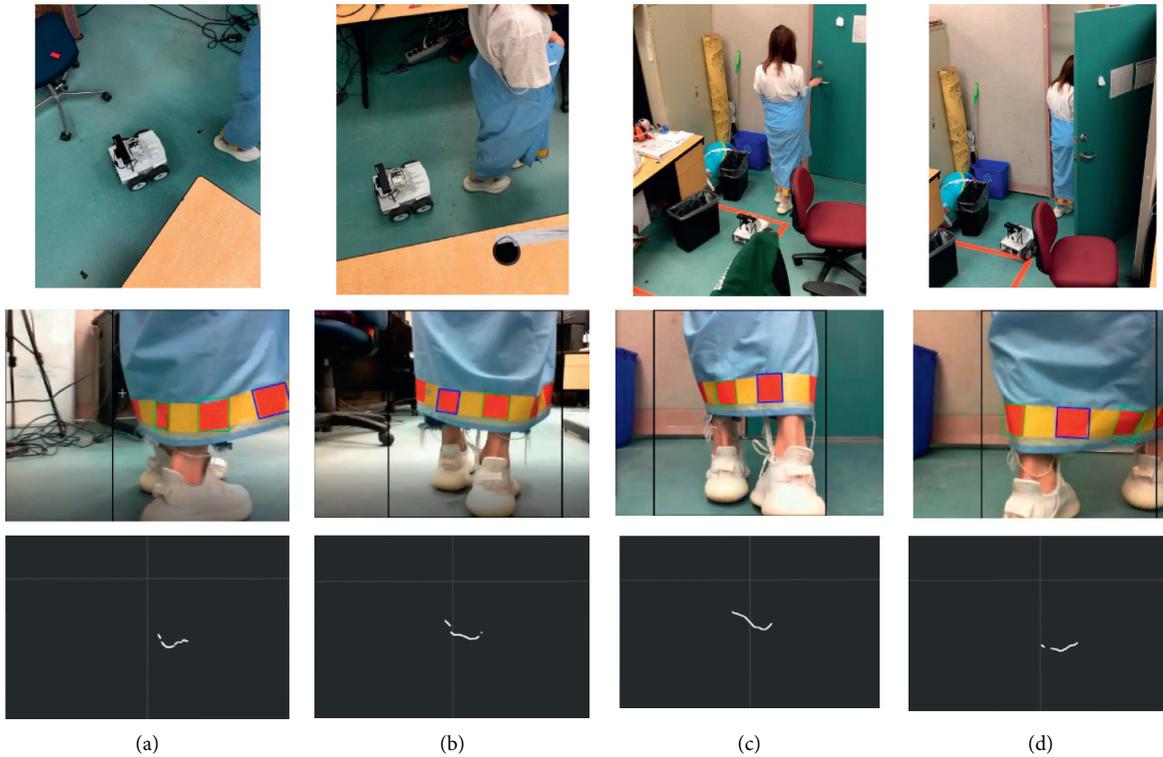


FIGURE 22: Example of tracking frames of the drip-stand patient follower robot for the case of the hospital gown. Each sample column shows a view of the patient and the robot, the segmented marker for computation of the actual distance of the patient to the robot, and the results of distance computation using depth sensor co-located with the RGB sensor. (a) Sample frame 1. (b) Sample frame 2. (c) Sample frame 3. (d) Sample frame 4.

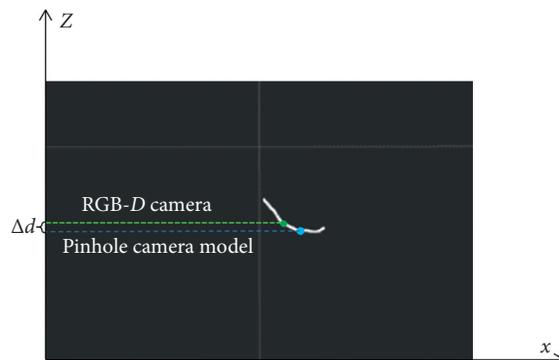


FIGURE 23: Example of computation of relative error in the marker-based RGB distance computation with the measurement obtained through depth sensor.

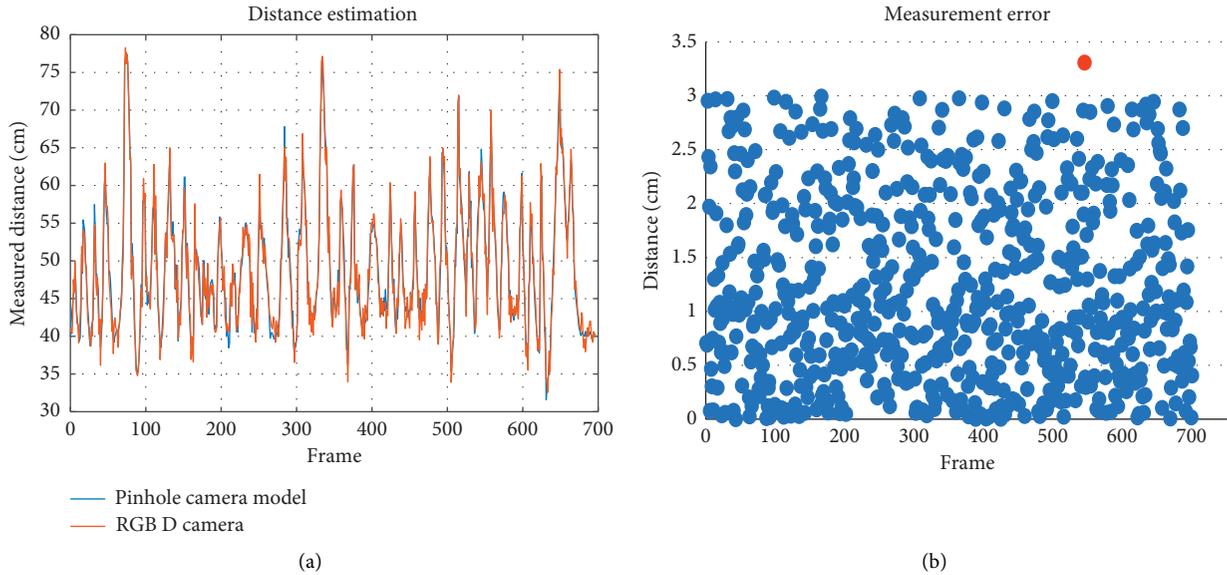


FIGURE 24: Tracking performance of the patient follower robot. (a) Comparison between the pinhole camera model and the depth sensor measurements. (b) Collection of measurement error during the tracking phase.

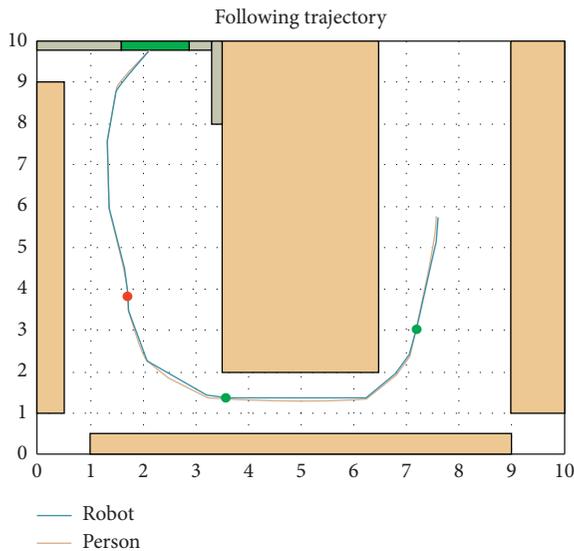


FIGURE 25: Trajectory of the patient and the drip-stand robot for the case when the patient is wearing hospital gown.

turning, and the robot would move too close to the patient and then move back to increase the distance. Also, different gait patterns can change the shape of the gown. In this case, the grid that was selected first as the reference grid for the pinhole camera model is relatively at the furthest distance among all segmented grids (e.g., see sample frame 1 of Figure 18). Hence, the robot controller continues setting reference commands for the robot, even if it is too close to the patient. Where the grid that is much closer to the camera as the reference grid, we observed the worst performance. The robot controller is now defining a new reference set point for the robot to increase its distance.

Figure 25 shows the overall performance of a drip-stand patient follower robot for this case.

5. Conclusions

This paper proposed a design of a drip-stand patient following robot. The design consists of a mobile platform with a connected supporting rod that holds the intravenous medications. We integrated a standard RGB camera located close to the floor level in our design. We can further preserve the privacy of the patient due to the constrained field of view and the proximity of the drip stand. The tracking information is obtained through visual processing of artificial markers placed on hospital clothing. Various designs of such placements on various types of hospital clothing are suggested, and performance of the drip-stand patient following robot was experimentally evaluated. To avoid the tipping action of the drip-stand patient follower robot, we proposed a modified trajectory profile for the mobile platform. This enhancement limits the velocity command and the associated maximum magnitude of the acceleration and deceleration.

The proposed novel design allows the robot to follow the patient from behind. The robot always keeps the designated distance to the patient, which is defined by the allowable slacks introduced by the intravenous medication tube. We can also modify the design to allow the robot to follow the patient from the side. Other researchers proposed similar design objectives. For example, the design in [14] suggests the robot following the patient by merely pulling a tether attached to the robot. To the best of our knowledge, our proposed design is one of its kind, which we can further enhance to explore its wide-spread acceptance.

Data Availability

The data generated during the experimental evaluation and used in the current study are available upon request.

Conflicts of Interest

The authors declare that they do not have any conflicts of interest.

References

- [1] M. Najib and M. Fahmi, "A study on techniques of person following robot," *International Journal of Computer Applications*, vol. 125, no. 13, pp. 27–30, 2015.
- [2] B. X. Chen, R. Sahdev, and J. Tsotsos, "Person following robot using selected online ada-boosting with stereo camera," in *Proceedings of the 14th Conference on Computer and Robot Vision*, pp. 48–55, Edmonton, Canada, May 2017.
- [3] B. Ilias, S. A. Abdul Shukor, S. Yaacob, A. H. Adom, and M. H. Mohd Razali, "A nurse following robot with high speed kinect sensor," *Journal of Engineering and Applied Sciences*, vol. 9, no. 12, pp. 2454–2459, 2014.
- [4] Aethon Automates Intralogistics, TUG Smart Autonomous Mobile Robot, 2020, <http://www.aethon.com/tug>.
- [5] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, vol. 2, pp. 674–679, Vancouver, Canada, 1981.
- [6] L. Shen, X. Huang, Y. Yan, and S. Bai, "An improved mean-shift tracking algorithm with spatial-colour feature and new similarity measure," in *Proceedings of IEEE International Conference on Multimedia Technology*, pp. 184–188, Hangzhou, China, July 2011.
- [7] L. Susperregi, J. M. Martínez-Otzeta, A. Ansuategui, A. Ibarguren, and B. Sierra, "RGB-D, laser and thermal sensor fusion for people following in a mobile robot," *International Journal of Advanced Robotic Systems*, vol. 10, no. 6, p. 9, 2013.
- [8] W. Choi, C. Pantofaru, and S. Savarese, "Detecting and tracking people using an RGB-D camera via multiple detector fusion," in *Proceedings of IEEE International Conference on Computer Vision Workshops*, pp. 1076–1083, Barcelona, Spain, November 2011.
- [9] S. Payandeh, *Visual Tracking in Conventional Minimally Invasive Surgery*, CRC Press, Boca Raton, FL, USA, 2016.
- [10] S. Tanimoto, *An Interdisciplinary Introduction to Image Processing*, MIT Press, Cambridge, MA, USA, 2012.
- [11] Trinamic Motion Control, 2020, https://www.trinamic.com/fileadmin/assets/Support/Appnotes/AN013Motion_control_profile_parametrization.pdf.
- [12] Parker Motion Control, 2020, https://www.parkermotion.com/manuals/a-ax/AX_Ch4p1.pdf.
- [13] Baldor Motion Products, http://www.multiprojekt.pl/ftp/abb_serwonapedy/archiwum%20baldor/application_note/an00118-000_motion_profiles_EN.pdf.
- [14] G. Endo, B. Allan, Y. Lemura et al., "Mobile follower robot as an assistive device for home oxygen therapy evaluation of tether control algorithms," *ROBOMECH Journal*, vol. 2, no. 1, 2015.