2016 IEEE International Conference on
Automation Science and Engineering (CASE)
Fort Worth, TX, USA, August 21-24, 2016

# Enhancing Robotic Unstructured Bin-Picking Performance by Enabling Remote Human Interventions in Challenging Perception Scenarios

Krishnanand N. Kaipa[1], Akshaya S. Kankanhalli-Nagendra[1], Nithyananda B. Kumbla[1], Shaurya Shriyam[1], Srudeep Somnaath Thevendria-Karthic[1], Jeremy A. Marvel[2], and Satyandra K. Gupta[1]

*Abstract*— **We present an approach that enables a robot to initiate a call to a remote human operator and ask help in resolving automated perception system failures during bin-picking operations. Our approach allows a robot to evaluate the quality of part recognition and pose estimation, based on a confidence-measure, and thereby determine whether to proceed with the task execution or to request assistance from a human in resolving the predicted perception failure. We present an automated perception algorithm that performs the joint task of part recognition and 6 degree-of-freedom pose estimation, and has built-in features to initiate the call to the human when needed. We also present the underlying mechanism for a rationalized basis for making the call to the human. If uncertainty in part detection leads to perception failure, then human intervention is invoked. We present a new user interface that enables remote human interventions when necessary. We report results from experiments with a dual-armed Baxter robot to validate our approach.**

## I. INTRODUCTION

Deploying robots in industrial applications requires high reliability of robotic task execution. This is accomplished by designing specialized hardware and software. Extensive system testing is needed to ensure all failure modes are understood and contingency plans are developed to handle them. Task execution failures typically require the assembly/manufacturing line to be paused if the fault is unrecoverable, and human intervention to clear the fault and restart the line. This type of intervention can be expensive, and hence robots are not used on a task until high level reliability can be achieved. Customized hardware and software costs can only be justified if the production volume is sufficiently high (e.g., automotive assembly lines).

Currently, robots have difficulty in assessing their own capability to complete a task. Consider the following case. A robot is capable of picking a part if it is presented to it at a certain location. However, if the part has shifted from its nominal location, the robot might not be able to pick it. The robot simply does not know where the transition boundary between task execution success and failure lies. As it attempts to pick the part, it might bump into it and push it further and jam the part into other parts. This can trigger a system fault and shut down the system.

To use robots in small production batch operations, robots are needed that can estimate the probability of task completion before beginning the task. This will enable robots to assess their confidence in doing a task. If the robot does not have a high confidence in completing a task, then it can call for help. This will enable human operators to provide the robot the needed assistance (e.g., better part pose estimation, invoke a different grasping strategy) and prevent major system faults that result from task execution failure. Providing task assistance help to robots is cheaper than recovering from a system shutdown.

These concepts are illustrated using robotic bin picking example in this paper. The experimental setup (Fig. 1) is based on RoboSAM, a ROBOtic Smart Assistant for Manufacturing. The RoboSAM system is built using a Baxter robot[1] and an Ensenso 3D camera. We are focused on a class of problems that manifest in the form of a part-order specifying multiple quantities of different parts to be singulated from a bin of randomly scattered pile of parts and transported to a destination location as rapidly as possible. Achieving this overall goal entails overcoming important challenges at various stages of task execution including part recognition and pose estimation, singulation, transport, and fine positioning. The singulation task involves picking only one part at a time.

This paper is focused on part recognition and pose estimation. This problem is challenging and still not fully solved due to conditions commonly found in factory environments [1], [2]. In particular, unstructured bins present diverse scenarios affording varying degrees of part recognition accuracies: 1) Parts may assume widely different postures, 2) parts may overlap with other parts, and 3) parts may be either partially or completely occluded. The problem is compounded due to factors like sensor noise, background clutter, shadows, complex reflectance properties of parts made of various materials, and poorly lit conditions. All these factors result in part recognition and pose estimation uncertainties.

In this paper, we present an approach that enables a robot to initiate a call to a remote human operator and ask help in resolving automated perception system failures. Our approach allows a robot to evaluate the quality of part recognition and pose estimation, based on a confidence-measure, and thereby determine whether to proceed with the task execution or to seek assistance from a human. In our previous work [3], we presented a preliminary approach to address this problem. We mainly studied the feasibility of a remote

[1]Maryland Robotics Center, University of Maryland, MD USA skgupta@umd.edu
[2] National Institute of Standards and Technology, Gaithersburg, MD USA jeremy.marvel@nist.gov

[1]DISCLAIMER: Any commercial product or company name in this paper is given for informational purposes only. Their use does not imply recommendation or endorsement by NIST or the University of Maryland.

978-1-5090-2408-7/16/$31.00 ©2016 IEEE 

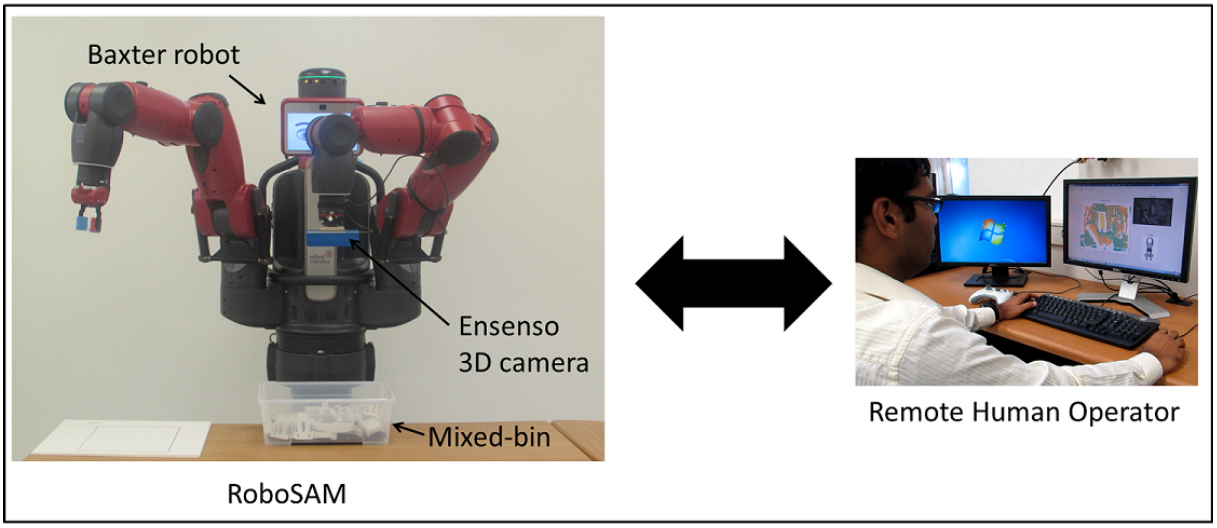


Fig. 1. Hybrid Cell showing the RoboSAM system built using the Baxter robot and the Ensenso 3D camera and the remote human operator

human using a simple joy-stick controlled graphical interface to perform part-matching and extract postural information that can be used by the robot to successfully singulate the desired part. However, the system lacked the ability to detect a perception failure and initiate a call to the human. This paper builds on our previous work and provides three new features:

1) A method to characterize uncertainty in pose estimation of a part match found by using an automated perception system.
2) A mechanism for the rationalized basis for making the call to the human
3) A new user interface that allows the remote human to provide distinguishing cues to resolve the part matching problem

## II. Related Work

Many research groups have addressed the problem of enabling robots, guided by machine-vision and other sensor modalities, to carry out bin-picking tasks [1], [3], [4], [5], [6]. Different aspects of robotic bin-picking include perception, grasp-planning, and motion planning. Each of these problems represents a vast area of research in itself and is usually treated separately.

Most previous attempts on a systems approach to bin-picking mainly focussed on the perception problem [1], [7], [8], [9], [10], [11], while assuming accurate robot grasping. However, model inaccuracies and sensor uncertainties make it difficult for a majority of the perception algorithms to provide reliable object recognition and localization estimates, thereby affecting overall bin-picking performance.

Given that pose estimation error impacts grasping performance in practice, many researchers have addressed the problem of grasp planning under perception uncertainty [12], [13], [14], [15], [16] and uncertainty in object shape due to manufacturing tolerances, and mechanics, due to limits on sensing during grasping [17]. However, perception failures are not explicitly addressed in most of the above approaches.

The robotic bin-picking system developed by Fuchs et al. [18] has built-in mechanisms to detect object localization failures. In particular, they assume significant uncertainty in object pose estimation and initiate grasping only when the reliability of the pose hypothesis falls below a given threshold. Otherwise, the localization is restarted from different view point of the camera. Our work is similar in terms of the perception failure detection capability. However, our approach differs in how the system responds to the failure. That is, rather than attempting the perception task again, the system calls a remote human for assistance.

Another relevant work is an algorithm, presented by Pronobis and Caputo [19], which measures its own level of confidence in performing a visual place recognition task. Taking a *support vector machine* approach, the authors propose a method for measuring the confidence level of the classification output based on the distance of a test image and the average distance of training vectors.

## III. Problem Formulation

Let $\ell \in \mathbb{R}^6 = \{x,y,z,\alpha,\beta,\gamma\}$ represent a general posture where $(x,y,z)$ and $(\alpha,\beta,\gamma)$ represent the position and orientation, respectively in three dimensions (3D).

**Definition 1.** A *mixed-bin* $\mathscr{B}(\kappa,n,\{n_i\})$ is a bin of randomly scattered pile of $n$ parts, comprising different multiple instances $n_i$ of $\kappa$ different part types:

$$\begin{aligned} \mathscr{B}(\kappa,n,\{n_i\}) &= \{p_i^{(j)} : j=1,\dots,n_i, i=1,\dots,\kappa\} \quad (1) \\ |\mathscr{B}| &= \sum_{i=1}^{\kappa} n_i = n \end{aligned}$$

where, part $p_i^{(j)}$ represents the $j^{\text{th}}$ instance of part type $i$.

**Definition 2.** *Position-gripper* is defined as an action performed by the robot to position its gripper at an appropriate posture above the bin just before approaching a part to be grasped.

**Definition 3.** *Approach-part* is defined as an action performed by the robot to move the gripper toward, and encompass, the part just before grasping takes place.

**Definition 4.** We say that a gripper *encompasses* a part when the intersection of the volume between the fingers with that of the part is non-zero and squeezing both the fingers in the pinch-direction results in a force-closure grasp.

**Definition 5.** *Grasp* refers to the act of grasping a part encompassed by the fingers. Note that by definition, force-closure is used as a constraint to evaluate candidate grasps.

**Definition 6.** *Extract* refers to the act of picking up a grasped part from the bin.

**Definition 7.** *Singulation* $\varepsilon$ is defined as the concatenation of the four stages of positioning the gripper, approach, grasping, and extraction.

**Definition 8.** A *Singulation plan* consists of a sequence of grasp postures used to singulate a part.

**Definition 9.** We define *tangle-free-singulation* $\varepsilon_{tf}$ as a singulation of a part from a bin such that it is not tangled with other neighboring parts in the bin during extraction, thereby ensuring singulation of only one part at a time.

Now, we formulate our problem as follows: Given a mixed bin $\mathscr{B}(\kappa, n, \{n_i\})$ and a desired part $p_d$ to be singulated: (1) find a part instance $p_d^j \in \mathscr{B}(\kappa, n, \{n_i\})$ with maximum probability of successful recognition $\wp(p_d^{(j)}|\mathscr{D})$, where $\mathscr{D}$ is the point-cloud data of the current scene, (2) find an estimate of its posture $\hat{\ell}_d^{(j)}$ and postural uncertainty $\sigma_d^{(j)}$ with a given confidence $\rho_d^{(j)}$, and (3) using the results from step 2, devise a mechanism to determine whether to proceed with the singulation task or seek human assistance in resolving a perception failure.

## IV. Approach

Given a CAD model of the desired part to be singulated and the three dimensional (3D) point cloud of the mixed-bin, the robot attempts to use an automated perception system to jointly solve the problem of identifying an instance of that part in the bin and its 6D posture. The system estimates its confidence in the part matching result, and thereby determines whether the robot can proceed with the task execution (part singulation), or to request help from a human in order to resolve the predicted failure. We have developed a new user interface that allows a remote human to perform pose estimation in scenes with high clutter where automated perception system may fail.

### A. Automated Perception Algorithm

Given a CAD model of the desired part to be singulated and the 3D point cloud of the mixed-bin, the automated perception system attempts to jointly solve the problem of identifying an instance of that part in the bin and its 6D posture. Let $\mathscr{P} = \{p_i : p_i \in \mathbb{R}^3,\}$ be the point cloud of the bin of parts captured from the 3D sensor. Let $\mathscr{Q} = \{q_i : q_i \in \mathbb{R}^3\}$ be the point cloud obtained by uniform surface sampling of the CAD model of the part to be identified. Our approach consists of extracting features (e.g., edges) available in the sensed data and exploiting these features to collapse the problem from a 6D search to a finite number of line searches. Feature extraction [20], [21], [22] is one of the preprocessing procedures used in many scene reconstruction tasks. The extracted features help in docking the CAD model of the desired part at possible postures in the point cloud of the scene where a part match is likely to be found. The algorithm steps are given below:

1) Estimate surface normals at each point in the point cloud
2) Cluster surface normals into a Gauss map to recognize planes
3) Use intersection of planes to extract oriented edges
4) For each oriented edge
    a) Align the part CAD model along the oriented edge
    b) Filter the CAD model to contain only the points perceivable from the camera for that orientation of the CAD model
    c) Obtain a part match by moving the filtered CAD model $\mathscr{Q}_f$ along the edge where it is docked as a function of a translation parameter $s$, and finding the $s^*$ that minimizes the mean point-to-point distance $\rho$ from the filtered CAD model to the point cloud from the sensor.

$$\rho = \min_s \frac{1}{|\mathscr{Q}_f|} \sum_{i=1}^{|\mathscr{Q}_f|} d(q_i, \mathscr{P})$$
$$\text{where}, d(q_i, \mathscr{P}) = \min_j ||q_i - p_j||, q_i \in \mathscr{Q}_f, p_j \in \mathscr{P}$$

5) Select the match that minimizes $\rho$.

Figure 2 shows the matching results by running the algorithm on some representative bin scenarios. In particular, this experiment reveals how the matching performance changes as a function of bin complexity−parts of same type not touching with each other (Fig. 2(a, b)), parts of same type overlapping with each other (Fig. 2(c, d)), and parts of different type overlapping with each other (Fig. 2(e, f)). Figure 3 illustrates a bin scenario that results in a part matching failure, where the desired part model (highlighted) is localized erroneously.

### B. Confidence Estimation

We estimate the confidence in the part matching result of the perception algorithm by using a signature based method. This involves obtaining four quantities: (1) ideal part match signature, (2) reference signatures based on synthetically generated point clouds, (2) probability distribution of dissimilarity between ideal and reference signatures, and (4) observed signature based on the test point cloud.

Given a sample point cloud of a single part and its CAD model, a part match signature is defined as the fraction

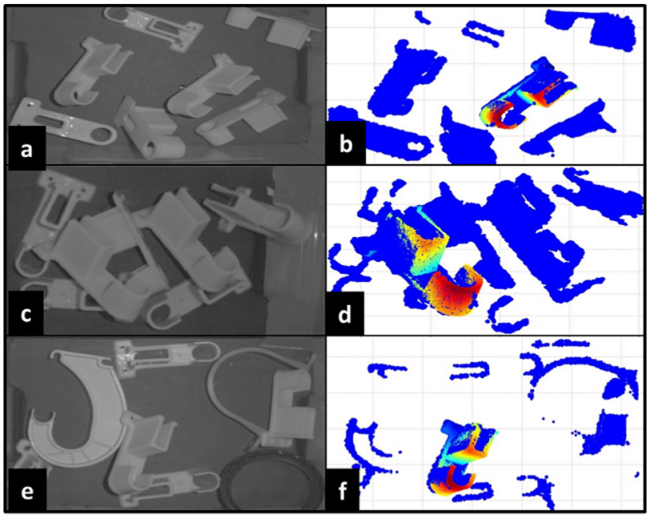


Fig. 2. Representative bin scenarios and corresponding matches: (a, b) Multiple parts of same type not touching with each other. (c, d) Multiple parts of same type overlapping with each other. (e, f) Multiple parts of different type overlapping with each other.

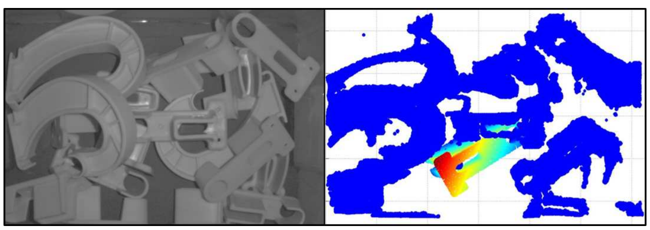

Fig. 3. Bin scenario that results in a part matching failure.

of points $\xi$ for which the minimum point-to-point distance $d(q_i, \mathscr{P})$ given in equation (2) is below a threshold distance $d_t$, plotted as a function of $d_t$. Note that this is a monotonically non-decreasing function.

The ideal signature is generated by performing calibration experiments to obtain the sensor noise model. Note that points from a sampled CAD model are used in the computation of $\rho$, which degrades the approximation of true $\rho$. To address this issue, we use a perfect cuboid-shaped object (Fig. 4(a)) in the calibration experiments. The CAD model of the object can be approximated by orthogonal planes. This

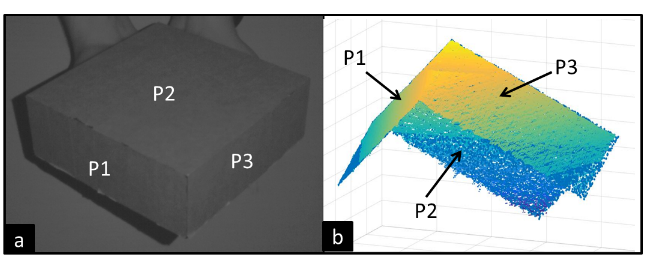


Fig. 4. (a) Object used for calibration. (b) Match obtained between the point cloud of the scene and the filtered CAD model

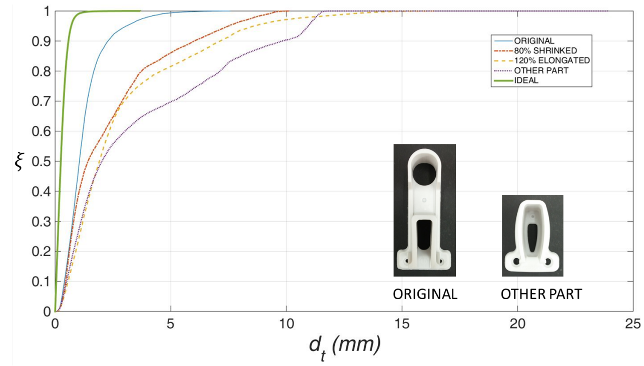


Fig. 5. Different signature curves: Ideal signature, signatures for original real part and its modified versions, and signature for a different part.

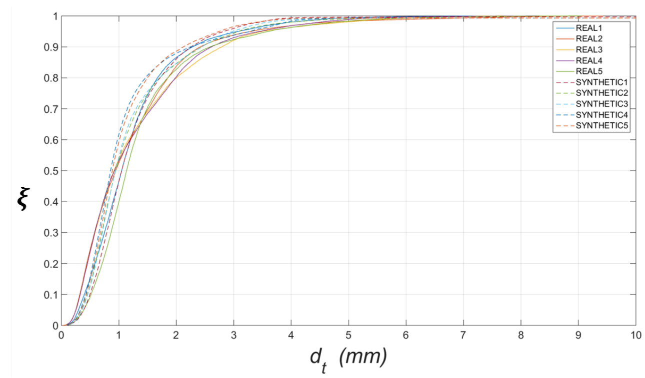


Fig. 6. Signature for a synthetic part along with signatures for a real part in 10 different postures

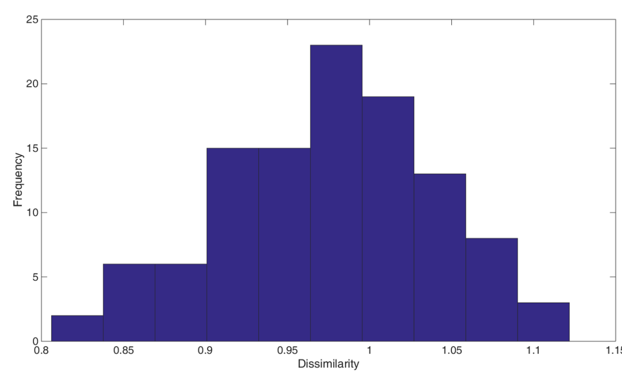


Fig. 7. Probability distribution of dissimilarity approximated as a normal distribution with mean $\mu$ = 0.9751 and standard deviation $\sigma$ = 0.0659 based on the histogram of dissimilarities

enables computing point-to-plane distances, which gives a better approximation of $\rho$ by isolating the sampling noise and discretization error and only accounting for sensor noise. The experiment is performed by placing the object in the scene such that three orthogonal planes are exposed to the sensor and obtaining a point cloud. Next, the automated perception algorithm described above is run to match the point cloud with the plane-fitted CAD model. The match is shown in Fig. 4(b). Now, $d(q_i, \mathscr{P})$ is computed as the minimum point-to-plane distance and used to generate an ideal part match signature.

Figure 5 shows an ideal signature and part match signature obtained by placing a real part in the scene. Note from the figure that the signature deviates as the part is modified (80% shrunk and 120% elongated). Also, the part match signature changes significantly for a different part. The dissimilarity of each part match signature from the ideal signature can be obtained by computing the corresponding difference in the area-under-the-curve of the two signatures.

Next, we must model the probability distribution of dissimilarity for a given part. First, a reference signature for the part of interest is obtained based on a synthetic point cloud that is representative of a real point cloud. This is generated by placing a part CAD model at an appropriate relative distance from a virtual camera in a simulated scene. There are mainly five sources of error that deviate the synthetic signature from the reference signature of the real part:

1) CAD model sampling error
2) Algorithm moves in discrete steps
3) The CAD model dimensions differ slightly from that of the real part
4) Gaussian sensor noise
5) Some points (mainy near part boundaries) are not visible due to sensor noise

The first two errors are taken care of by using the same CAD model sampling and the same discretization steps of the matching algorithm as used for the real part. The third source of error is accounted for by manually measuring the dimensions of the real part and using them to create a better approximation of the real part. The fourth error is addressed by adding Gaussian noise into the synthetic point cloud. The final source of error is accounted for by randomly culling a few percent of points such that points near boundaries have much higher probability of removal than interior points. The signatures for the synthetic part and a real part, each in five different postures are shown in Fig. 6. Note from the figure that the synthetic signatures closely approximate the signatures of the real part.

By using the above procedure, a set of 100 synthetic signatures were obtained and a histogram of the corresponding dissimilarities, along with dissimilarities for real part in 10 different postures, was used to approximate the probability distribution of dissimilarity between ideal and reference signatures (Fig. 7). The resulting dissimilarity distribution can be approximated as a normal distribution with a mean $\mu$ = 0.9751 and standard deviation $\sigma$= 0.0659. The standard deviation in position $\sigma_p$ = 0.51 mm and standard deviation in orientation $\sigma_o = 0.43^o$. Given an observation, which is a point cloud of the bin and the filtered CAD model of the desired part, the observed signature is obtained and its dissimilarity with ideal is computed. This observed dissimilarity is used in conjunction with the dissimilarity probability distribution for the purpose of confidence estimation. If the measured dissimilarity is not in the range $[\mu - 3\sigma, \mu + 3\sigma]$ (= [0.77, 1.17]), then it implies that the confidence in the part match is low, thereby declaring the part match as a failure.

Another parameter that influences matching performance, and thereby the confidence measure, is the percentage of points in the point cloud of the CAD model that are filtered either due to self occlusions or occlusions due to other neighboring parts. Therefore, whenever the filtered points are above a certain threshold (arbitrarily, we chose 70 %), we declare the part match as a failure.

## V. Design of User Interface to Enable Remote Human Interventions

We have developed a new user interface (Fig. 8) that allows a remote human to perform pose estimation in scenes with high clutter where the automated perception system may fail. The system makes a Skype call to a remote human when help is needed and sends information consisting of the raw camera image of the scene, the corresponding point cloud, and the CAD model of the part to be picked.

The human operator selects features (edges) from the 2D image and shows a correspondence in the CAD model (Fig. 9). The algorithm uses these features to estimate the part location and orientation in 3D and dock the CAD model at this pose. The user can do minor adjustments to the pose using a joystick. The x and y information in the image space is transformed to point cloud co-ordinates using scaling and translation operations.

### A. Evaluation of the User Interface

In these experiments, we considered complex bin examples where the failure rate of the automated perception system was more than 50 %. Figure 8 shows the user interface used by the human to perform part matching on the case where the automated perception system failed (Fig. 3). Figure 8(c) shows the CAD model docking using the edge-selection method. The user interface provides different functions that allow the human operator to achieve the part matching task. We conducted experiments to analyze the influence of different combinations of these features on the time taken to solve the problem and the overall success rate of the singulation task. Accordingly, the effectiveness of the user interface was evaluated across three experimental regimes:

1) Usage of only joystick to move the CAD model and dock it at an appropriate posture in the point cloud
2) Usage of only the edge selection method to directly dock the CAD model
3) Usage of the edge selection method to dock the CAD model, and subsequently the joystick to do any fine adjustments if necessary

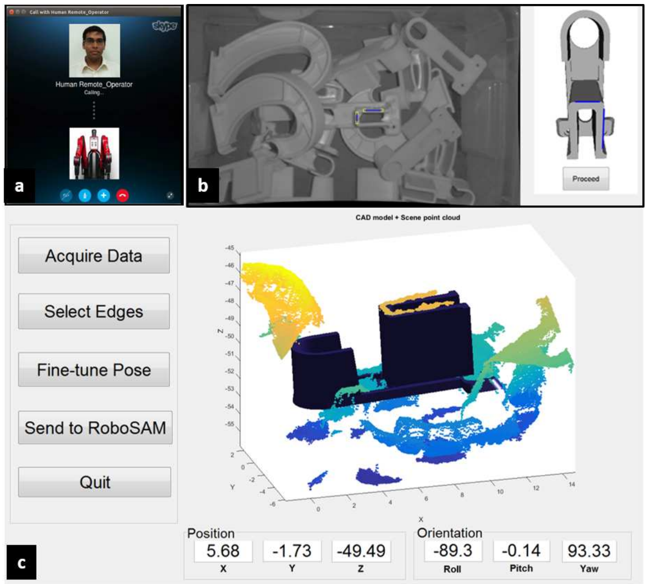


Fig. 8. (a) Robot sending Skype call to a remotely operating human requesting for help. (b) 2D image of the bin and CAD model of the part to be identified. (c) User interface used by the remote human to resolve part recognition and pose estimation failures.

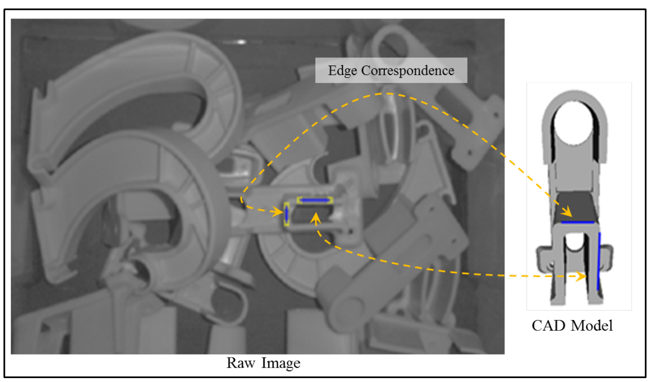


Fig. 9. Illustration of human identifying correspondences between edges in the image with those in a CAD model

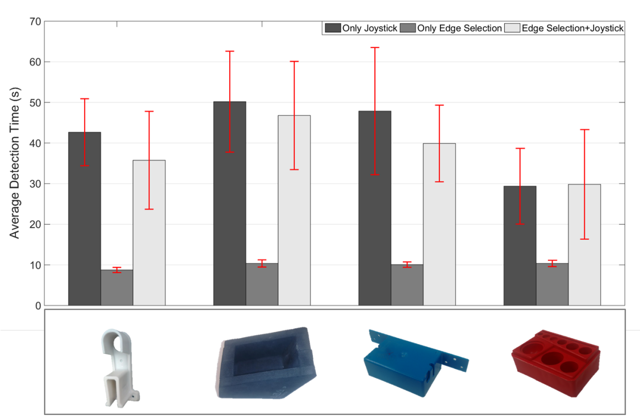


Fig. 10. Average time taken (in seconds) by the human to complete the perception task for 10 trials in each regime across four different parts. Success rates of 100 %, 80 %, and 100 % were achieved in the first, second , and third regimes, respectively, for all the four parts.

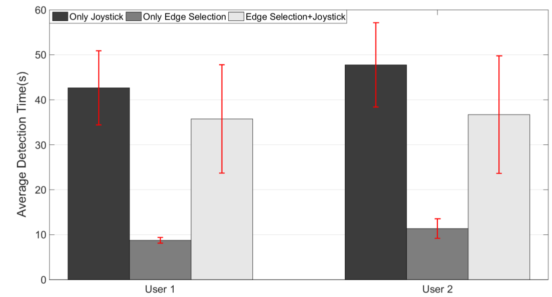


Fig. 11. Comparison of average time taken (in seconds) by two users to complete the perception task for 10 trials in each regime for the white part

We conducted the experiments for a total of 120 trials. Each trial consisted of the human using one of the three methods to perform the part matching task. The trial was validated by sending the extracted postural information to the robot and verifying whether or not the robot could singulate the specified part by using this information. We conducted ten trials for each regime and across four parts with different geometries. We expect that this task will be performed by experts in real industrial settings. Therefore, all trials were carried out by a well-trained user. The singulation success rate was 80 % in the third regime where only edge selection was used to register the part. In the first and third regimes, the success rate was 100 %. Because of high success rates, 10 user trials per regime was sufficient to validate the effectiveness of the user interface. The time taken (in seconds) by the human to complete the perception task over ten trials in each regime for all the four parts is shown in Fig. 10. Similar performance was observed across the parts for all the regimes. The edge-selection only took the least time for all the parts, but with some failure rate. Therefore, the third regime that ranked second in terms of time, and with 100 % success rate was chosen as the best solution.

In the third regime, the user spends about 10 s in edge selection and subsequently about 25 s using the joystick to improve the estimated posture. Note that about 80 % success rate can be expected with only joystick (from second regime). This information can be exploited by the user to reduce the time spent in using joystick to achieve a level of accuracy, which may be redundant.

Next, we tested the trainability of the interface. For this purpose, we trained a second user and conducted ten trials for the white part to compare the user's performance with that of the first user in all the three regimes. To have a common benchmark, the same data used by the first user was presented to the second user. The comparison was only limited to the part matching task in these experiments as the same bin settings were no more available to proceed with the singulation task. Instead, difference in transformations was computed and used as a comparison metric. Figure 11 shows a comparison of time taken by the two users to complete the perception task. The second user took an average of 36.7 sec to complete the perception task for the white part, in the third regime, which is very close to that of the first user. Similar performance was observed in first and second regimes.

## VI. Conclusions

We presented an approach that treats coping with uncertainty as a key step to handling failures and enhancing performance in robotic unstructured bin-picking. We used different experimental regimes to evaluate the effectiveness of the user interface. Currently, we have tested the effectiveness of the user interface for four parts. Empirical evaluations across a variety of part shapes are needed to perform further testing of the methods presented in the paper. Future work also includes analyzing the frequency of calls to human and comparing the cycle time of the human-in-the-loop concept with the classical way of bin-picking when exception occurs. Integration of the methods presented in the paper with a singulation planning method [23] and a fine-positioning method [24] is currently under progress. In our previous work, we have developed other modules including ontology for task partitioning in human-robot collaboration for kitting operations [25], sequence planning for complex assemblies [26], instruction generation for human operations [27], and ensuring human safety [28]. Future work consists of investigating how to integrate them to realize hybrid work cells where humans and robots collaborate to carry out industrial tasks.

## References


[1] M.-Y. Liu, O. Tuzel, A. Veeraraghavan, Y. Taguchi, T. K. Marks, and R. Chellappa, "Fast object localization and pose estimation in heavy clutter for robotic bin picking," *The International Journal of Robotics Research*, vol. 31, no. 8, pp. 951–973, 2012.

[2] J. A. Marvel, K. Saidi, R. Eastman, T. Hong, G. Cheok, and E. Messina, "Technology readiness levels for randomized bin picking," in *Proceedings of the Workshop on Performance Metrics for Intelligent Systems*, PerMIS '12, (New York, NY, USA), pp. 109–113, ACM, 2012.

[3] K. N. Kaipa, S. S. Thevendria-Karthic, S. Shriyam, A. M. Kabir, J. D. Langsfeld, and S. K. Gupta, "Resolving automated perception system failures in bin-picking tasks using assistance from remote human operators.," in *Proc. of IEEE International Conference on Automation Science and Engineering*, August 2015.

[4] D. Buchholz, S. Winkelbach, and F. M. Wahl, "Ransam for industrial bin-picking," in *Robotics (ISR), 2010 41st International Symposium on and 2010 6th German Conference on Robotics (ROBOTIK)*, pp. 1–6, June 2010.

[5] S. Balakirsky, Z. Kootbally, C. Schlenoff, T. Kramer, and S. K. Gupta, "An industrial robotic knowledge representation for kit building applications," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1365–1370, Oct 2012.

[6] A. Schyja, A. Hypki, and B. Kuhlenkotter, "A modular and extensible framework for real and virtual bin-picking environments," in *Proc. of IEEE International Conference on Robotics and Automation*, pp. 5246–5251, May 2012.

[7] B. K. Horn and K. Ikeuchi, "Picking parts out of a bin," tech. rep., Massachussetts Institute of Technology, 1982.

[8] S. Kristensen, S. Estable, M. Kossow, and R. Brsel, "Bin-picking with a solid state range camera," *Robotics and Autonomous Systems*, vol. 35, no. 34, pp. 143 – 151, 2001. Seventh Symposium on Intelligent Robotic Systems - SIRS'99.

[9] F. Boughorbel, Y. Zhang, S. Kang, U. Chidambaram, B. Abidi, A. Koschan, and M. Abidi, "Laser ranging and video imaging for bin picking," *Assembly Automation*, vol. 23, no. 1, pp. 53–59, 2003.

[10] S. Leonard, A. Chan, E. Croft, and J. J. Little, "Robust motion generation for vision-guided robot bin-picking," in *ASME 2007 International Mechanical Engineering Congress and Exposition*, pp. 651–658, American Society of Mechanical Engineers, 2007.

[11] C. Papazov, S. Haddadin, S. Parusel, K. Krieger, and D. Burschka, "Rigid 3d geometry matching for grasping of known objects in cluttered scenes," *The International Journal of Robotics Research*, 2012.

[12] V.-D. Nguyen, "Constructing force-closure grasps," *The International Journal of Robotics Research*, vol. 7, no. 3, pp. 3–16, 1988.

[13] M. Roa and R. Suarez, "Computation of independent contact regions for grasping 3-d objects," *IEEE Transactions on Robotics*, vol. 25, pp. 839–850, Aug 2009.

[14] Y. Zheng and W.-H. Qian, "Coping with the grasping uncertainties in force-closure analysis," *The International Journal of Robotics Research*, vol. 24, no. 4, pp. 311–327, 2005.

[15] D. Berenson, S. Srinivasa, and J. Kuffner, "Addressing pose uncertainty in manipulation planning using task space regions," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1419–1425, Oct 2009.

[16] J. Weisz and P. Allen, "Pose error robust grasping from contact wrench space metrics," in *Proc. of IEEE International Conference on Robotics and Automation*, pp. 557–562, May 2012.

[17] B. Kehoe, D. Berenson, and K. Goldberg, "Toward cloud-based grasping with uncertainty in shape: Estimating lower bounds on achieving force closure with zero-slip push grasps," in *Proc. of IEEE International Conference on Robotics and Automation*, pp. 576–583, May 2012.

[18] S. Fuchs, S. Haddadin, M. Keller, S. Parusel, A. Kolb, and M. Suppa, "Cooperative bin-picking with time-of-flight camera and impedance controlled dlr lightweight robot iii," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*, pp. 4862–4867, Oct 2010.

[19] A. Pronobis and B. Caputo, "Confidence-based cue integration for visual place recognition," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*, pp. 2394–2401, Oct 2007.

[20] C. Weber, S. Hahmann, and H. Hagen, "Sharp feature detection in point clouds," in *Shape Modeling International Conference (SMI), 2010*, pp. 175–186, June 2010.

[21] K. Demarsin, D. Vanderstraeten, T. Volodine, and D. Roose, "Detection of closed sharp edges in point clouds using normal estimation and graph theory," *Computer-Aided Design*, vol. 39, no. 4, pp. 276 – 283, 2007.

[22] S. Gumhold, X. Wang, and R. Macleod, "Feature extraction from point clouds," in *In Proceedings of the 10 th International Meshing Roundtable*, pp. 293–305, 2001.

[23] K. N. Kaipa, S. Shriyam, N. B. Kumbla, and S. K. Gupta, "Automated plan generation for robotic singulation from mixed bins," in *IROS Workshop on Task Planning for Intelligent Robots in Service and Manufacturing*, 2015.

[24] K. N. Kaipa, N. B. Kumbla, and S. K. Gupta, "Characterizing performance of sensorless fine positioning moves in the presence of grasping position uncertainty," in *IROS Workshop on Task Planning for Intelligent Robots in Service and Manufacturing*, 2015.

[25] A. G. Banerjee, A. Barnes, K. N. Kaipa, J. Liu, S. Shriyam, N. Shah, and S. K. Gupta, "An ontology to enable optimized task partitioning in human-robot collaboration for warehouse kitting operations," in *Proc. SPIE, Next-Generation Robotics II; and Machine Intelligence and Bio-inspired Computation: Theory and Applications IX, 94940H*, 2015.

[26] C. Morato, K. N. Kaipa, and S. K. Gupta, "Improving assembly precedence constraint generation by utilizing motion planning and part interaction clusters," *Computer-Aided Design*, vol. 45, no. 11, pp. 1349 – 1364, 2013.

[27] K. N. Kaipa, C. Morato, B. Zhao, and S. K. Gupta, "Instruction generation for assembly operations performed by humans," in *ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 1121–1130, American Society of Mechanical Engineers, 2012.

[28] C. Morato, K. N. Kaipa, B. Zhao, and S. K. Gupta, "Toward safe human robot collaboration by using multiple kinects based real-time human tracking," *Journal of Computing and Information Science in Engineering*, vol. 14, no. 1, p. 011006, 2014.