Reflective workpiece detection and localization for flexible robotic cells

# Reflective workpiece detection and localization for flexible robotic cells

Sergey Astanin, Dario Antonelli*, Paolo Chiabert, Chiara Alletto

[a] *Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, Italy*

**Abstract**

A smart vision system for industrial robotic cells is presented. It can recognize and localize a reflective workpiece, and allows for automatic adjustments of the robot program. The purpose of the study is to improve industrial robots awareness of the environment and to increase adaptability of the manufacturing processes where full control over environment is not achievable. This approach is particularly relevant to small batch robotic production, often suffering from only partial control of the process parameters, such as the order of jobs, workpiece position, or illumination conditions.

A distinguishing aspect of the study is detection of workpieces made of diverse materials, including shiny metals. Reflective surfaces are common in the industrial manufacturing, but are rarely considered in the research on object recognition because they hinder many of the object recognition algorithms. The proposed solution has been qualified and tested on a selected benchmark in realistic workshop environment with standard artificial light conditions. The training of the object recognition software is an automatic process and can be executed by non-expert industrial users to allow for recognition of different types of objects.

*Keywords:* Machine Vision, Human-Robot Cooperation

## 1. Introduction

Flexibility in manufacturing is a multifaceted concept. Its definition varies according to the context. In [1] twelve different definitions are reported. [2] proposes several quantitative definitions of the flexibility. This work is mostly concerned with product flexibility, which is the ability of a manufacturing system to make a variety of part types with the same equipment [3]. It is referred to as reconfigurability of a manufacturing system on cell and work-piece levels, or its ability to switch with minimal effort to a particular family of workpieces.

This kind of flexibility is particularly relevant in collaborative human-robot scenarios for small and medium enterprises, where some operations, usually handling, are executed by the human operator, while the others, usually assembling or processing, are performed by the robot [4]. Introduction of the human factor brings the advantage of more dexterity in the execution of tasks but introduces a source of variability within the workplace. The cell may not be considered a structured environment anymore, so this scenario hinders offline robot programming [5]. Therefore an adaptive robotic system should be aware of the environment, analyze it, and change its behavior dynamically [6, 7].

In a hypothetical scenario, where workpieces are handled and positioned by a human worker, and robot is supposed to do some other operations on them, the variable aspects are workpiece readiness, type, state and location.

## 2. State of the Art

In this section some methods to solve the above mentioned problems of object detection, recognition and localization are presented with a focus on industrial applications.

*Detection of unknown objects* is a frequent task in video processing and surveillance. Background subtraction is a very popular approach to detect unknown objects [8]. More elaborate methods may rely on image segmentation [9, 10] or saliency detectors [11].

Background subtraction allows to highlight the areas of the image which are significantly different from the previous (background) images. In the industrial settings with a fixed camera, this allows to see what areas of the work space are different with respect to some reference state, or, in other words, to see where new objects have appeared or the work area was modified.

Object detection and the wider problem of *object recognition and localization* do not necessarily imply that vision-based techniques should be used. Autonomous vehicles often use laser scanners to detect unknown objects [12, 13]. Some industrial applications opt for non-visual

*Corresponding author.
*Email address:* `dario.antonelli@polito.it` (Dario Antonelli)

methods to recognize and localize objects too. RFID is a common means to carry object type and identity [14]. Three-dimensional range imaging and scanners like Kinect are becoming increasingly popular too [15, 16, 17, 18]. For a full review of 3D data acquisition techniques see [19]. But machine vision still remains one of the most popular methods for parts identification [20].

Reflective objects are rarely considered in visual recognition research. Sometimes the object is only slightly reflective and the traditional methods are still applicable [21], but this is rarely the case for textureless objects. Many authors employ workarounds like relying on contours [22] or using 3D scanners[23, 24]. However, 3D scanners often fail to obtain reliable 3D depth images from non-Lambertian surfaces, and few point-to-point correspondences can be found in a stereo pairs [24]. Impressive results based on matching the contours were achieved in [25, 26]. Unfortunately, the method relies on a custom multi-flash camera [27]. To the extent of our knowledge at the time of writing, no commercial implementation of such hardware exists, and the image processing code is not widely available, neither in Open Source nor commercially.

Many industrial applications rely on relatively simple marker-based [28, 29, 30] and convolution-based methods [22, 31]. Recognition and localization problems can be completely bypassed when fiducial markers used. Their practical applications are limited by the necessity to attach and detach markers, operations which complicate the production process. Convolution-based method is a case of an example-driven search. They are especially sensitive to variability of the object, and usually are not scale- and rotation-invariant [32, 33]. Thus both approaches are poorly suited for localization of manually positioned reflective objects in industrial environment.

More elaborate methods rely on feature-based matching. The features may be derived from CAD models [34] or chosen manually in an ad-hoc manner [35]. Unfortunately, the choice of good invariant features is not straight-forward and requires human expertise.

General object recognition methods are very helpful if the system is supposed to deal with highly variable object appearances. Many such methods are able to learn or select good features automatically, given a good training set [36, 37, 38]. These methods have been successfully used for face or pedestrian detection and general object recognition.

In this paper a system fordetection, recognition and localization of a workpiece made of reflective metal is proposed. It does not rely on fiducial markers and can cope with high variability of the illumination conditions and object appearance.

The proposed solution is build upon a simple background subtraction to detect unknown objects, and a Viola-Jones classifier to recognize the workpiece. 3D scanners were avoided because their performance is known to degrade when working with the reflective surfaces. Automatic feature selection based on actual workpiece appearance was preferred over the manual feature selection or features design from the object design. The output of the system is used to adjust and run the robot program, and the performance of the method is validated experimentally.

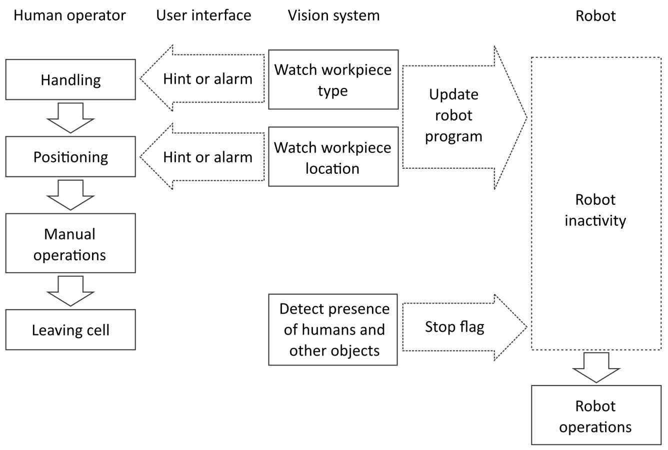



Figure 1: A potential role of a vision system in an adaptive robotic cell, where some operations like handling and positioning are carried out by a human operator. The vision system may communicate to the human (by suggesting what to do or warning about irregularities) and with the robot (by adjusting the robot program and vetoing execution until the environment is ready).

## 3. Case Study

To demonstrate our implementation of the adaptive robotic cell, a case study was planned which consisted of 1) manually positioning of an aluminum CNC machined workpiece on the bench, 2) detection, recognition and localization of the workpiece using a vision system, 3) adjustment of the robot program according to the type and the location of the workpiece, or sending a signal to delay or prevent program execution, if the workpiece is not present or it is a wrong one (Fig. 1). This section describes various aspects of the experiment, such as workpiece design and material selection,

**Workpiece**. The workpiece used in the experiment has geometrical features typical of most mechanical components, such as, cylindrical holes and pins, prismatic ribs.

**Material**. The choice of the workpiece material is the fundamental decision in the design of the experiment. Image interpretation and object recognition depend on many properties, such as texture, color, visible edges which may vary according to the material [39]

Metals and polymers are both frequently present in manufacturing, but metals represent a challenge for many vision methods. On a reflective material we may see false specular edges or edges highlighted according to particular orientation between light source, object surface and observer [39]. Specular reflections are often overexposed, thus texture and color information may be easily lost. Surface roughness is another factor to consider. Rough surfaces may appear to have more diffuse reflection. Smooth surfaces tend to produce more specular reflections. Polished reflective material appears to be the worst case in terms of reflections. The proposed method is tuned specifically to deal with objects made of reflective metal.

**Illumination conditions**. The trials were conducted in a laboratory, with artificial fluorescent lights typical for a workshop. The illumination is stationary and constant but is not uniform nor diffuse.

**Work area**. The experiment is carried out in a robot cell which includes a robot, a work bench, a camera, a computer to run image processing software, a local area network to communicate with the robot, and a physical barrier that enclosed all elements. The work bench has a white shiny surface, completely reachable by the robot arm.

**Equipment**. To carry out the trials, a USB camera, Logitech C920, operated in VGA resolution, and an anthropomorphic robot with six degrees of freedom, Comau Smart5 NS, were used. The camera was located over the work bench so that the entire workarea was visible.

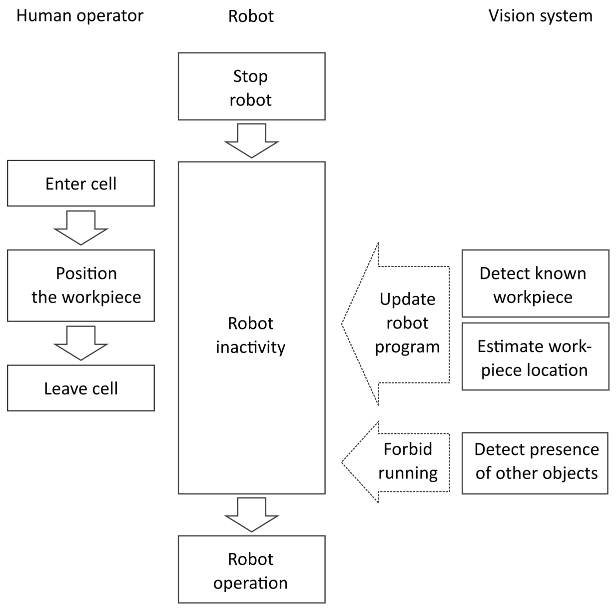



Figure 2: Usage scenario in our case study.

**Usage scenario**. The full cycle of the system is supposed to start with a stopped robot, so that the operator may enter the work cell. Once he or she positions the workpiece on the bench, the vision system should detect the workpiece, calculate its location, and update the robot program. The operator leaves the cell to start the robot from outside. The vision system may prevent it from running if it detects unknown objects on the work bench.

The following outcomes were considered:

- No object is detected at all; the robot program cannot be generated.
- One and only one item of the correct type is placed on the bench and nothing is left in the visible work area; a new robot program may be executed when the operator leaves the cell.
- An item of the correct type and one or more additional objects of any type are detected in the visible work area; the program can be generated, but its execution is delayed until the additional objects are removed.

## 4. Algorithms

The problem was decomposed into three tasks: detection and recognition of the known workpiece, estimation of the object location in the real-world coordinate frame, and detection of unknown objects via background subtraction. The output of these tasks drives the final decision making.

### *4.1. Workpiece detection and recognition*

A workpiece made of reflective metal may look very differently depending on its location with respect the camera and the light source. In Fig. 3 ten randomly selected samples of the same item are presented. The images were cropped from a frame taken with a stationary camera and under constant artificial illumination. Only location of the workpiece on the bench was varied. It can be also noted, that when the camera is observing a wider area, the amount of detail and resolution of the smaller object of interest could be severely reduced.

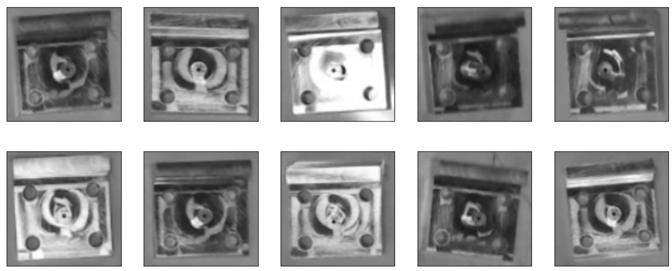


Figure 3: Variability of the workpiece appearance with respect to small ($< 50$ cm) displacements observed from $h \approx 1$ m.

Manual selection of the distinctive features or their combinations is not trivial. For this reason the Viola–Jones method [37, 40] was used, which can identify good features automatically. The method relies on four ideas:

- Build an integral image which allows to calculate integrals of image intensity over an arbitrary rectangular area in $O(1)$ time. Given the original image $i(x, y)$, the integral image can be defined as

$$I(x, y) = \sum_{\substack{0 \le x' \le x \\ 0 \le y' \le y}} i(x', y')$$

- Consider only the so called Haar-like features: the differences in the integral intensity between two or three adjacent rectangular areas of the image. Only four references to the integral image are sufficient to calculate an integral over an arbitrary rectangular area of the original image.

- Use Haar-like features as "weak classifiers" (choose a threshold $\theta$ and polarity for every feature $f$). Such classifiers usually have very high error rates (up to 0.4) individually, but they can be combined using a boosting method (like AdaBoost or Gentle AdaBoost [41]) to construct stronger classifiers.

- Build a degenerate decision tree (a "cascade") which rejects negative candidates as early as possible. At every node of the cascade, the threshold of the strong classifier is adjusted to *minimize false negatives*. If the first node yields a negative result, the processing is terminated. If the result is positive, the classifier of the second node is used and so on. The cascade tries to reject as many negatives as possible as early as possible. The subsequent classifiers are trained on the data samples which pass through the previous stages

This method allows for fast real-time recognition of objects. Notably, the method, by its nature relying on integral macro features, is relatively forgiving to low resolution and out of focus images. According to [37] the accuracy of the cascading classifier can be tuned by choosing a maximum acceptable rate of false positives $f_{FP}$ at every stage (it decides how many false positives remain after every stage), a maximum acceptable miss rate $f_{FN}$ at every stage (it decides how many positive samples are irrevocably missed at every stage), and the number of the stages $n$. The false positive rate of the cascading classifier can be estimated as

$$F_{FP} = \prod_{i=1}^{n} f_{FP} = f_{FP}^{n}.$$

And the recall (the ratio of true positive outcomes to the number of positive samples) of the classifier can be estimated as

$$F_{FN} = \prod_{i=1}^{n} (1 - f_{FN}) = (1 - f_{FN})^{n}.$$

Assuming that a single stage classifier has a false positive rate of only 0.4, and miss rate of 0.005, the cascading classifier with 10 stages will have a false positive rate of $0.4^{1}0 \approx 1/1000$, and will manage to detect $(1 - 0.005)^{1}0 \approx 95\%$ of the positive samples.

In applications, the cost of type I errors (false positives, the workpiece was reported when it was not there) and type II errors (false negative, the workpiece was not detected) errors may be different. The objectives on both depend on the chosen risk management policy. The desired overall false positive and detection rates and can be achieved by imposing the number of classifier stages, varying the number of features used on each stage and choosing thresholds of each stage. Also the importance of a sufficiently large and diverse training set cannot be underestimated.

330 positive samples and 183 background samples were randomly selected for a training set. To produce background samples, the object was artificially removed from the image by inpainting [42] and a random area of the image was selected. To mitigate effects of manual labeling of the objects' region of interest, the training set was further increased tenfold by doing ten random crop and random rotations per each sample (Fig. 4). So the effective size of the training set was over 5000 samples produced from 330 distinct video frames. A test set of the same size was produced from the remaining labeled frames.

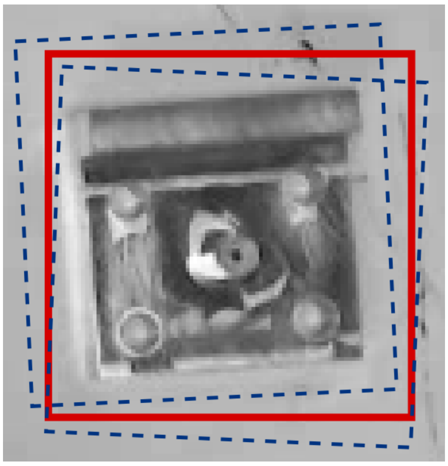


Figure 4: Manually labeled region of interest (solid line) and two possible random crops (dashed line).

The recognition method developed for our case study is a binary classifier. The common approach to evaluate its performance is to apply the classifier to a labeled test set and calculate the confusion matrix, i.e. the number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) results. This allows to calculate various performance metrics [43]:

- Precision $= TP/(TP + FP)$,
- Recall $= TP/(TP + FN)$,
- F-score $F_1 = 2TP/(2TP + FP + FN)$.

Recall, also known as sensitivity, reflects the effectiveness of the classifier to identify positive values (to never miss the object). Precision reflects a degree of agreement of the data with the positive labels with the labels given by the classifier (its ability to avoid producing false positives). $F_1$, also known as F-score, is the harmonic mean of the precision and the recall.

In OpenCV implementation of the multi-scale Viola-Jones detector, the user may choose a sensitivity parameter `minNeighbors`. Different values of the parameter represent different trade-offs between precision and recall of the classifier (Fig. 5). Even with only fixed-scale detector, it was possible to achieve 98% recall with optimal precision. Such an object detector was adequate for our case study.

### *4.2. Location estimation*

The position of the detected workpiece is defined in the reference frame of the image, in pixels. To trigger robot program the location of the workpiece in the three-dimensional real world reference frame has to be calculated.

A common camera calibration procedure was used with a 9×6 chessboard-like pattern observed in different locations of the frame [44, 45]. This procedure is widely known in the field of computer vision, and allows to estimate camera projection matrix, distortion coefficients, and rotation-translation matrices for each of the chessboard positions. One of the chessboard positions was related to the real-world (robot's) reference frame. Camera calibration allows to estimate parameters of the nonlinear transform from the real-world coordinates $X$, $Y$, $Z$ to the image coordinates $u$, $v$. The inverse transform is generally not unique, unless an additional constraint is imposed. The region of interest of a two-dimensional image provides only two degrees of freedom, $(u, v)$. Assumption, that the workpiece is positioned on the surface of the table $Z = 0$ is the additional constraint. The inverse transform was calculated by an iterative solver. Thus for any point of the image three-dimensional real-world coordinates of the corresponding point on the bench can be calculated and passed to the robot.

Coordinates $(X, Y, Z)$ of the approximate center of the object could be sufficient to execute many manufacturing operations. Whenever the complete location information is required, an additional step should be taken. A higher resolution image of the region of interest was used in this case to find (predict) positions of 5 or more reference points on the workpiece. The predictor could be as simple as a random forest regressor [46, 47] trained on the same training set (if the reference points were properly marked). Though our preliminary results show that neural networks could have better generalization properties. These points and camera calibration data could be used to estimate all six degrees of freedom of the workpiece location [48] (Fig. 6).

## 5. Background subtraction

Object detection method and classifier may recognize instances of the known object. To detect *unknown* objects, the foreground objects have to be separated from the background image.

Many of the publications on background subtraction are concerned with dynamically updating background models. Mixture of Gaussians [49] appears to be one of the mostly used background models [50, 51]. In our case, on a work bench under artificial illumination, the background is always the same, and a simple Gaussian model should suffice.

The average background is subtracted from every frame, and a binary threshold is applied. The threshold parameter $t$ is global and constant. It was set just above the level where the method detects reflections on the bench surface (Fig. 7).

### *5.1. Decision-making*

The output of the Viola–Jones classifier is a set of bounding boxes (rectangles) of likely locations of the workpiece. Usually just one such location is expected. In our case there is no indication if they are good matches.

The output of the background subtraction is a binary image mask where all pixels different from the normal background are assigned a positive intensity value.

This information is merged together to calculate:

- $A_i$, the area of the foreground objects within every detected bounding box $i$,
- $A_o$, the area of the foreground objects outside of the detected bounding boxes.

All area values are normalized to keep the implementation resolution-independent. The value of $A_o$ indicates how big or how many unknown objects are present in the frame. $A_i$ could be optionally used to discard some of the false positives. This technique may be useful to increase sensitivity of the classifier at the cost of loss of precision.

A positive decision is triggered if and only if:

- There is just one detected object
- The area $A_o$ of unknown objects is below some threshold level $t_a$
- (Optional condition) The area $A_1$ is bigger than a half of the bounding box area

The complete workflow is depicted in Fig. 8. In Fig. 9 different outcomes of the decision process are shown. Only in Fig. 9 (c) the decision is positive.

## 6. Robotic cell implementation

The camera was mounted on a statically placed photographic tripod directly over the working bench. The bench was completely within the operating range of the robot and remained easily accessible to the human operator from one of the sides. The video stream was continuously analyzed by a computer program, and a new workpiece location was uploaded to the robot when the positive decision triggered. This event was displayed to the operator (Fig. 9 c). As the prototype did not implement security measures, the operator had to leave the cell and manually start robot program from the outside. A token operation was executed: the

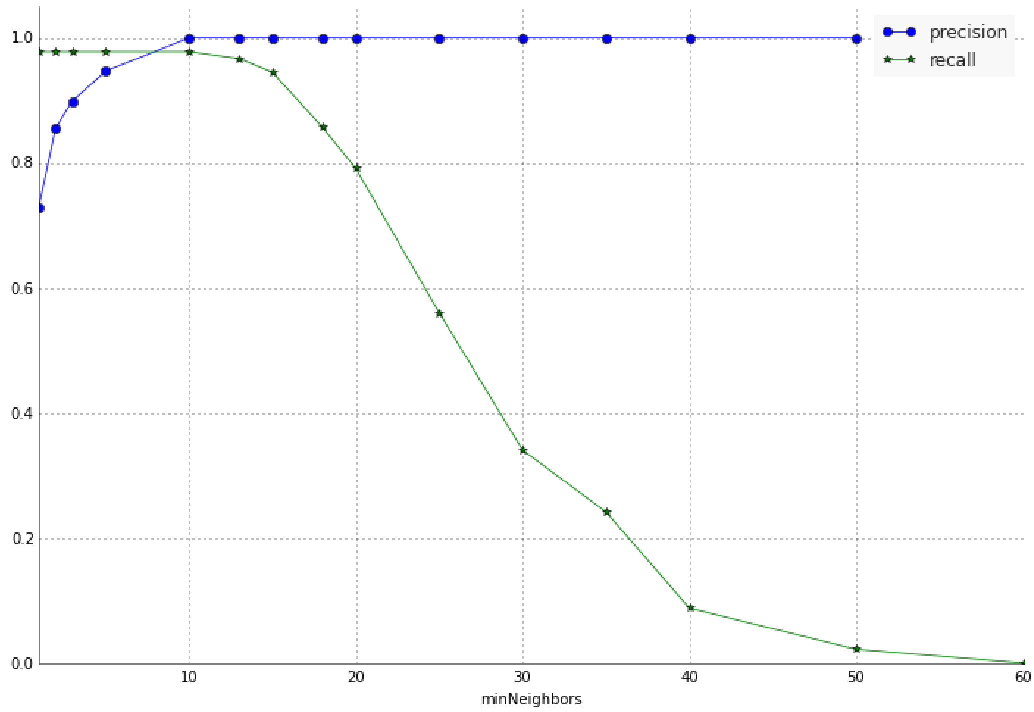



Figure 5: Precision and recall of the trained classifier.

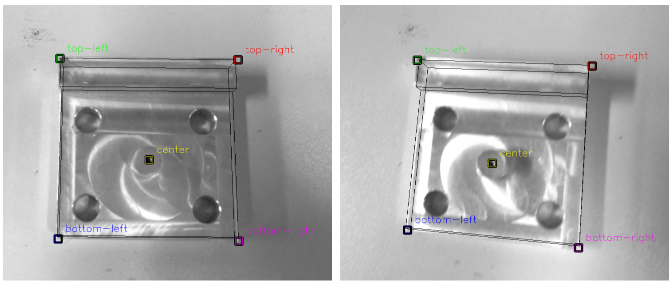


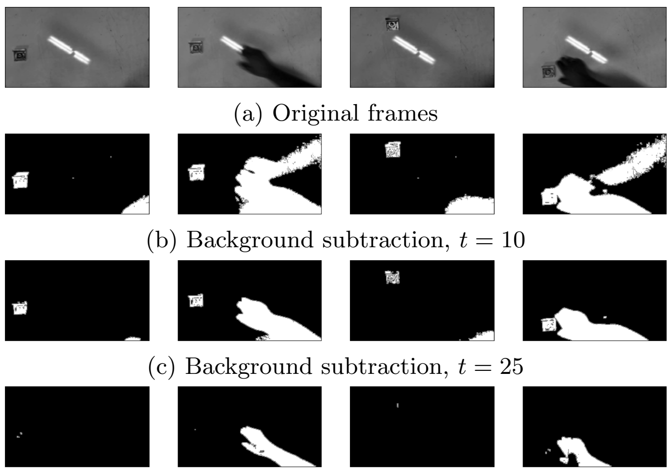

Figure 6: Reference points of the workpiece were predicted by a random forest regressor, and the pose was estimated using EPnP algorithm [48]. It was used to project a wireframe model over the image.

robot was supposed to approach and touch the workpiece. The correctness of the execution was visually controlled by analyzing the record or by examining a paint mark left by the robot (Fig. 10).

## 7. Conclusions

The proposed augmentation of the robot cell with a vision system allowed to recognize known workpiece and estimate its location, as well as to detect presence of unknown objects. This information can be used to adjust robot program automatically in the working environment where the exact position of the workpiece is unknown. The feature may be particularly useful in collaborative human-robot work cells, where handling and positioning are responsibility of the human.

We have shown that a generic object recognition method, initially developed for face detection, can be easily trained to recognize objects made of reflective metal even in unfavorable light conditions. The training of the object detector is an automatic process. Even a non-expert end user may teach the system to recognize new workpiece types if he or she supplies some labeled images of the new



(a) Original frames

(b) Background subtraction, $t = 10$

(c) Background subtraction, $t = 25$

(d) Background subtraction, $t = 80$

Figure 7: Foreground masks of a sequence of frames for various values of the threshold parameter $t$. In (b) $t = 10$ and reflections of the hand and some spurious bright spots are detected on the table surface. In (c) $t = 25$ and only the workpiece, the hand and its shadow are detected. In (d) $t = 80$ and the shadows are not detected anymore, but the masks of the hand and the workpiece are incomplete.

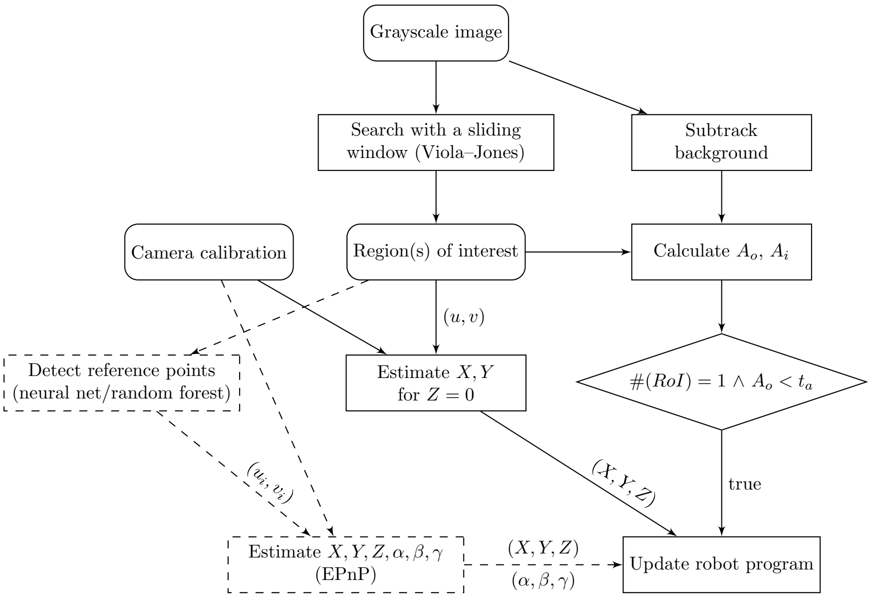
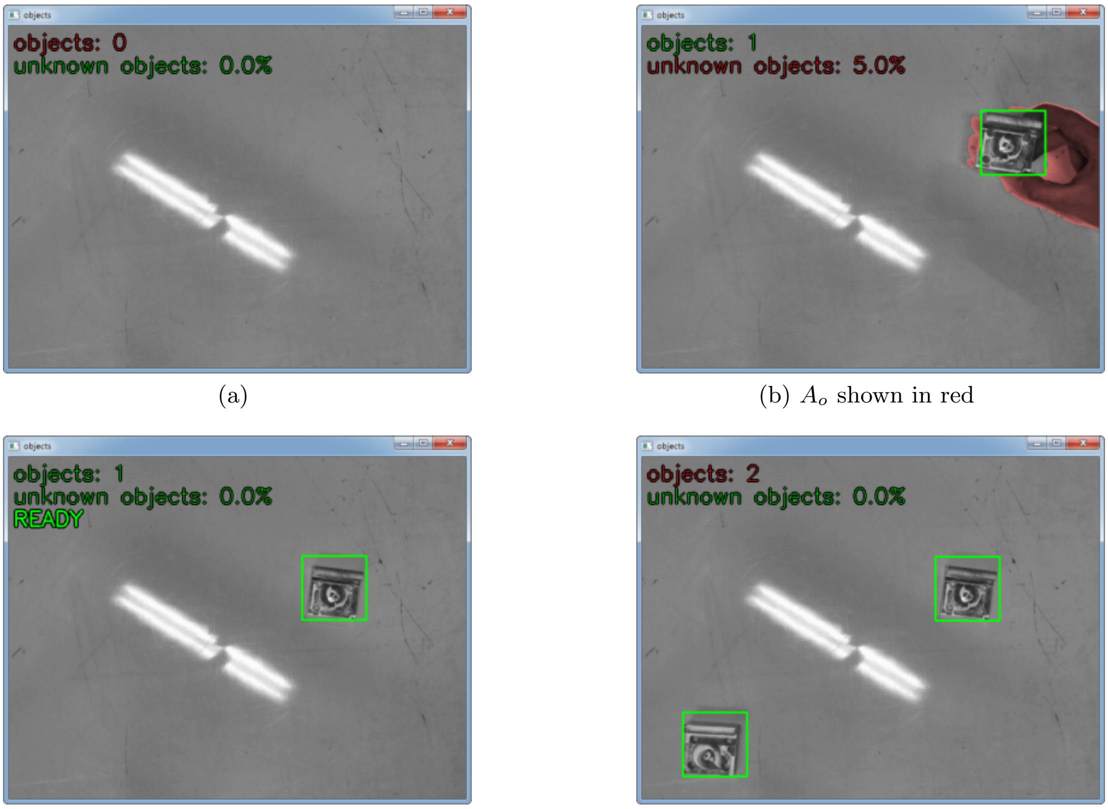



Figure 8: Complete workflow for workpiece detection, recognition and localization in a robotic cell where the location of the workpiece is subject to variability.




(a) (b) $A_o$ shown in red

(c) (d)

Figure 9: Screenshots of the workpiece detector in action. (a) Neither the workpiece nor other objects are visible (negative decision). (b) The workpiece is detected, but there is an unknown object (the hand). The decision is negative. (c) Only one workpiece is detected and no other object is detected (positive decision). (d) Two copies of the workpiece are detected (negative decision).

workpiece.

An advantage of the developed system is its easy integration with the existing robot cells: it doesnt require to replace the robotic equipment already installed in the work plant. Only a basic video camera is required, and all the algorithms have multiple Open Source and commer-

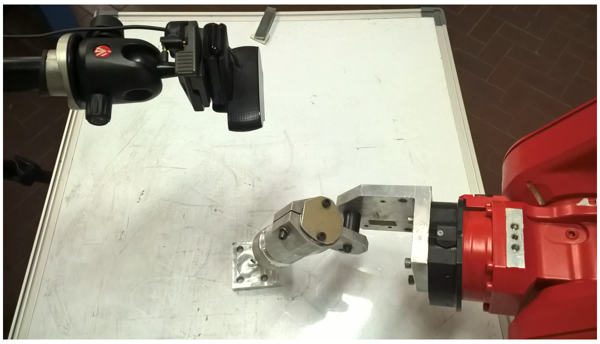


Figure 10: Experimental setup and execution of the token operation on the workpiece (touching a point with a marker).

cial implementations. Localization and pose estimation do not require a complete CAD model (only few points are used). Detection, recognition, and partial localization with only three degrees of freedom could be done without CAD. This property could be particularly relevant if collaborative robotics is ever pushed to small-scale or even artisan manufacturing. In terms of hardware and software availability and ease of setup this method compares favorably to other methods that can deal with reflective workpieces (either marker-based solutions or edge-matching systems, which rely on custom cameras, and need CAD models even for recognition).

The method can be scaled to few (up to $N \sim 5$) workpiece types by simply training multiple Viola–Jones detector and running them in parallel. Our preliminary testing shows that when the number of workpiece types is large ($N > 10$), a multi-class classifier, like a deep neural network, is a more effective approach. However, the questions of multi-class recognition and the issues of optimal training set construction have to be addressed in a separate study.

## References


[1] R. S. Wadhwa, Flexibility in manufacturing automation: A living lab case study of Norwegian metalcasting SMEs, Journal of Manufacturing Systems 31 (2012) 444–454.

[2] G. Chryssolouris, Manufacturing Systems: Theory and Practice, Mechanical Engineering Series, Springer-Verlag, New York, 2006.

[3] H.-P. Wiendahl, H. A. ElMaraghy, P. Nyhuis, M. F. Zh, H.-H. Wiendahl, N. Duffie, M. Brieke, Changeable manufacturing-classification, design and operation, CIRP Annals-Manufacturing Technology 56 (2007) 783–809.

[4] K. N. Kaipa, C. Morato, J. Liu, S. K. Gupta, Human-robot collaboration for bin-picking tasks to support low-volume assemblies, in: Human-Robot Collaboration for Industrial Manufacturing Workshop, held at Robotics: Science and Systems Conference (RSS 2014).

[5] Z. Pan, J. Polden, N. Larkin, S. Van Duin, J. Norrish, Recent progress on programming methods for industrial robots, Robotics and Computer-Integrated Manufacturing 28 (2012) 87–94.

[6] C. Raibulet, Facets of adaptivity, in: Software Architecture, Springer, 2008, pp. 342–345.

[7] A. M. Petrina, Advances in robotics (Review), Automatic Documentation and Mathematical Linguistics 45 (2011) 43–57.

[8] A. Bayona, J. SanMiguel, J. Martinez, Comparative Evaluation of Stationary Foreground Object Detection Algorithms Based on Background Subtraction Techniques, in: Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, 2009. AVSS '09, pp. 25–30.

[9] S. Gould, T. Gao, D. Koller, Region-based segmentation and object detection, in: Advances in neural information processing systems, pp. 655–663.

[10] L. Li, W. Huang, I. Y. Gu, Q. Tian, Foreground object detection from videos containing complex background, in: Proceedings of the eleventh ACM international conference on Multimedia, ACM, 2003, pp. 2–10.

[11] A. Borji, D. N. Sihite, L. Itti, Salient object detection: A benchmark, in: Computer Vision–ECCV 2012, Springer, 2012, pp. 414–429.

[12] F. Fayad, V. Cherfaoui, Tracking objects using a laser scanner in driving situation based on modeling target shape, in: Intelligent Vehicles Symposium, 2007 IEEE, IEEE, 2007, pp. 44–49.

[13] K. C. Fuerstenberg, K. C. Dietmayer, V. Willhoeft, Pedestrian recognition in urban traffic using a vehicle based multilayer laserscanner, in: Intelligent Vehicle Symposium, 2002. IEEE, volume 1, IEEE, 2002, pp. 31–35.

[14] S. Makris, G. Michalos, G. Chryssolouris, RFID driven robotic assembly for random mix manufacturing, Robotics and Computer-Integrated Manufacturing 28 (2012) 359–365.

[15] A. Ferraro, M. Indri, I. Lazzero, Dynamic update of a virtual cell for programming and safe monitoring of an industrial robot (2012).

[16] P. M. da Costa, P. Costa, P. Costa, J. Lima, G. Veiga, Part Alignment Identification and Adaptive Pick-and-place Operation for Flat Surfaces, in: Robotics in Smart Manufacturing, Springer, 2013, pp. 203–212.

[17] K. Elashry, R. Glynn, An Approach to Automated Construction Using Adaptive Programing, in: Robotic Fabrication in Architecture, Art and Design 2014, Springer, 2014, pp. 51–66.

[18] L. F. Rocha, M. Ferreira, V. Santos, A. P. Moreira, Object recognition and pose estimation for industrial applications: A cascade system, Robotics and Computer-Integrated Manufacturing 30 (2014) 605–621.

[19] Z. M. Bi, L. Wang, Advances in 3d data acquisition and processing for industrial applications, Robotics and Computer-Integrated Manufacturing 26 (2010) 403–413.

[20] H. Golnabi, A. Asadpour, Design and application of industrial machine vision systems, Robotics and Computer-Integrated Manufacturing 23 (2007) 630–637.

[21] P. Piccinini, A. Prati, R. Cucchiara, Real-time object detection and localization with SIFT-based clustering, Image and Vision Computing 30 (2012) 573–587.

[22] P. Bellandi, F. Docchio, G. Sansoni, Roboscan: a combined 2d and 3d vision system for improved speed and flexibility in pick-and-place operation, The International Journal of Advanced Manufacturing Technology 69 (2013) 1873–1886.

[23] A. Pochyly, T. Kubela, M. Kozak, P. Cihak, Robotic vision for bin-picking applications of various objects, in: Robotics (ISR), 2010 41st international symposium on and 2010 6th German conference on robotics (ROBOTIK), VDE, 2010, pp. 1–5.

[24] J. Lee, T.-W. Kim, S. Kang, K. Kim, J. Kim, J. B. Kim, Bin picking for the objects of non-Lambertian reflectance without using an explicit object model, in: 2014 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), pp. 489–493.

[25] A. Agrawal, Y. Sun, J. Barnwell, R. Raskar, Vision-guided robot system for picking objects by casting shadows, The International Journal of Robotics Research (2009).

[26] M.-Y. Liu, O. Tuzel, A. Veeraraghavan, Y. Taguchi, T. K. Marks, R. Chellappa, Fast object localization and pose estimation in heavy clutter for robotic bin picking, The International Journal of Robotics Research 31 (2012) 951–973.

[27] R. Raskar, K.-H. Tan, R. Feris, J. Yu, M. Turk, Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging, in: ACM transactions on

graphics (TOG), volume 23, ACM, pp. 679–688.
[28] P. Maric, V. Djalic, Automatic detection of robot's manipulator end-effector position using SIFT algorithm, Annals of DAAAM & Proceedings (2011).
[29] D. Kruse, J. Wen, R. Radke, A sensor-based dual-arm telerobotic system, IEEE Transactions on Automation Science and Engineering 12 (2015) 4–18.
[30] E. Azar, C. Feng, V. Kamat, Feasibility of in-plane articulation monitoring of excavator arm using planar marker tracking, Journal of Information Technology in Construction 20 (2015) 213–229.
[31] S. Datta, R. Ray, D. Banerji, Development of autonomous mobile robot with manipulator for manufacturing environment, The International Journal of Advanced Manufacturing Technology 38 (2008) 536–542.
[32] J. P. Lewis, Fast normalized cross-correlation, in: Vision interface, volume 10, pp. 120–123.
[33] S. Sun, H. Park, D. R. Haynor, Y. Kim, Fast template matching using correlation-based adaptive predictive search, International Journal of Imaging Systems and Technology 13 (2003) 169–178.
[34] B. Babic, N. Nesic, Z. Miljkovic, A review of automated feature recognition with rule-based pattern recognition, Computers in Industry 59 (2008) 321–337.
[35] P. Rummel, W. Beutel, Workpiece recognition and inspection by a model-based scene analysis system, Pattern Recognition 17 (1984) 141–148.
[36] A. Leonardis, H. Bischof, Robust recognition using eigenimages, Computer Vision and Image Understanding 78 (2000) 99–118.
[37] P. Viola, M. J. Jones, Robust real-time face detection, International journal of computer vision 57 (2004) 137–154.
[38] C. Szegedy, A. Toshev, D. Erhan, Deep neural networks for object detection, in: Advances in Neural Information Processing Systems, pp. 2553–2561.
[39] A. Koschan, M. Abidi, Detection and classification of edges in color images, Signal Processing Magazine, IEEE 22 (2005) 64–73.
[40] R. Lienhart, J. Maydt, An extended set of haar-like features for rapid object detection, in: Image Processing. 2002. Proceedings. 2002 International Conference on, volume 1, IEEE, 2002, pp. I–900.
[41] J. Friedman, T. Hastie, R. Tibshirani, others, Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors), The annals of statistics 28 (2000) 337–407.
[42] A. Telea, An image inpainting technique based on the fast marching method, Journal of graphics tools 9 (2004) 23–34.
[43] M. Sokolova, G. Lapalme, A systematic analysis of performance measures for classification tasks, Information Processing & Management 45 (2009) 427–437.
[44] OpenCV dev team, OpenCV 2.4.11.0 API Reference. Camera Calibration and 3D Reconstruction, Technical Report, 2014.
[45] Z. Zhang, A flexible new technique for camera calibration, IEEE Transactions on pattern analysis and machine intelligence 22 (2000) 1330–1334.
[46] L. Breiman, Random forests, Machine learning 45 (2001) 5–32.
[47] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.
[48] V. Lepetit, F. Moreno-Noguer, P. Fua, Epnp: An accurate o (n) solution to the pnp problem, International journal of computer vision 81 (2009) 155–166.
[49] Z. Zivkovic, Improved adaptive Gaussian mixture model for background subtraction, in: Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, volume 2, IEEE, 2004, pp. 28–31.
[50] T. B. Moeslund, A. Hilton, V. Krger, A survey of advances in vision-based human motion capture and analysis, Computer vision and image understanding 104 (2006) 90–126.
[51] T. Bouwmans, F. El Baf, B. Vachon, Background modeling using mixture of gaussians for foreground detection-a survey, Recent Patents on Computer Science 1 (2008) 219–237.