Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computers in Industry 56 (2005) 943–957
COMPUTERS IN INDUSTRY
www.elsevier.com/locate/compind

Salient feature extraction of industrial objects for an automated assembly system

Yuichi Motai *

Department of Electrical and Computer Engineering, University of Vermont, Burlington, Vermont 05405, USA

Received 1 December 2004; received in revised form 31 March 2005; accepted 31 May 2005
Available online 26 September 2005

Abstract

This paper presents a vision based human–robotic interaction (HRI) framework for the modeling and localization of industrial objects typically found in an assembly task. Automating robotic vision for complicated industrial objects is an important, yet still difficult task, especially in the stage of extracting object features. To tackle this specific problem, we have developed a new HRI system consisting of an off-line vision model acquisition, in which the object's salient features are acquired through a human-in-the-loop approach. Subsequently, two feature extraction algorithms; region-growing and edge-grouping, are applied to the model development through collaboration between the human and robot. Finally, using a Kalman filter estimation with a proper ellipse representation, our object localization system generates ellipse hypotheses by grouping edge fragments in the scene, driven by the acquired vision model of objects. The proposed system is validated by experiments using actual industrial objects for both HRI-based object modeling and automated object localization.

Keywords: 3D robot vision system; Salient feature extraction; Human-in-the-loop segmentation; Elliptic edge grouping; Kalman filter estimation

1. Introduction

Robust localization of industrial objects for assembly tasks is an important issue in the robot vision community. A major component of automobile assembly tasks is bin-picking of clustered objects as shown in Fig. 1. The objects in automobile industries, such as several types of alternator covers, hub rotors, and tires have the shape of a large class of curved objects. Each of these objects can be uniquely defined by a salient geometrical feature of a 3D ellipse. It is now generally believed that model-based vision of the kind reported in [1,2,5,6,9] cannot always be driven by mechanical CAD models of objects. Since CAD models must by definition be a complete geometric representation of an object, the resulting representations can be excessively complex and geometrically too rich for robotic vision work.

In Comparison, the representation for a vision system needs to be sufficient only for the purposes of recognition and pose estimation. Obviously, if a vision

* Tel.: +1-802-656-9660; fax: +1-802-656-3358.
E-mail address: ymotai@uvm.edu.

0166-3615/$ – see front matter
doi:10.1016/j.compind.2005.05.020

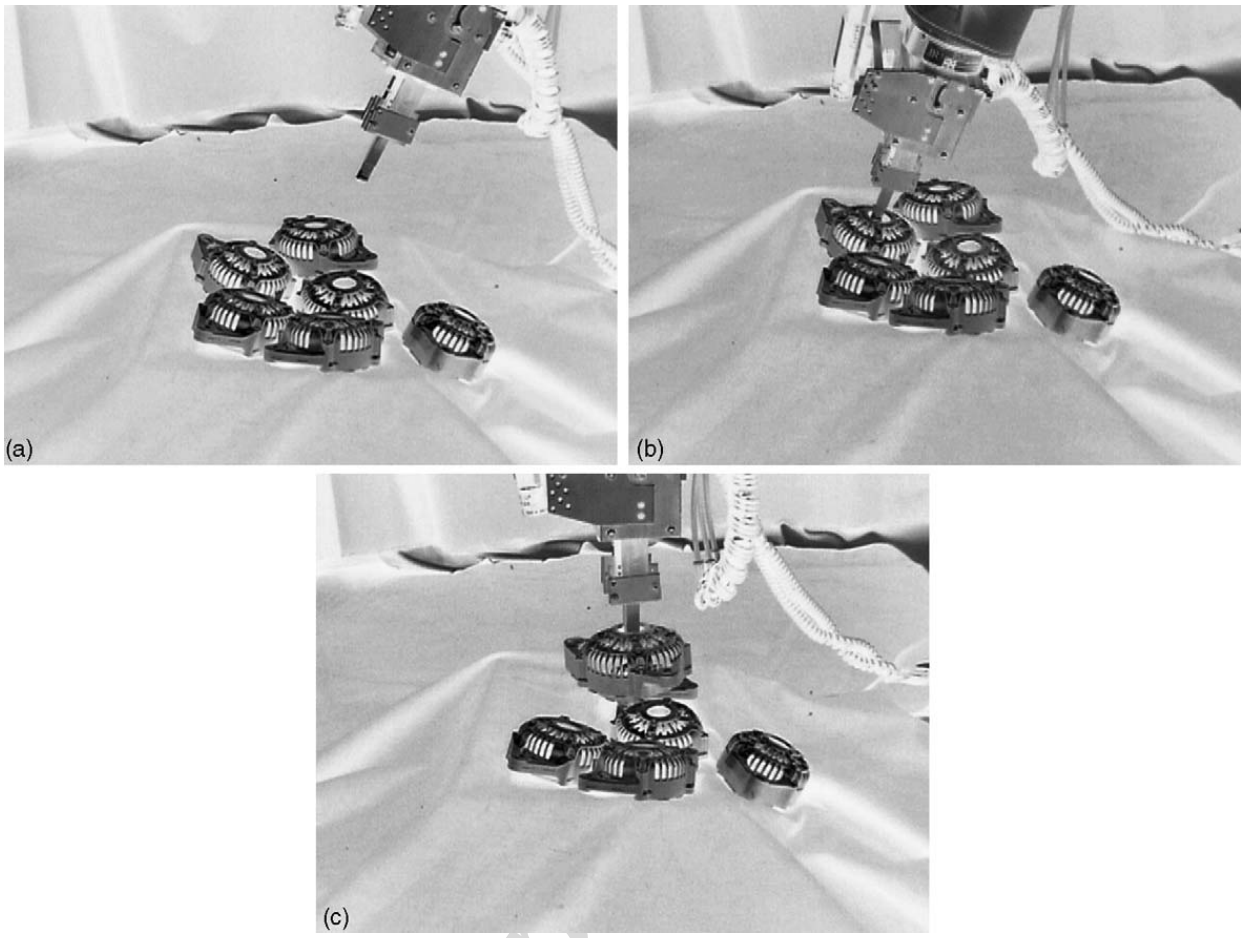

Fig. 1. Robotic manipulation for the localized Alternator Covers 1. The industrial robotic vision computes the pose of complex industrial parts by a single camera mounted on a robot hand that carries the camera to multiple viewing points.

system is expected to see only one type of a rigid object in a clutter-free environment, the needed representation can be very simple and consist of just a small number of features. We will refer to the representation needed for recognition and pose calculation as a vision model.

While it may be possible to derive a vision model from a CAD model, a superior strategy is constructing a vision model directly from the object itself. For the foreseeable future, one would want to develop easy-to-use graphical interfaces to help the computer build the model. In this paper, we will show a new efficient visual-based human–robot interaction (HRI), which would allow a computer to construct vision models with human assistance, at the same time a human could simultaneously verify its inputs with consistency constraints.

In the rest of this paper, Section 2 describes the main issues in this paper regarding feature extraction modules, which essentially determine robustness and accuracy of the object localization. We then present a superior way to represent an ellipse as a salient feature in Section 3. The model feature acquisition method is described in Section 4, and the scene feature extraction method is described in detail with a new edge-grouped feature extraction in Section 5. The two-phase feature matching method will be then presented in Section 6. Finally, the experiment's evaluation results will be shown in Section 7.

2. Specific problems and contributions

Numerical model-based vision systems have been developed to estimate object pose for robotic

manipulation. In 3D vision systems [1,6], the correspondence search between model features and scene features needs to be solved first, and the precise estimation of 3D object pose needs to be then accomplished second. To reliably automate tasks, it is essential to develop robust feature extraction and correspondence matching. In our previous approach [18], for example, the split and merge feature segmentation method was applied for feature extraction of the scene images. Although those extracted features were prominent enough to represent the object, the previous system had limitations for the particular industrial objects, due to lighting illumination, object shading, surface material cracks, and imaging noise (Fig. 2).

Although many other studies have attempted to solve feature extraction problems, a completely successful algorithm is still not available, particularly

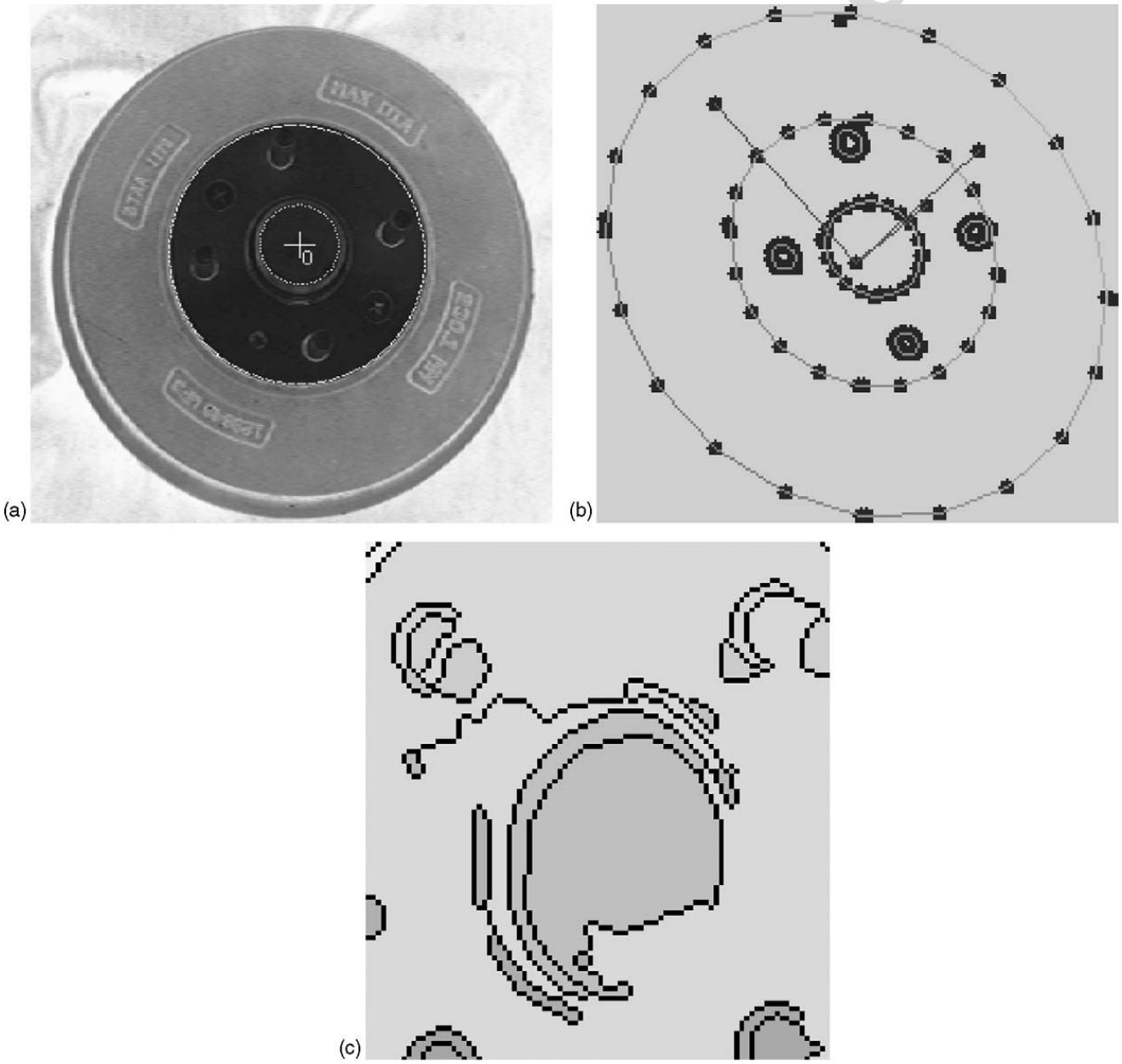

Fig. 2. (a) 2D view of an automobile part Hub. (b) 3D object model generated by Human Computer Interaction editor with grasping coordinates. (c) Example of broken smooth convex curves.

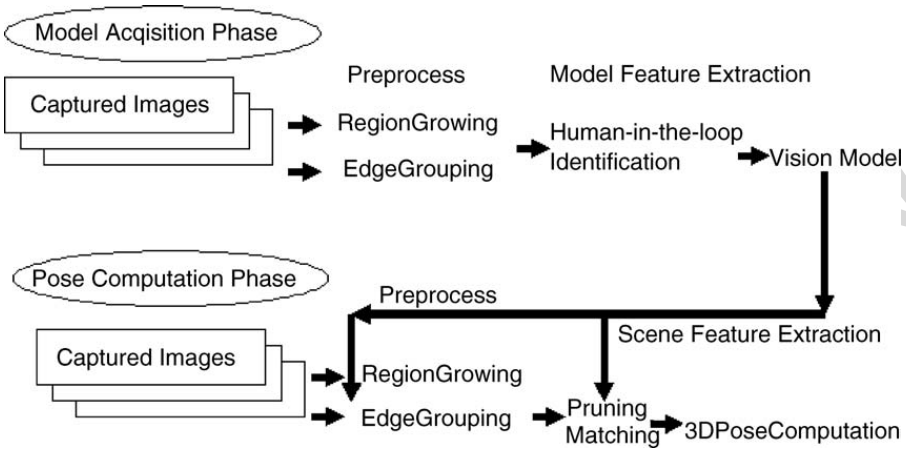

Fig. 3. System diagram of our proposed robotic vision modules. The model acquisition phase in the half upper section is an off-line learning phase (or robot teaching phase) with an HRI framework, and the pose computation phase in the half lower section is an on-line testing phase for automate object localization.

for objects composed of many circular or elliptical features. In ellipse extraction, the least mean squares method is frequently used to fit an accurate ellipse boundary [4,7], although this method is very weak for outliers. The main alternative methods such as Hough transformation [14,22] RANSAC [11], and moments [28], have shown an ability to extract ellipses under some limited cluttered scenes.

Studies of salient features extraction have indicated that edge-grouping methods are promising [15,16]. Such algorithms utilize global salient structures of object contours, inspired by perceptual organization [15]. However, when a target object is too complicated, such edge-grouping has difficulties in extracting the features. Our contribution in this module of image processing is the new edge-grouping method using Kalman filter estimation, which can provide salient ellipse features representing complex object classes.

Another contribution of this paper is to develop a model acquisition system in the framework of HRI, using a human-in-the-loop process to extract the object features for the vision model. We have extended our existing framework [20], or other human-in-the-loop approaches for acquiring vision models [8,12,25–27], by applying it to a fundamental robotic vision system. More specifically, we have developed an interactive framework by emphasizing human perceptual skill. As we will describe later in detail, a human operator supervises the selection of the feature segmentators to construct the visual descriptions of the salient features. More specifically, *the human's perception is embedded in the acquired vision model*. This human perception includes knowledge of the salient feature extracting methods: (1) region-growing segmentation and (2) edge-grouping segmentation. Our comprehensive vision model for the target object helps to solve the localization problem by providing the feature extraction strategy.

The overall modules are illustrated in Fig. 3. In the model acquisition phase, HRI is conducted as an off-line learning process, producing a vision model to be used as a target object. In the pose computation phase, automated robotic vision and robotic manipulation are conducted as on-line processes. Based on the vision model, salient feature extraction and matching are used to compute the 3D pose of the object.

To localize the target objects, the human-in-the-loop vision model *instructs /drives* the feature extraction methods and corresponding matching for the cluttered scene. Our proposed feature-matching approach has the advantage that the system is capable of compensating for the loss of salient feature extraction in the scene which also verifies the localization results. Two major feature-extraction approaches, edge-grouping and region-growing, will increase the likelihood of attempting feature-matching between model and scene.

3. Representation of ellipses

One of our edge-grouping strategies gathers fragments based on attributes such as size, convexity,

and gray-level similarity. Note that *fragment* is a minimum entity of edges that are finely divided into local segments. To group fragments, the computer system checks the elliptic curves by the number of fragments that participate in formation of the ellipses. For each group of fragments, the system estimates the parameters of the hypothesized ellipses using iterative Kalman filtering. In order to formalize the criteria of the parameters, we will first describe the representation of the ellipse.

A parametric elliptic representation and estimation of objects composed of circular parts is very important for edge-based feature extraction. We reconsider the shortcomings in these areas of previous approaches [10,23], and propose a proper parametric form for Kalman filter estimation [17]. Many previous researchers have represented ellipses with parametric equations such as

$$cu^2 + duv + ev^2 + fu + gv + h = 0 \tag{1}$$

or

$$\frac{(u\cos\theta - v\sin\theta - u_0)^2}{a^2} + \frac{(u\sin\theta + v\cos\theta - v_0)^2}{b^2} = 1 \tag{2}$$

where (u, v) in Eqs. (1) and (2) represents any arbitrary points on an ellipse in the 2D image. (c, d, e, f, g, h) in Eq. (1) are coefficients. In Eq. (2), $a$and b represent the lengths of the major and minor axes of the ellipse, θ the orientation of the ellipse, and (u_0, v_0) the center of the ellipse. The first representation Eq. (1) includes arbitrary quadratic forms other than ellipses, and therefore it is difficult to deal with geometric constraints over an ellipse by a set of parameters (c, d, e, f, g, h). Due to the wide range of inexplicitly bounded parameters, elliptic curves may or may not be generated. The second representation equation (2) is more intuitive, but we have observed an instability in estimating the orientation parameter θ, when a and b approach the same value, since small changes in either a or b cause large changes in θ.

Consequently, we have exploited a different approach to represent ellipses which will estimate all proper parameters of elliptic features extracted from the image. As shown in Fig. 4, we represent the ellipse by two focal points P_1 and P_2 and the sum s of

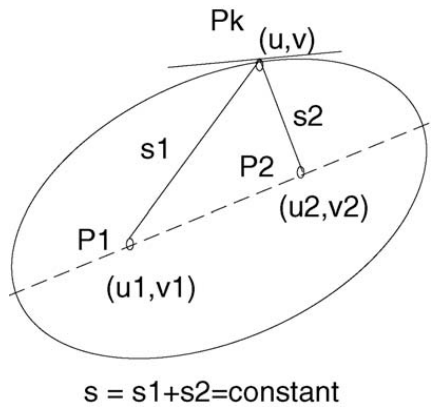

Fig. 4. Ellipse model for parameter estimation.

the distances (s_1, s_2) between the two focal points and any boundary points P_k. Let (u_1, v_1) and (u_2, v_2) be the image coordinates of two focal points P_1 and P_2, and (u, v) be the image coordinate of an arbitrary boundary point P_k. Then

$$f \equiv \sqrt{(u-u_1)^2 + (v-v_1)^2} + \sqrt{(u-u_2)^2 + (v-v_2)^2} - s = 0 \tag{3}$$

Our contribution is to derive the criterion in Eq. (3) using this proper ellipse representation to fit a Kalman filter approach. The parameterization in (3) is very important when applying Kalman filtering. In this specific parametric representation, if the two focal points closely approach each other, the two focal points become simply coincident; therefore no instability for the parameter estimation can be observed. So our problem is the estimation of five ellipse parameters $\boldsymbol{p} = (u_1, v_1, u_2, v_2, s)$ extracted from the images. We will apply Eq. (3) to Kalman filter estimation in Section 5.2.

4. Model feature extraction with the HRI framework

To acquire *salient* features through the HRI graphical user interface (GUI) editor, a human identifies the features representing the object. He/she first selects one of the dominant shapes shown in the grey-highlighted top portion of Fig. 5, and

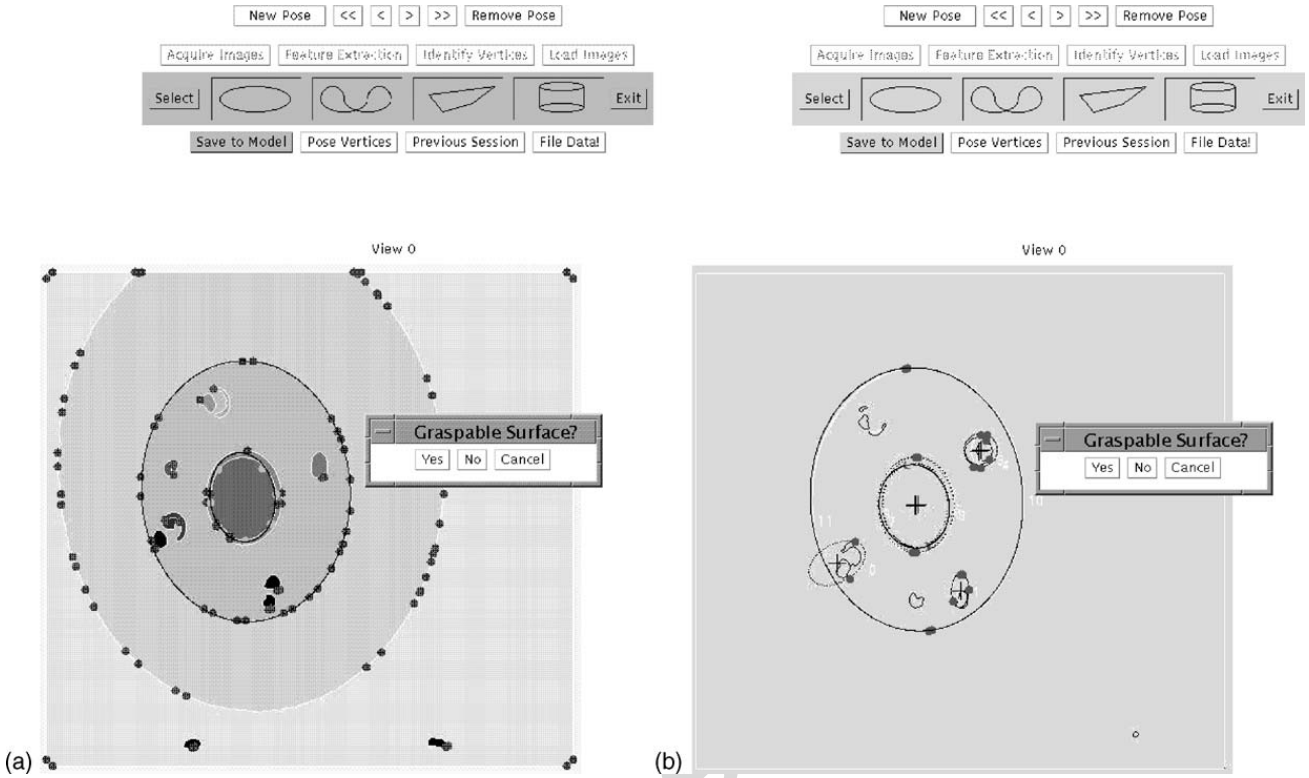

Fig. 5. Teaching features to represent an object model through the HRI graphical user-interface (GUI) editor with reference to robotic vision outputs of (a) region-growing segmentation and (b) edge-grouping segmentation.

specifies the points along the preprocessed contours by clicking the mouse. This process makes use of the assumption that a human perceptually knows which salient feature represents the object well, based on the segmentation of the images with reference to the candidate features extracted by the computer. These salient candidate features are generated by the region-growing and edge-grouping algorithms (explained details in the next section). For example, Fig. 5(a) shows a region-growing segmentation of a Hub object, while Fig. 5(b) shows an edge-grouping segmentation of the same object. For the first case in Fig. 5(a), based on the human's observation of the image processing results, two ellipses are selected for salient features, and region-growing segmentation clearly extracts those regions. The edge-grouping in Fig. 5(b) also extracts those two ellipses, therefore the human chooses these identical ellipses extracted by the edge-grouping segmentation method as a feature verification. These features which are acquired in this manner have several important attributes, such as perimeter, area, shape complexity, and gray level mean. The 2D primitives are then integrated by a human-assisted stereo matching process so that the 3D salient primitives can be computed using stereo triangulation [19,20].

In our human-in-the-loop feature extraction procedure, as shown by another complex object in Fig. 6 (a), a human observes and chooses which feature extraction method the computer should apply for the specific object. Again, using the reference to the segmented images Fig. 6(b) and (c), he/she may choose to represent ellipses by drawing on the top of the edge-grouping result. Note that in Fig. 6(c), black dot curves result from the system's edge-grouping procedure, white edges indicate fragments, and bright white points show end-points of the fragments. Using Fig. 6(b), the vision model also includes supporting features in addition to salient features. These secondary features are used for the verification of salient feature matching later in the localization phase. The vision model used is specific to the target object. For example, the strategy for Alternator Cover 2 includes extracting salient

features using edge-grouping segmentation and applying secondary feature extraction using region-grouping segmentation. In the case of Alternator Cover 1, the salient features are extracted by region-growing through the human's inspection of the segmentation results. The edge-grouping is then used to extract supporting features by aggregating the fragmental edges in the scene until generating compatible salient features. In this model-driven manner, we propose a new strategy of combining the two-feature extraction methods in order to increase the robustness of feature extraction and matching which is addressed in Section 6.

As such, our model acquisition system for learning the object model not only produces a 3D geometrical shape of the object, but also a strategy for the feature extraction methods. The human and computer have different responsibilities in our HRI framework. Our comprehensive vision model for the target object helps to solve the localization problem by providing a feature extraction strategy, which is described in the next section.

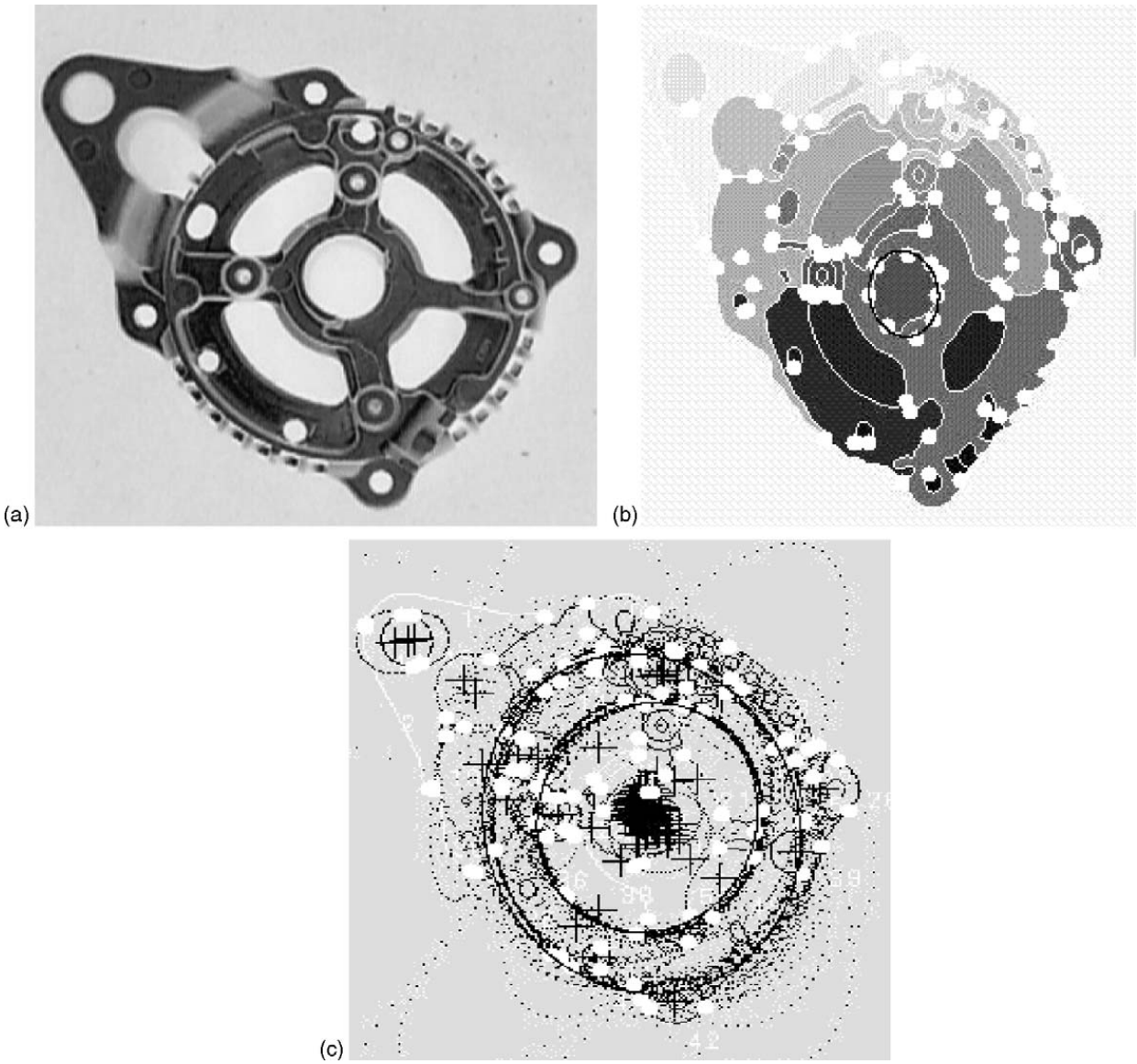

Fig. 6. Outputs displayed to a human operator in order to identify a feature extraction strategy for an image scene of Alternator Cover 2: (a) captured image; (b) region-growing segmentation result (grey level regions); (c) edge-grouping segmentation result (black dot curves).

5. Scene feature extraction in automated robotic vision

For the object scene, the system automatically extracts the salient features using either the region-growing segmentation method or the edge-grouping segmentation method based on the vision model and its embedded strategy. In this section, we will describe the following two segmentation methods for extracting object features from the scene images.

5.1. Region-growing method

The region-growing method extracts the segmented areas with both simple edge tracking and edge linking in order to close contours by performing splitting and merging [13] as shown in Fig. 7 (a). This region extraction algorithm has the advantage that the shape of the regions is well-preserved, especially on the surface boundaries. The outline of our developed algorithm [18,24] can be described as followings:

- Canny edge detector [3].
- Edge linking, especially around T-junctions to produce continuous edge contours.
- Split-and-merge segments of homogeneous regions when edge pixels do not exist within their interior.

The outputs of the above procedures are regions with their complete description given by the quadtree structure and the required attribute measures that was specified by the elliptic object model. After applying attribute constraints based on the vision model, such as area, circularity, shape complexity, perimeter, and average gray level with deviation, many segmented regions in the scene are pruned out as shown in Fig. 7(b). The boundary under the constraints, if it exists, is fitted to a 2D elliptic shape in the primitive model.

5.2. Edge-grouping method

Edge-grouping is an important process since most of the edge extraction methods provide broken edges or curves when the objects are located in an uncontrolled environment. Our edge-grouping method described in this section will compensate for this breakage of edges and form potential ellipses in the image. Matching individual fragments with model features such as ellipse boundaries is difficult. We

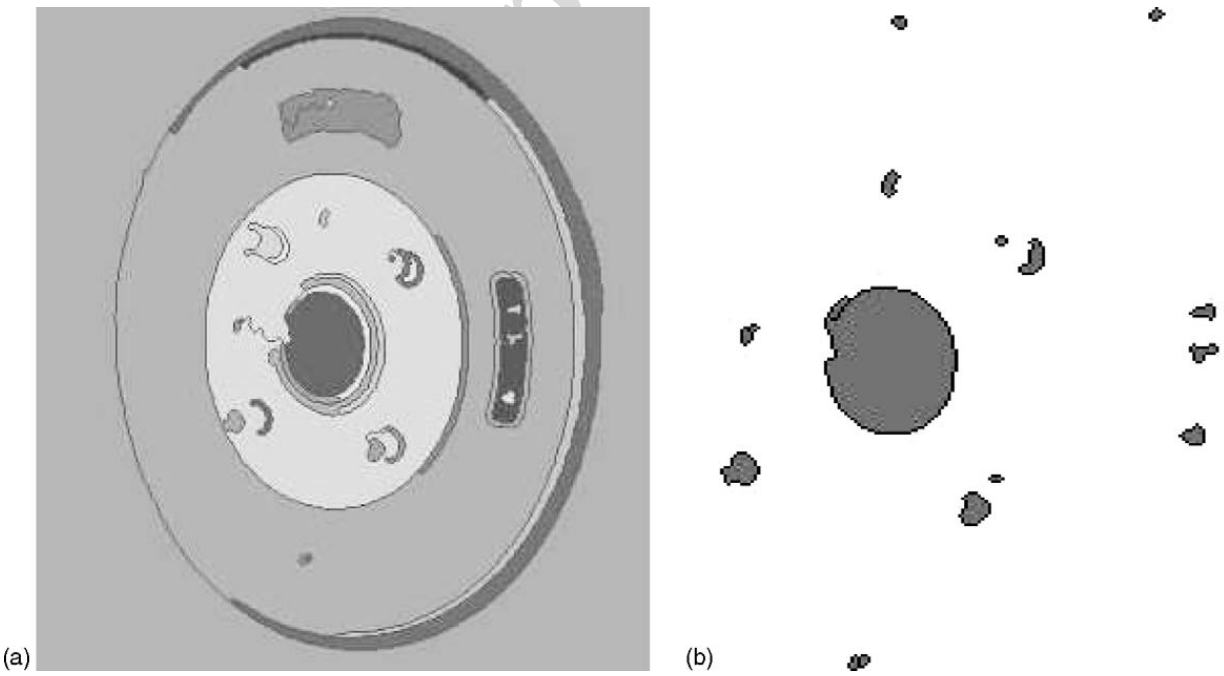

Fig. 7. Feature extraction procedure for the object scene: (a) applying region-based segmentation; (b) pruning the segmented regions by attribute constraints [region-growing method].

would like to group the fragments so that the generated groups will match the model features. Fig. 8 illustrates the procedure to extract ellipse candidates in the scene image based on grouping fragments. The system checks the elliptic curves by the number of fragments participating in forming the ellipses. Our method of edge grouping consists of the following three procedures: (i) fragment extraction, (ii) hypothesis generation for ellipses from grouped fragments, and (iii) hypothesis verification for ellipses by parameter fitting.

In the first process (i) fragment extraction, the following steps are taken:

- Canny edge detector.
- Extraction of curved fragments by examining endpoints and junctions.
- Division of curved fragments into smooth convex curves.

In the second process (ii) hypothesis generation, the following geometric constraints are taken for pair-wise grouping of fragments in order to check the initial discard of hypothesis:

(1) *Size constraint*: The size of the ellipse in the image is limited. For each group of fragments, the combined curves must be smaller in size than a certain threshold based on the object model.
(2) *Convexity constraint*: Any pair of fragments must not violate convexity when these fragments are combined. The fragments are checked at two endpoints and the center of mass, and compared to the other fragments.
(3) *Gray-level similarity constraint*: Any pair of fragments must possess gray level similarity, in either the internal or the external region of the elliptic curve. If the object region is homogeneous along the elliptic curve, then two fragments must satisfy the similarity gray level.

In the third process (iii) hypothesis verification, the following steps are taken:

- Generation of initial estimate for potential ellipses: Given a set of points along a group of fragment curves, $\boldsymbol{p} = (u_1, v_1, u_2, v_2, s)$ in Eq. (3) is to be estimated. To compute the initial estimate of $\boldsymbol{p}$, we first compute the centroids and the moment of inertia for the image points participating in the fragment set. The initial estimates of (u_1, v_1) and (u_2, v_2) are computed as the above centroids. The initial estimate of s is computed by the sum of the

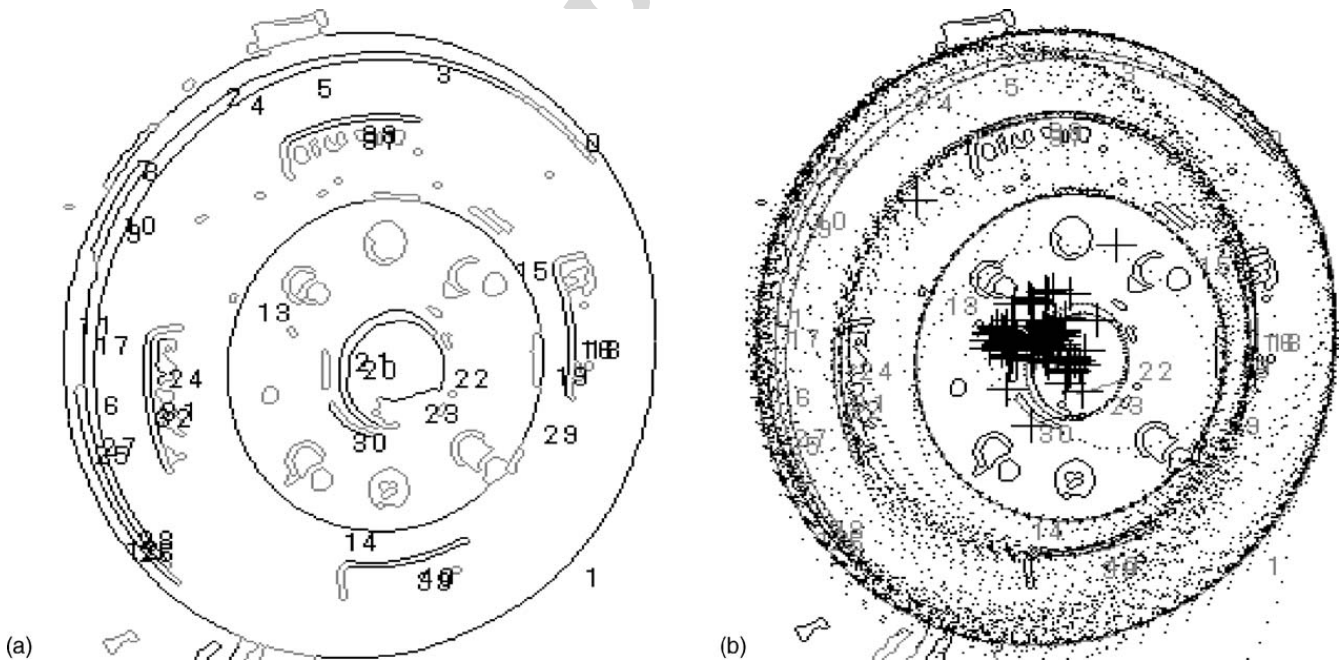

Fig. 8. Feature extraction procedure for the object scene: (a) curve fragments extraction (indexed) along edges; (b) group of fragments extraction (dotted) by aggregation through elliptical parameter estimation [edge-grouping method].

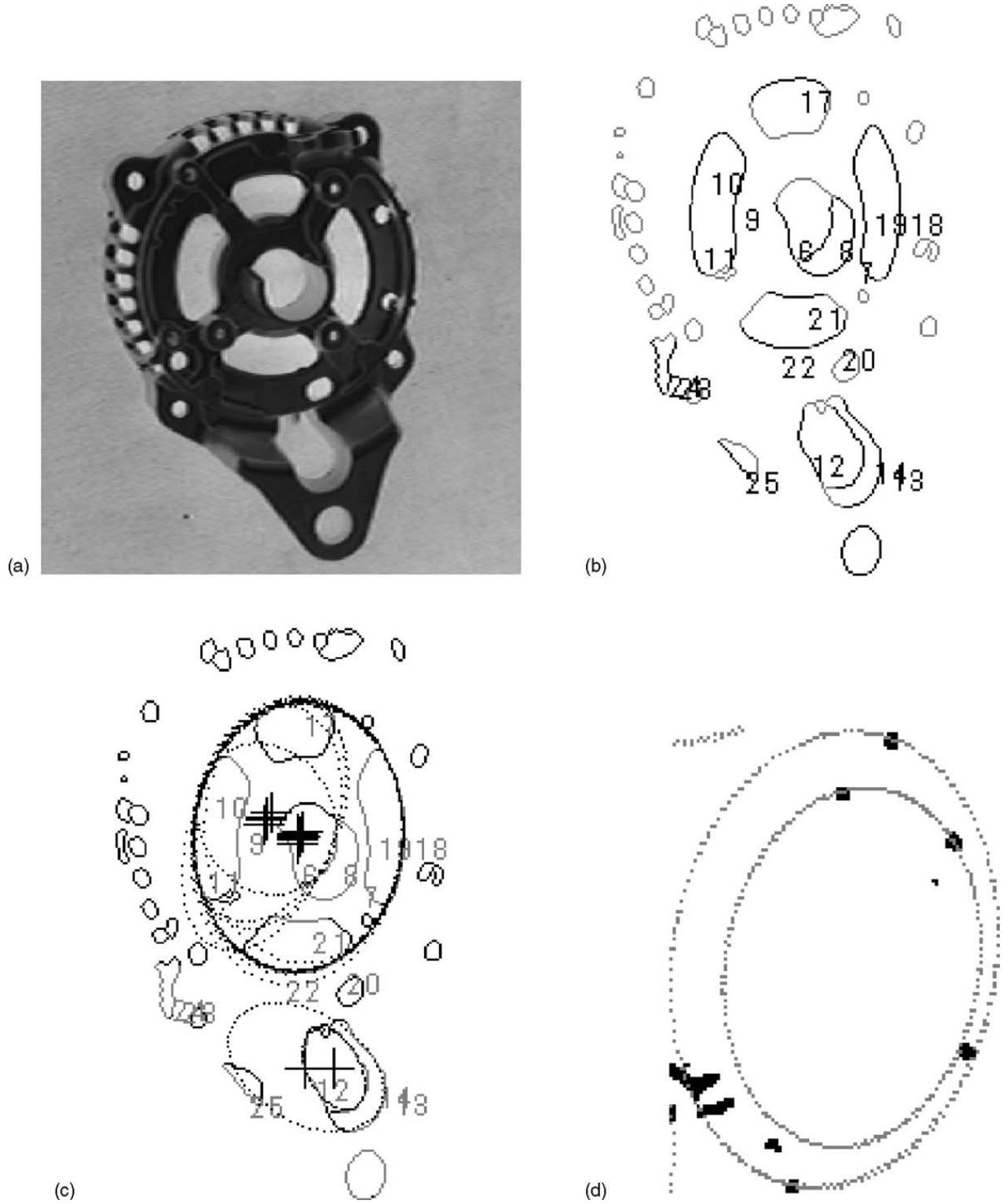

Fig. 9. Edge-grouping sequential process: (a) object scene, (b) curve fragments extraction along edges, (c) group of fragments extracted by aggregation through elliptical parameter estimation, and (d) small holes extracted along hypothesized regions.

lengths of two axes, $s1 + s2$, spanned by the moment of inertia.

- Verification by fine parameter fitting using the Kalman filter: The system selects representative points $z = (u^i_j, v^i_j)$ from the fragments by selecting equally spaced j points along the boundary curves. The system selects at least 16 points $(j = 0, 1, \ldots, 15)$ for each fragment i. By applying the Kalman filter to the constraint equation of (3) for every selected boundary point (u, v), the system updates the ellipse parameter $\boldsymbol{p}$. Once we obtain an estimate of $\boldsymbol{p}$, the next step is to compute the fitting error by computing the constraint equation of (3). The fitting error is the average sum of $f$for the representative points. If the fitting error exceeds a certain threshold, then we discard this hypothesis.

6. Two-phase feature matching strategy using vision models

The segmentation strategy embedded in the vision model determines the feature matching procedure. Using a few representative numbers of salient features in the model, the automated robotic vision system searches for a correspondence by checking the intrinsic attributes of the representative salient features in the object scene. The robot vision system then applies a second feature extraction method to the supporting features in order to verify the hypotheses created by the first matching. We call this approach two-phase feature matching. The vision model specifies which feature extraction modules are first applied and which matching modules are subsequently applied. Since the main focus in this paper is the HRI framework, we will briefly show a matching flow of the Alternator Cover 2, using Fig. 9.

7. Experimental results of feature extraction and matching

We implemented the off-line HRI GUI editing process for learning object models using a SUN Workstation. It took several minutes for a human and a computer to interactively generate each object model to learn features. Our wrist-mounted robotic vision systems consisted of a Sony DC-47 monocular 1/3 in. CCD camera with Pulnix Lens of focal-length 16 mm, an industrial robot PUMA 761 or Kawasaki JS10, and an Intel Personal Computer running Linux. We mounted a monocular camera on the robotic manipulation gripper for capturing multiple images of the object from different viewpoints. A calibration matrix was computed for stereo reconstruction from arbitrary viewpoints [21].

We used two or three views to automatically localize the object in the testing phase, and five views

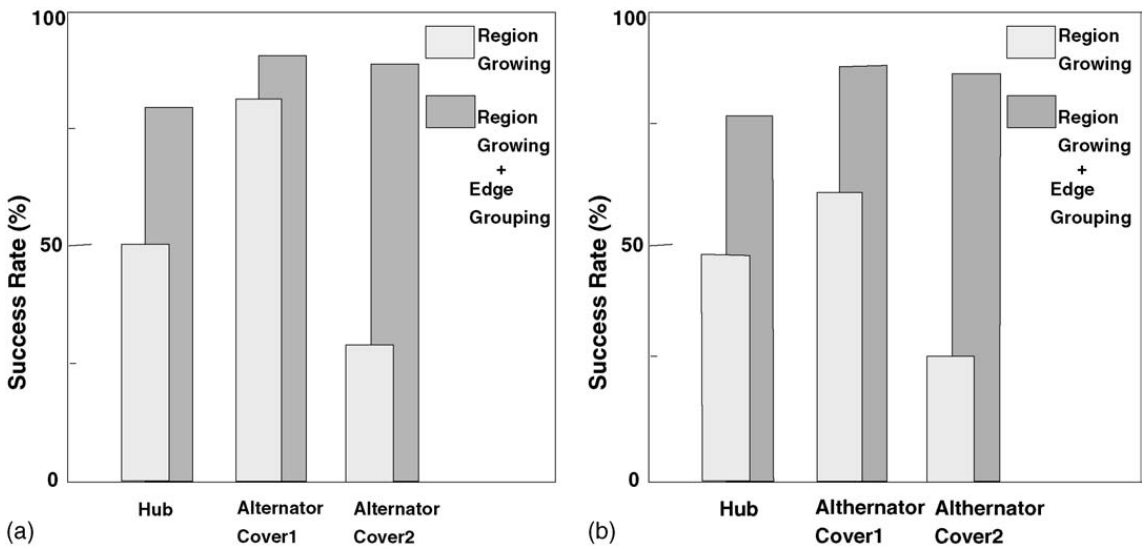

Fig. 10. (a) Feature extraction success rate, (b) feature matching success rate. Previous method (region-growing only) vs. our proposed method (region-growing + edge-grouping)for the objects: Hub and Alternator Covers 1 and 2.

to generate the object model in the learning phase. After the vision model was acquired in this manner, the same object in a random pose was captured from two or three viewpoints for automatic 3D pose calculation. Once the feature matching between model and scene was achieved, then the 3D translation and rotation from the model coordinate to the scene coordinate was computed using the quaternion approach [6]. The robot gripper picked up the object or performed a peg-in-hole alignment through its 3D localization. Three types of automobile industrial objects were analyzed for localization: Hub and two different types of Alternator Covers (Types 1 and 2). Thirty-five scenes were captured, each containing 4–6 pieces of randomly placed objects. The total number of objects across all the scenes was proximately 140 pieces. We evaluated the *success rate* and the *accuracy* of our two-phase feature matching algorithm.

The *success rate* of feature extraction was 81% for Hub, and 91% and 89% for Alternator Covers 1 and 2

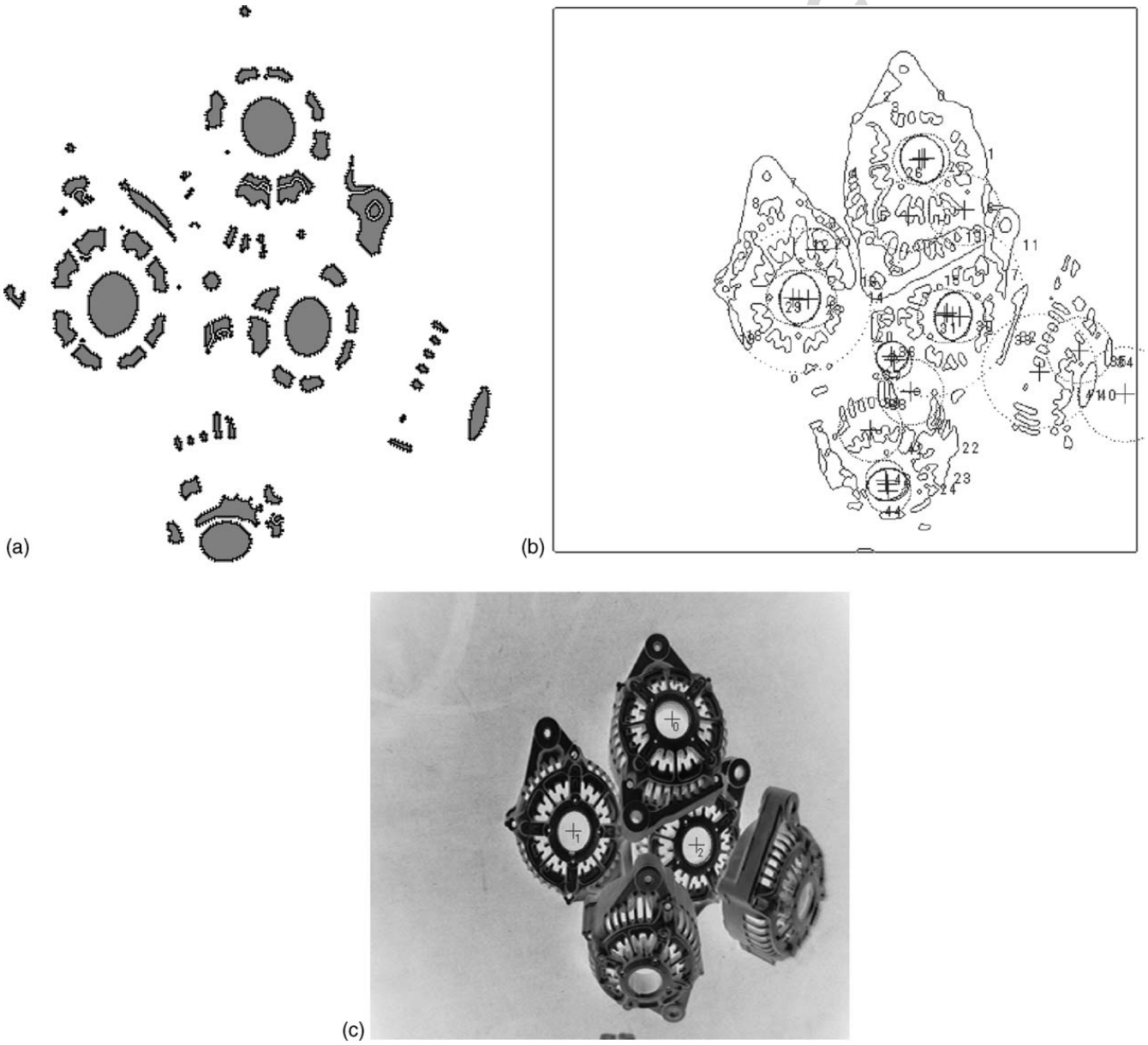

Fig. 11. (a) region-base feature boundary (enhanced contour), (b) edge-grouping extraction, (c) model superimposition based on region-match [Alternator Cover1].

(out of 140 pieces for overall trials) as shown in Fig. 10 (a). Also our final localization success rate in Fig. 10(b) was 77% for Hub, 87% for Alternator Cover 1, and 85% for Alternator Cover 2. Compared to the single feature extraction method (our existing system [18]), our proposed dual feature extractions with HRI GUI framework have made improvements over previous success rates. If the automated robotic vision system is allowed to recapture the object scene, the overall success rate would be close to the most reliable rate ($\sim$100%).

The localization *accuracy* was also evaluated by the translation and rotation in the world coordinates. The quantitative error, shown in Table 1, was computed with the results of Fig. 10(b) by showing the average error with the standard deviation. The outline diameter was 240 mm for Hub, and 30 and 65 mm for Alternator Covers 1 and 2 of a symmetric

Table 1
Localization error result of three objects

Object	Process	Translation (mm)	Rotation (°)
Hub	Region-match	3.4 ± 1.9	11.3 ± 5.7
	Group-match	3.5 ± 2.0	12.0 ± 5.1
Alternator Cover 1	Region-match	1.4 ± 0.5	4.3 ± 1.6
	Group-match	1.1 ± 0.4	4.0 ± 1.0
Alternator Cover 2	Group-match	1.9 ± 0.7	4.4 ± 1.8
	Region-match	2.2 ± 1.1	5.1 ± 2.2

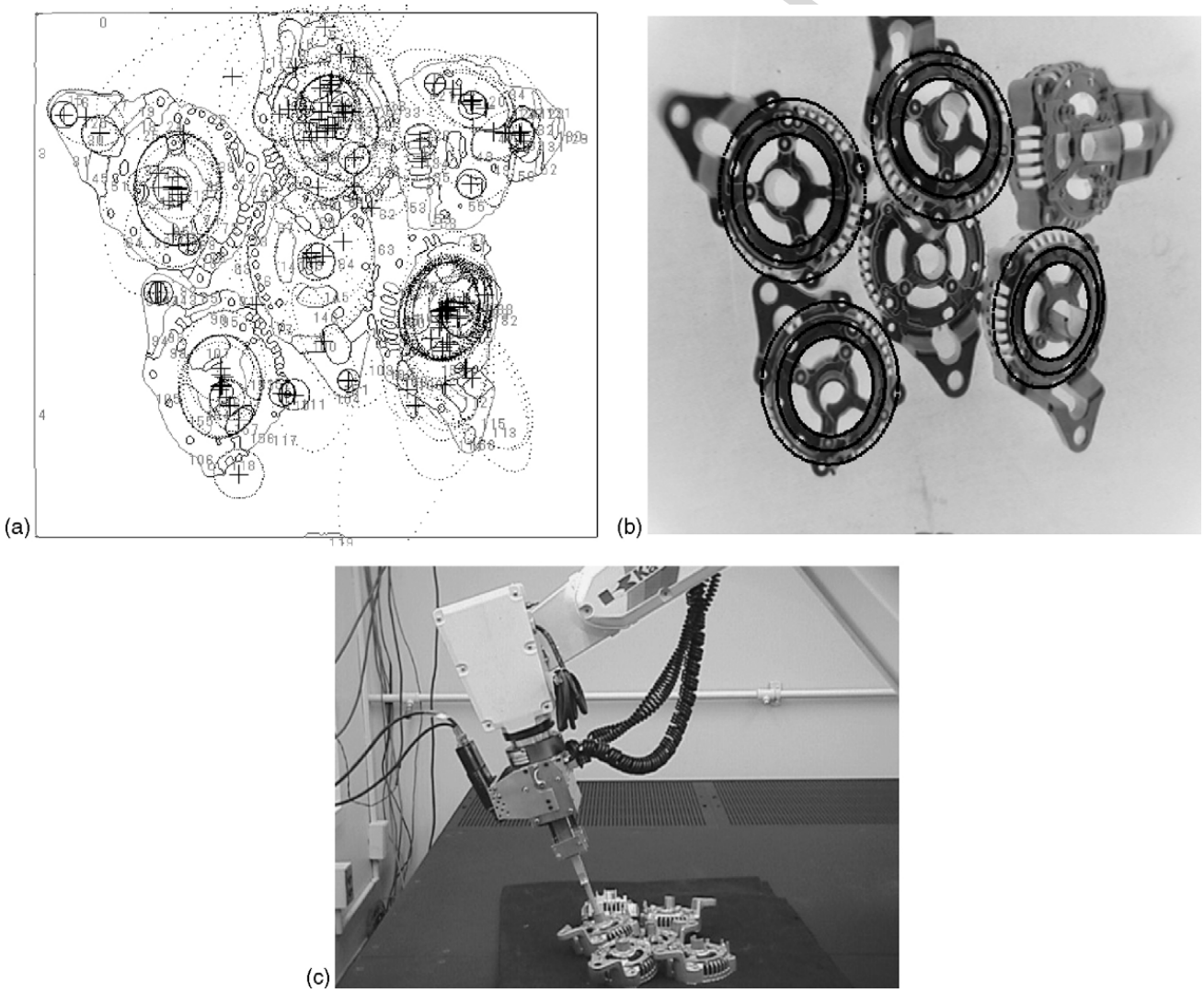

Fig. 12. (a) edge grouping, (b) 2D superimposition of model onto scene, and (c) robot manipulation demonstration from a close view [Alternator Cover 2].

circular outline, respectively. For Alternator Cover 1, Fig. 11 (a) shows a typical result of region-match, and Fig. 11(b) is a result of Group-match. These results provide the localization parameters are superimposed by the object 3D model onto the 2D scene. As for Alternator Cover 2, Fig. 12 (a) shows ellipses hypotheses, which are formed by edge grouping. Obviously for these multiple complex objects, our proposed system achieved improvements in both robustness and accuracy. These localization results were verified again through the final robotic manipulation. Since we have not optimized the automated robotic vision modules, we cannot determine the final time cost necessary for implementing an overall real-time vision algorithms. The proposed new framework of HRI for the robotic vision system discussed in this paper is used routinely for a reliable demo in Purdue's Robot Vision Lab.

8. Conclusion

We have developed a new visual-based HRI for extracting salient features of an industrial object for an automated assembly system. This system consists of the following two phases.

In the phase of visual-based HRI, human and robot vision acquired a vision model for the target object to be manipulated, through a human-in-the-loop GUI editor. A human operator supervised the selection of the feature segmentators in order to construct visual descriptions of the salient features. Thus, the human's perception was embedded in the acquired vision model. The human perception included knowledge of the extracting methods of salient features: (1) region-growing segmentation and (2) edge-grouping segmentation. Our comprehensive vision model for the target object helped to solve the localization problem by providing a feature extraction strategy.

In the phase of object pose computation driven by the acquired vision model, the automated robot vision: (1) derived an efficient ellipse representation for Kalman filter estimation, (2) built a competent edge-grouping method to generate ellipse hypotheses in a complex object, and (3) established a two-phase feature matching algorithm.

The contributions in this paper are as follows: A proper parametric representation of an ellipse feature was achieved for optimizing the criterion. Our human-in-the-loop vision model provided a salient feature acquisition strategy so that the computer system can compensate the loss of salient feature extraction in the object scene. Using each group of edge fragments, the system estimated the parameters of ellipse hypotheses using an extended Kalman filtering. The salient features were extracted from the object scene based on the feature acquisition strategy embedded in the vision model by a concatenation of both region-growing and edge-grouping methods. The evaluation results of three different objects verified that the two-phase feature matching method improved the success rate of object localization.

For further extension of our work, the human-in-the-loop module for the salient feature extraction would be enhanced. Because a human "draws" features on the top of the images, the system can determine which feature extraction method generated by the computer approximates the choice of the feature extraction perceptually generated by the human. In this adaptive manner of further HRI, the feature extraction should modify itself to the human's choices for updating the threshold and parametric setting.

Acknowledgment

The author wishes to express his gratitude to A. Kosaka and A.C. Kak for their comments and help during his research in the Robot Vision Laboratory at the Purdue University.

References

[1] R.C. Bolles, P. Horaud, 3DPO: a three-dimensional part orientation system, International Journal of Robotics Research 5 (3) (1986) 3–26.
[2] C. Bradley, Rapid prototyping models generated from machine vision data, Computers in Industry 44 (2) (2001) 159–173.
[3] J. Canny, A computational approach to edge detection, IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (6) (1986) 679–698.
[4] C. Chatterjee, E.K.P. Chong, Efficient algorithms for finding the centers of conics and quadrics in noisy data, Pattern Recognition 30 (5) (1997) 673–684.

[5] C.H. Chen, A.C. Kak, Robot vision system for recognizing 3D objects in low-order polynomial time, IEEE Transactions on Systems, Man and Cybernetics 19 (6) (1989) 1535–1563.
[6] O.D. Faugeras, Three-dimensional Computer Vision, MIT Press, 1993.
[7] A. Fitzgibbon, M. Pilu, R.B. Fisher, Direct least square fitting of ellipses, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (5) (1992) 476–480.
[8] A.R.J. François, G.G. Medioni, Interactive 3D model extraction from a single image, Image and Vision Computing 19 (6) (2001) 317–328.
[9] L. Grewe, A.C. Kak, Interactive learning of a multiple-attribute hash-table classifier for fast object recognition, Computer Vision and Image Understanding 61 (3) (1995) 387–416.
[10] F. Guichard, J. Tarel, Curve finder combining perceptual grouping and a Kalman like fitting, Proceedings of IEEE International Conference on Computer Vision 2 (1997) 1003–1009.
[11] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2000.
[12] S. Heuel, R. Nevatia, Including interaction in an automated modeling system, Proceedings of the International Symposium on Computer Vision (1995) 383–388.
[13] S.L. Horowitz, T. Pavlidis, Picture segmentation by a tree traversal algorithm, Journal of ACM 23 (2) (1976) 368–388.
[14] P.V.C. Hough, Method and means for recognizing complex patterns, US Patent 3,069,654 (1962).
[15] D.P. Huttenlocher, P.C. Wayner, Finding convex edge groupings in an image, International Journal of Computer Vision 8 (1) (1992) 7–27.
[16] D.W. Jacobs, Robust and efficient detection of salient convex groups, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (1) (1996) 23–37.
[17] A. Kosaka, A.C. Kak, Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertainties, Computer Vision, Graphics, and Image Processing-Image Understanding 56 (3) (1991) 271–329.
[18] A. Kosaka, A.C. Kak, Stereo vision for industrial applications, in: Handbook of Industrial Robotics, Nof SY *edit*, John Wiley & Sons, Inc., 1999, pp. 269–294.
[19] Y. Motai, Ph.D thesis, School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, 2002.
[20] Y. Motai, A.C. Kak, An interactive framework for acquiring vision models of 3D objects from 2D images, IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics 34 (2004) 566–578.
[21] Y. Motai, A. Kosaka, SmartView: hand-eye robotic calibration for active viewpoint generation and object grasping, Proceedings of IEEE International Conference of Robotics and Automation (2001) 2183–2190.
[22] C.F. Olson, Improving the generalized Hough transform through imperfect grouping, Image and Vision Computing 16 (1998) 627–634.
[23] J. Porrill, Fitting ellipses and predicting confidence envelopes using a bias corrected Kalman filter, Image and Vision Computing 8 (1990) 37–41.
[24] K. Rahardja, A. Kosaka, Vision based bin-picking: recognition and localization of multiple complex objects using simple visual cues, Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems 3 (1996) 1448–1457.
[25] C.R. Shyu, C.E. Brodley, A.C. Kak, A. Kosaka, A.M. Aisen, L.S. Broderick, ASSERT: A physician-in-the-loop content-based image retrieval system for HRCT image databases, Computer Vision and Image Understanding, Special Issue on Content-Based Access of Image and Video Libraries 75 (1/2) (1999) 111–132.
[26] M.L. Smith, G. Smith, T. Hill, F. Meyer, An environment for off-line configuration and programming of a vision based inspection system using CAD data, The Caledonian International Engineering Journal (1) (1996) 31–39.
[27] L.N. Smith, P.S. Midha, A knowledge based advisory system for optimum and concurrent design and manufacture of powder metallurgy components, International Journal of Production Research 37 (1) (1998) 125–137.
[28] K. Voss, H. Suesse, Invariant fitting of planar objects by primitives, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (1) (1997) 80–84.

Yuichi Motai is currently an assistant professor of Electrical and Computer Engineering at the University of Vermont, USA. He received a Bachelor of Engineering degree in Instrumentation Engineering from Keio University, Japan in 1991, a Master of Engineering degree in Applied Systems Science from Kyoto University, Japan in 1993, and a Ph.D. degree in Electrical and Computer Engineering from Purdue University, USA in 2002. He was a tenured Research Scientist at the Secom Intelligent Systems Laboratory, Japan from 1993 to 1997. His research interests are in the broad area of computational intelligence; especially of computer vision, human–computer interaction, ubiquitous computing, sensor-based robotics, and mixed reality.