Proc. IEEE 5th Int. Conf. on Computer Vision
pp.1106-1113, Cambridge, MA, June 1995

Task-Oriented Generation of Visual Sensing Strategies*

Jun Miura

Department of Mechanical Engineering
for Computer-Controlled Machinery
Osaka University
jun@cv.ccm.eng.osaka-u.ac.jp

Katsushi Ikeuchi

School of Computer Science
Carnegie Mellon University
ki@cs.cmu.edu

Abstract

In vision-guided robotic operations, vision is used for extracting necessary information for achieving the task. Since visual sensing is usually performed with limited resources, visual sensing strategies should be planned so that only necessary information is obtained efficiently. This paper describes a method of systematically generating visual sensing strategies based on knowledge of the task to be performed. The generation of the appropriate visual sensing strategy entails knowing what *information to extract,* where *to get it, and* how *to get it. This is facilitated by the knowledge of the task, which describes what objects are involved in the operation, and how they are assembled. Our method has been implemented using a laser range finder as the sensor. Experimental results show the feasibility of the method, and point out the importance of task-oriented evaluation of visual sensing strategies.*

1 Introduction

In vision-guided robotic operations, visual sensing strategies should be planned so that only necessary information is obtained efficiently. To determine an efficient visual sensing strategy, knowledge of the task is necessary. Without knowledge of the task, it is often difficult to select the appropriate visual features to be observed. In addition, resources may be wasted in tracking uninformative features.

From this standpoint, research on *task-oriented vision*, *active vision*, or *purposive vision* has been actively investigated [1] [3] [6]. By using knowledge of the task, the vision system can be designed to be fast and robust. However, the designing process tends to be task-specific and requires a significant amount of effort. Thus, it is desirable to develop a systematic method which can generate *task-oriented* visual sensing strategies automatically, namely a method that optimizes each visual sensing strategy according to a given task.

The generation of task-oriented sensing strategy is decomposed into the following three subproblems to be solved successively:

- determine what visual information is necessary for the current task;
- determine which visual features carry such necessary information; and
- determine how to get necessary information with the sensors used.

The first two subproblems are concerned with focusing the attention to informative visual features; the last problem is concerned with evaluation of sensing strategies.

The ability of focusing attention is important to realize efficient visual sensing strategies. There have been several approaches to this problem. Some work [5] [12] is concerned with exploratory visual sensing planning under uncertainty of the knowledge of the scene. in these approaches, visual feature set, from which the observed features are selected, is given in advance; it is not automatically derived from the task description. Some work [4] [9] is concerned with the usage of the task, including the constraints on the environment, for concentrating the visual processing to only necessary portion of image. These approaches are, however, highly task-specific and are based on the careful *a priori* examination of the task.

The third subproblem (i.e., how to get necessary information) is decomposed further into two more specific problems of determining a set of feasible sensing strategies and subsequently selecting the most appropriate one among them. The goal of the former is to determine sensing condition which satisfies several requirements on imaging such as resolution, field of view, focus and visibility [15]. The ability of solving such a problem would be necessary for any sensor planners as a subroutine.

As mentioned above, the second more specific problem is to determine the best sensing strategy which maximizes some "goodness" function. The minimum uncertainty criterion (i.e., to select the sensing strategy which minimizes the uncertainty of information

*This research was sponsored by the Advanced Research Projects Agency under the Department of the Army, Army Research Office under grant number DAAH04-94-G-0006.

to be obtained) has often been used (e.g., [8]). In certain types of tasks, however, some part of information may need to be more accurate than the rest. Thus, some appropriate function should be automatically designed for each task which can evaluate how each sensing strategy contributes to the proper execution of the current task.

This paper proposes a novel method of systematically generating visual sensing strategies based on knowledge of the task. We deal with visual sensing strategy generation in assembly tasks, in which the environment is *known*, that is, the shape, the size, and the *approximate* location of every object is known to the system. In this situation, the role of visual sensors is to determine the position of the currently assembled object with sufficient accuracy so that the object can be, with a high degree of certainty, mated with other objects. In the following sections, we describe new methods in: (1) determining necessary visual information, (2) determining visual features to observe, and (3) evaluating sensing strategies.

2 Determining What Information is Necessary

This section first describes the task analysis based on face contact relations between objects. Then, using the result of the analysis, a theory is derived which determines what visual information is necessary for each assembly operation.

2.1 Task Analysis Based on Face Contact State

We analyze a state of the environment in terms of face contacts between object surfaces [7]. The analysis deals with the case where polyhedral objects perform only translational motions. We assume that each assembly operation involves one manipulated object, manipulated by a robot for the current operation, and several stationary environmental objects which have face contacts with the manipulated object. We also assume that the goal of each assembly operation is to establish the required face contact state.

2.2 Representation of Face Contact States

Let us suppose a surface patch of the manipulated object have a face contact to a surface patch of an environmental object. This surface contact pair constrains the manipulated object's possible translation motion by:

$$\boldsymbol{N} \cdot \Delta \boldsymbol{T} \geq 0,$$

where $\Delta \boldsymbol{T}$ denotes possible translational motion vectors of the manipulated object and $\boldsymbol{N}$ denotes the normal direction of an environmental surface patch.

Possible translation vectors of the object can be represented by using points on the Gaussian sphere

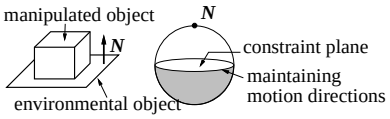


Fig. 1: Constraint inequality depicted on the Gaussian sphere.

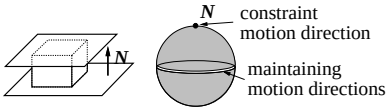


Fig. 2: A bidirectional constraint.

(see Fig. 1). The constraint from a patch pair defines several regions in the Gaussian sphere. Assuming that the normal, $\boldsymbol{N}$, points to the north pole of the Gaussian sphere without loss of generality, the northern hemisphere corresponds to possible motion directions; the southern hemisphere corresponds to prohibited motion directions.

When several surface patches of different orientations make contact, possible motion directions are constrained through simultaneous linear inequalities. These constraints are represented as a combined region in the Gaussian sphere.

In Fig. 1, motions of the directions corresponding to the boundary of the southern hemisphere (the equator) maintain the current face contact state. The degrees of freedom of the maintaining the contact state (*maintaining DOF*) is two. Motions of the directions corresponding to the inside of the detaching hemisphere break the contact state, and is referred to as the *detaching* motion. A pure detaching motion is the detaching motion which does not contain any maintaining motion component. The pure detaching motion in Fig. 1 is along the constraint normal $\boldsymbol{N}$; its degrees of freedom (*detaching DOF*) is one.

Fig. 2 shows the case where two normal vectors of environmental objects have the opposite directions. The possible motion directions of the manipulated object can be represented as *the entire great circle* perpendicular to the axis connecting the two poles. There are no detaching motions; the detaching DOF is zero. One direction along the surface normals is completely constrained; the degrees of freedom of the constraint directions (*constraining DOF*) is one.

We can specify a face contact state by using a triplet of maintaining, detaching, and constraining DOFs. Using this triplet, for example, the states of Figs. 1 and 2 are represented as $(2,1,0)$ and $(2,0,1)$, respectively. Each assembly operation is considered as a *transition* from one face contact state to another. By extracting possible contact states and transitions between them, we get the *transition graph* as shown in Fig. 3.

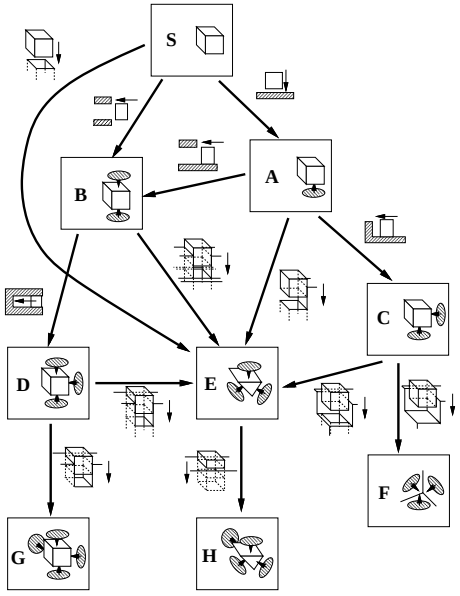


Fig. 3: Contact state transitions represented as a directional graph [7].

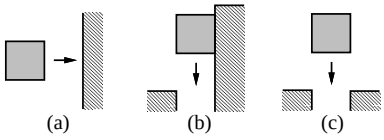


Fig. 4: Three typical cases of increase of constraint on a degree of freedom. Type and transition of the triplet is as follows:
(a): maintaining → detaching $((3,0,0) \rightarrow (2,1,0))$.
(b): detaching → constraining $((2,1,0) \rightarrow (2,0,1))$.
(c): maintaining → constraining $((3,0,0) \rightarrow (2,0,1))$.

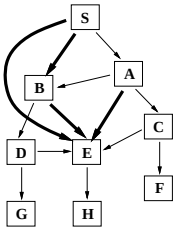


Fig. 5: Classification of state transitions. Bold lines indicate the transitions that require visual information. Thin lines indicate the transitions that do not require visual information.

2.3 Determining What Visual Information is Necessary

An assembly operation always increases constraints on some degrees of freedom of the manipulated object. This increase of constraint is classified into three cases: from maintaining DOF to detaching DOF, from detaching DOF to constraining DOF, and from maintaining DOF to constraining DOF. Fig. 4 shows typical situations corresponding to the three cases.

Let us examine how the type of the degree of freedom for horizontal motion changes in these cases, and how that change is realized. In case (a), the degree of freedom changes from maintaining DOF to detaching DOF. Since the approaching direction of the block is parallel to the direction of the pure detaching motion at the final state (i.e., the normal vector of the wall), this operation is realized by moving the block until the face contact occurs. Thus, this operation can be performed by compliant motion without visual information. In case (b), the degree of freedom changes from detaching DOF to constraining DOF. Although the horizontal degree of freedom is constrained at the final state, visual information is also unnecessary because the desired horizontal position can be kept by maintaining the contact between the block and the right wall.

In case (c), the degree of freedom changes from maintaining DOF to constraining DOF. The horizontal position of the block needs to be adjusted with visual information before mating so that both the left and the right face contact are achieved simultaneously. Since there is no contact before mating, force information cannot be used.[1]

To summarize, if a degree of freedom becomes constraining DOF from maintaining DOF, that degree of freedom should be observed. By applying this theory to thirteen possible state transitions enumerated in Fig. 3, four transitions (S→B, S→E, A→E and B→E) were found to require visual information (see Fig. 5).

3 Selection of Features to be Observed

In each vision-guided assembly operation, a relevant set of visual features to be observed needs to be selected so that necessary degrees of freedom of the assembled objects are observed. This selection is performed as follows (see Fig. 6). From the task description, the degrees of freedom to be constrained by the current assembly operation is obtained. On the other hand, a set of observable features comes from the face contact information in the task description. By consulting *sensing primitives* [10], which describe the relationships between observed features and degrees of freedom to be measured, a feasible set of features is selected. Necessary geometric information is retrieved from a CAD-based world model.[2]

[1] A sophisticated force control-based manipulation strategy may be employed to perform this kind of assembly operation without visual feedback [14]. Even in such a case, reducing errors by visual information would be useful.

[2] The object recognizer determines each object configuration in the real world and generates a CAD-based world model.

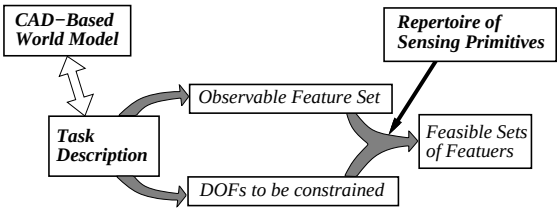


Fig. 6: Selection of features to be observed using task description and sensing primitives.

The selection problem is formalized as follows. We take the Gaussian sphere G and use points on the sphere to represent *movable directions* under a certain set of constraints. Let us use the following notation:

- A_{goal}: A set of points on G which represents desired constraints which need to be achieved after the next sensing is completed.
- A_{curr}: A set of points on G under the current set of constraints. This includes constraints realized by a motion control during the current assembly operation (e.g., a contact-maintaining operation in Fig. 4(b)) as well as geometric constraints achieved so far.
- A_i: A set of points on G which represents constraints to be obtained by observing the ith feature; this information is described in each sensing primitive.

By observing n different features $\{feature_i | i = 1, \ldots, n\}$, the following resultant point set, A_{obsd}, is obtained:

$$A_{obsd} = A_{curr} \bigcap \left(\bigcap_{i=1}^{n} A_i \right).$$

In order that this set of features is sufficient for providing enough constraints, the following condition must be satisfied:

$$A_{goal} \supseteq A_{obsd}. \quad (1)$$

This condition says that the constraints to be obtained after the observation must be stronger than or equal to the desired constraints.

As an example, let us consider two cases (a) and (b) of peg-in-hole operation shown in Fig. 7. Suppose we are localizing the hole by observing the position of its edges, e_1 and e_2. Since the degrees of freedom to be adjusted are limited on a plane perpendicular to the insertion direction, we use the Gaussian circle instead of the Gaussian sphere.

First, as shown in Fig. 7, A_{curr} for case (a) is represented by the full circle, while that for case (b) is represented by two points on the x axis. A_{goal} is the Gaussian circle with no points for both cases. Then, as shown in Fig. 8, A_i for each edge, e_1 or e_2, is represented by two points corresponding the direction

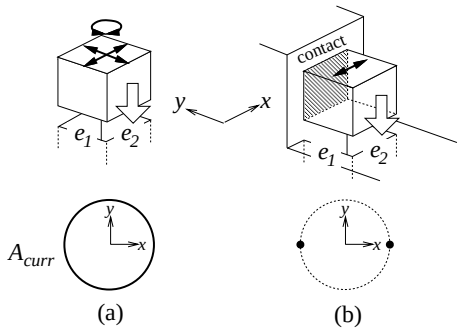


Fig. 7: Two insertion operations and movable directions.

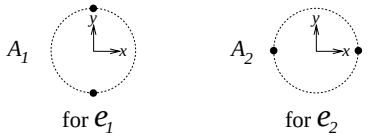


Fig. 8: Constraints obtained by observing edges: A_1 for edge e_1 and A_2 for edge e_2.

of the edge. By applying equation (1) to these data, we can decide that observing only e_1 is sufficient for case (b), and that observing both e_1 and e_2 is necessary for case (a).

Once a set of features to be observed is obtained, a set of feasible sensor positions is enumerated from which all of the selected features are observable. Similar to the prior works [15], the candidates for sensor positions are enumerated by considering visibility of features and the possibility of collision between sensors and other objects, as well as the configuration of sensors. The selection of the final sensing strategy from the candidates is described in the next section.

4 Task-Oriented Evaluation of Visual Sensing Strategies

4.1 Uncertainty-Based Evaluation

We evaluate a sensing strategy based on the accuracy of visual information to be obtained by the sensing strategy. We assume that sensor data consist of position measurements, and that each measurement has the uncertainty, which can be calculated using the uncertainty model of the sensor. The problem of estimating the object position is described as follows. Let $\boldsymbol{a}$ be the parameter vector representing the object position. Suppose that the parameter vector is estimated from a set of 3D measurements $\{\boldsymbol{x}_i | i = 1, \ldots, N\}$, and that the following equation $\boldsymbol{f}_i$, which represents constraints on measurements (shape of the measured feature) and comes from the CAD-based world model, should ideally be satisfied for each measurements:

$$\boldsymbol{f}_i(\boldsymbol{x}_i, \boldsymbol{a}) = \mathbf{o} \tag{2}$$

The initial estimate of the parameters are also obtained from the CAD-based world model. Since this equation is non-linear in general, by applying the Extended Kalman Filter theory [2] to this estimation problem, the covariance matrix S of $\boldsymbol{a}$ is given by

$$\begin{aligned} S^{-1} &= \sum_{i=1}^{N} M_i^T W^{-1} M_i, \\ M_i &= \frac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{a}}, \\ W_i &= \frac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{x}_i} \Lambda_i \frac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{x}_i}^T, \end{aligned} \tag{3}$$

where Λ_i is the covariance matrix representing the uncertainty in the ith measurement $\boldsymbol{x}_i$.

Accuracy of an estimate can be evaluated based on the covariance matrix S. Note that the uncertainty of each element of $\boldsymbol{a}$ has a different effect to the success of the current operation, and that this relative effects of the elements depends on the assembly operation currently being carried out. The next subsection describes a general method to define an appropriate evaluation function for each assembly operation.

4.2 Evaluation Based on Predicted Success Probability

One of the most direct criteria that measures the "goodness" of a sensing strategy is whether the current operation will succeed with the selected sensing strategy. Therefore, we use the *predicted success probability* of the operation as the criterion; the sensing strategy that is selected is the one most likely to result in successful task execution.

To calculate the success probability, we first calculate a region in the space of the error of the position parameter vector such that if the error is inside the region, the current operation succeeds. We call this region a *success region*. Then, the predicted success probability is given by calculating how much portion of the uncertainty distribution of the error, which is predicted from the sensor model, is inside the success region of the current operation (see Fig.9). This success probability is numerically calculated by quantizing the space of the error vector. After calculating the probabilities for all feasible sensing strategies, the one with the highest probability is selected as the final sensing strategy.

4.3 Calculation of Success Region

A success region is a representation of the clearance of the operation, and is formed by a set of simultaneous inequalities in the space of unconstrained degrees of freedom in each operation. In case of the operation shown in Fig. 7(a), for example, the success region is formed in a three-dimensional space composed of two translational and one rotational degrees of freedom on the plane perpendicular to the insertion direction. A method of automatically calculating success regions is described in [10].

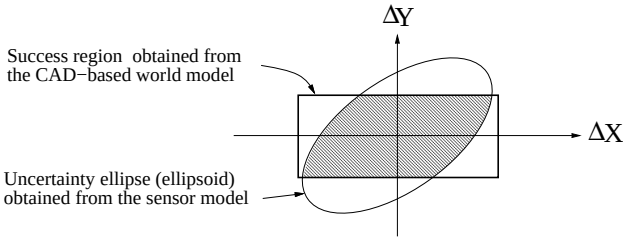


Fig. 9: Calculation of the predicted success probability. This figure shows the case where the position parameter of the object is two-dimensional, (X, Y).

Here, as an example, we derive the success region for the case where the shape of the cross-section is rectangle. Fig. 10 shows a top view of the operation. Edges of the hole are aligned to the X and the Y axes. Let W_X and W_Y be the widths of the peg in the X and Y axes, respectively. Also, let k denote the clearance ratio of the hole. These values come from the CAD model. We need to adjust the position and the orientation of the peg, (X, Y, θ).

Let $(\Delta X, \Delta Y, \Delta\theta)$ denote the error of (X, Y, θ). We can obtain eight inequalities for this operation: two for each vertex of the peg. For the upper-right corner of the peg and the hole, for example, the following two inequalities are obtained:

$$\begin{aligned} \Delta X + \frac{W_X}{2}\cos\Delta\theta - \frac{W_Y}{2}\sin\Delta\theta &\leq \frac{k+1}{2}W_X, \\ \Delta Y + \frac{W_X}{2}\sin\Delta\theta + \frac{W_Y}{2}\cos\Delta\theta &\leq \frac{k+1}{2}W_Y. \end{aligned}$$

We calculated the actual success regions for two sets of geometric values. Fig. 11 shows the resultant success regions. As shown in the figure, the tolerance in ΔX in case (b) is larger than that in case (a), while the tolerance in $\Delta\theta$ is smaller. If the uncertainty distribution of the position parameter is the same to both cases, the resultant success probabilities should differ from each other. Thus, the effect of the uncertainty in the parameter vector to the task execution needs to be evaluated by considering the success region.

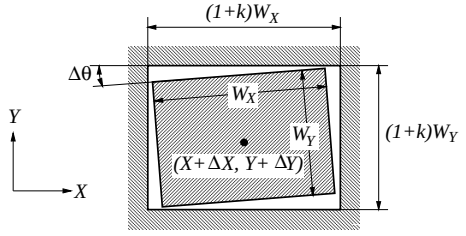
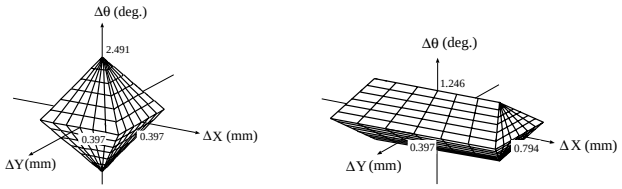


Fig. 10: Rectangular peg-in-hole operation.



(a) $W_X = 19.05$ (mm), $W_Y = 19.05$ (mm), $k = 0.043$.

(b) $W_X = 38.10$ (mm), $W_Y = 19.05$ (mm), $k = 0.043$.

Fig. 11: Example success regions.

5 Implementation of the Method using Line Laser Range Finder

5.1 Laser Range Finder and General Sensing Strategy

The proposed method has been implemented using a Toyota line laser range finder (LRF) [11] as the sensor. The LRF emits slit laser, detects highlighted portion of the object by a TV camera, and obtains a line of 3D measurements (see Fig. 12). The LRF is attached to an overhead platform of the RobotWorld [13]. This platform has four degrees of freedom: three degrees of freedom for translation and one for rotation about the vertical axis.

Every assembly operation that requires visual information is a kind of "peg-in-hole" operation. The location of a peg is measured by observing its side faces; the location of the hole is measured by observing several points (currently, five points) on its edges. Thus, we prepare sensing primitives for the following two geometric features: a straight edge and a planar face. We use a sensing strategy as shown in Fig. 13; data for one assembly operation are collected at several position by moving the LRF in parallel with the insertion direction because the relative displacement on the plane perpendicular to the insertion direction is important for the operation. We also control the position of the range finger so that each measured point is kept within a certain area of the slit laser; the uncertainty of the measurement with the LRF is considered to be almost constant in this area. Thus, the only parameter that specifies the position of the range finder is the angle ϕ between the direction of the slit laser and some axis of the plane perpendicular to the insertion direction (see Fig. 14).

5.2 Assembly Operation with Visual Feedback

The actual vision-guided assembly operation is performed in a "stop and sense" mode. First, a peg is moved by a manipulator to the position just before a hole. Then, the LRF is placed to the planned position, and measures the position of the hole and the peg. If the error in the relative position between the peg and the hole is within the success region (see Section 4.2), the peg is inserted. Otherwise, the peg position is adjusted and the peg is observed again. This final step is repeated until the relative position becomes satisfactory, and then the peg is inserted.

6 Experimental Results

This section describes the experimental results of a peg-in-hole operation. Experimental results for other operations can be found in [10].

To validate the selection of visual sensing strategy based on predicted success probability (see Section 4.2), we compared the predicted success probability, which is predicted from the object models and sensor models, with the actual success ratio, which is statistically obtained through a number of actual trials of the same operation by the actual robot and the sensor.

6.1 Uncertainty Model of Laser Range Finder

The laser range finder we used provides quite accurate measurement: less than $0.1(mm)$ in depth and less than $0.3(mm)$ in the horizontal position. In order to stress the effect of uncertainty to the success probability, we artificially added a relatively large Gaussian noise to the measurement; we added Gaussian of standard deviation $0.12(mm)$ to the depth measurement and that of standard deviation $0.30(mm)$ to the horizontal position measurement; uncertainties of these two measurements are set to be independent of each other. Fig. 15 shows the distribution of 500 measurements of the same point. The resultant uncertainties are reasonably Gaussian with almost the desired standard deviations.

The reason why we used this uncertainty model is that the purpose of this paper is not to construct an uncertainty model of our laser range finder but to demonstrate that our method can generate the optimal sensing strategy if the uncertainty model of the sensor is given.

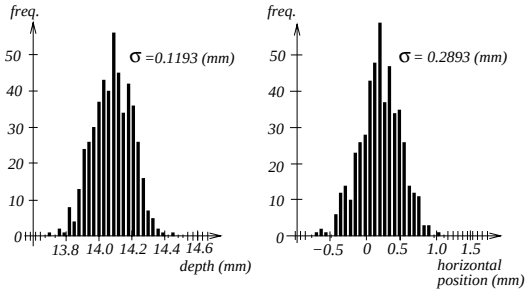


Fig. 15: Distribution of measurements and estimated standard deviation.

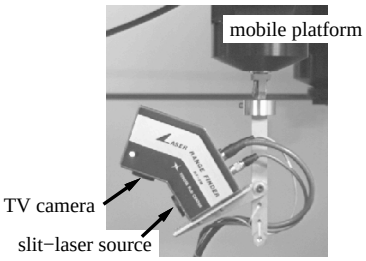


Fig. 12: Laser range finder.

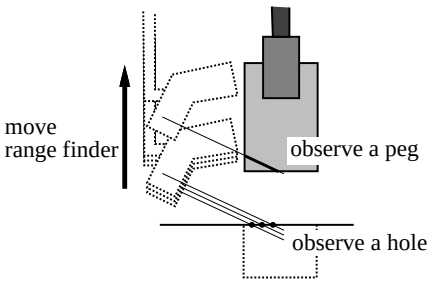


Fig. 13: A strategy for observing a peg and a hole.

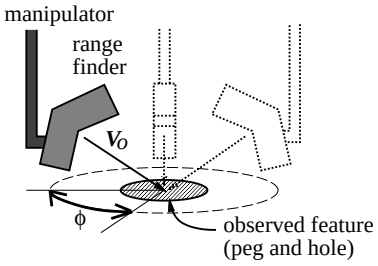


Fig. 14: Candidate positions.

6.2 Face Contact Analysis and Actual Operation

Fig. 16 shows the face contact analysis of the operation. This operation results in two translational degrees of freedom being constrained. The face contacts to be achieved are (f_1-f_1'), (f_2-f_2'), (f_3-f_3') and (f_4-f_4'). The candidates for observed features are f_1, f_2, f_3 and f_4 for the peg, and e_1', e_2', e_3' and e_4' for the hole. To obtain sufficient information for localization, two neighboring faces and edges were observed.

Considering the conditions that five points are completely observed on an edge, and that the LRF does not collide with the robot manipulating the peg, the range of the feasible viewing direction (ϕ) were determined as shown in Fig. 17. The center of the circular trajectory of the LRF was placed on the vertex at the intersection of the two neighboring edges. Fig. 18 shows a successful peg-in-hole operation.

6.3 Comparison of Predicted Success Probability with Actual Success Ratio

We compared the predicted success probability with the actual success ratio in the following two sets of the objects:

Case (a): The cross-section of the peg is a square of $19.05(mm) \times 19.05(mm)$. The clearance ratio of the hole is 0.043. The success region for this case is shown in Fig. 11(a).

Case (b): The cross-section of the peg is a rectangle of $38.1(mm) \times 19.05(mm)$. The clearance ratio of the hole is 0.043. The success region for this case is shown in Fig. 11(b).

In each case, several viewing angles (ϕ in Fig. 17) were selected; for each angle, the same operation was performed 50 times and the numbers of success and of failure were accumulated to calculate the actual success ratio.

Fig. 19 shows the comparison results; the results of actual experiments coincide with the predicted results quite well. We think this result shows the importance of task-oriented evaluation of sensing strategies, i.e., the appropriate sensing strategy should be selected by considering the task to be performed.

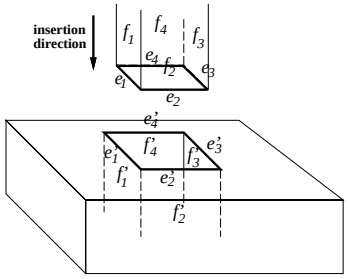


Fig. 16: Face contact analysis of the rectangular peg-in-hole operation. The triplet of DOFs (see Section 2.2) changes from $(3,0,0)$ to $(1,0,2)$.

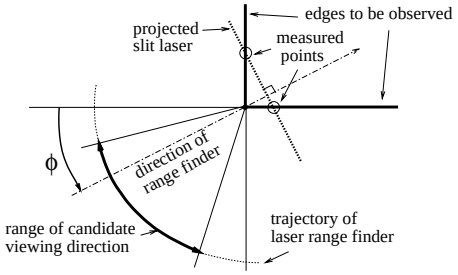


Fig. 17: Top view of candidate viewing directions for observing two edges and faces.

7 Conclusion

We have described a method of systematically generating visual sensing strategies in assembly tasks using the knowledge of the task to be performed. Based on the result of the analysis of the assembly tasks in terms of the transition of face contacts between object surfaces, we can determine which degrees of freedom of the assembled objects should be measured. A set of visual features to be observed are then selected by which the necessary degrees of freedom are measured. Finally, among feasible visual sensing strategies, the one with the highest predicted success probability is selected as the final sensing strategy.

The proposed method has been implemented using

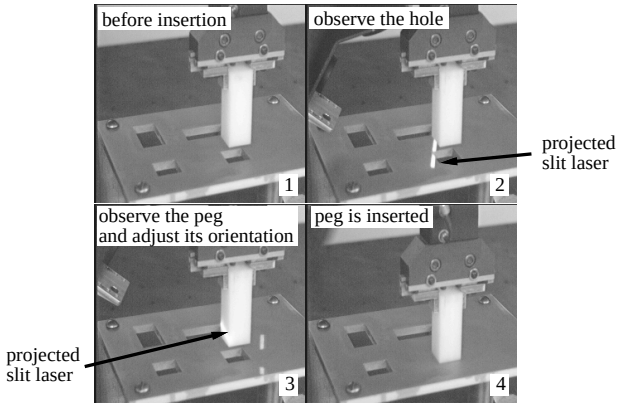


Fig. 18: A successful peg-in-hole operation.

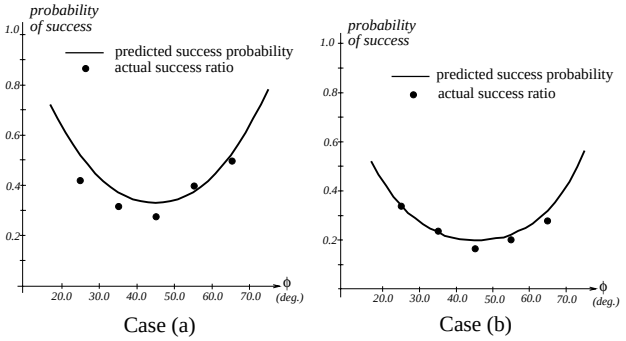


Fig. 19: Comparison of the predicted success probability with the actual success ratio.

a line laser range finder as the sensor. The experimental results show the feasibility of the method, and point out the importance of task-oriented evaluation of visual sensing strategies.

In [10], we have extended the proposed method to more general assembly tasks where an object can be composed of planar or cylindrical surfaces, and where, in addition to three translational degrees of freedom, one rotational degree of freedom is allowed. Since most of assembly operations in practical situations are realized by these four degrees of freedom, this extension would be useful for practical use. We have successfully applied this extended method to several assembly operations such as a gear-mating operation [10].

Acknowledgement

The authors would like to thank Dr. Sing Bing Kang for helpful comments on the draft of this paper.

This research was conducted while the first author was with Carnegie Mellon University. His stay was supported by The Telecommunications Advancement Foundation, Tokyo, Japan.

References

[1] Y. Aloimonos, editor. *Active Perception.* Lawrence Erlbaum Associates, Inc., New Jersey, 1993.

[2] N. Ayache and O.D. Faugeras. Maintaining Representations of the Environment of a Mobile Robot. *IEEE Trans. on Robotics and Automat.*, RA-5(6):804–819, 1989.

[3] R. Bajcsy. Active Perception. *Proceedings of IEEE*, 76(8), 1988.

[4] I. D. Horswill. *Specialization for Perceptual Processes.* PhD thesis, Massachusetts Institute of Technology, 1993.

[5] S.A. Hutchinson and A.C. Kak. Planning Sensing Strategies in a Robot Work Cell with Multi-Sensor Capabilities. *IEEE Trans. on Robotics and Automat.*, RA-5(6):765–783, 1989.

[6] K. Ikeuchi and M. Hebert. Task Oriented Vision. In *Proc. of Int. Conf. on Intelligent Robots and Systems*, pages 2187–2194, Raleigh, NC, 1992.

[7] K. Ikeuchi and T. Suehiro. Toward an Assembly Plan from Observation Part I: Task Recognition With Polyhedral Objects. *IEEE Trans. on Robotics and Automat.*, 10(3):368–385, 1994.

[8] K. Kemmotsu and T. Kanade. Sensor Placement Design for Object Pose Determination with Three Light-Stripe Rnage Finders. In *Proceedings of 1994 IEEE Int. Conf. on Robotics and Automat.*, pages 1357–1364, San Diego, CA, 1994.

[9] Y. Kuniyoshi and H. Inoue. Qualitative Understanding of Ongoing Human Action Sequences. In *Proceedings of the 13th Int. J. Conf. on Artificial Intelligence*, pages 1600–1609, Chambéry, France, 1993.

[10] J. Miura and K. Ikeuchi. Task-Oriented Generation of Visual Sensing Strategies in Assembly Tasks. Technical Report CMU-CS-95-116, School of Computer Science, Carnegie Mellon University, February 1995.

[11] O. Ozeki, K. Higuchi, and S. Yamamoto. Automated Dimension Inspection System for Automotive Plastic Parts with a Laser Probe. In *Proc. of ROBOTS 12 and Vision '88 Conference*, pages 5–51–5–60, Detroit, MI, 1988.

[12] R.D. Rimey. Control of Selective Perception Using Bayes Nets and Decision Theory. Technical Report 468, Computer Science Department, The University of Rochester, December 1993.

[13] V. Scheinman. A Multiple Robot Vision Guided Assembly System. In R. Bolles and B. Roth, editors, *Robotics Research 4*. The MIT Press, 1987.

[14] T. Suehiro and K. Takase. Representation and Control of Motion in Contact and Its Application to Assembly Tasks. In H. Miura and S. Arimoto, editors, *Robotics Research 5*, pages 367–374. The MIT Press, 1990.

[15] K.A. Tarabanis, P.K. Allen, and R.Y. Tsai. A Survey of Sensor Planning in Computer Vision. *IEEE Trans. on Robotics and Automat.*, 11(1):86–103, 1995.