

A Survey of Sensor Planning in Computer Vision

Konstantinos A. Tarabanis, Peter K. Allen, *Member, IEEE*, Roger Y. Tsai

Abstract—A survey of research in the area of vision sensor planning is presented. The problem can be summarized as follows: Given information about the environment (e.g., the object under observation, the available sensors) as well as information about the task that the vision system is to accomplish (i.e., detection of certain object features, object recognition, scene reconstruction, object manipulation), develop strategies to automatically determine sensor parameter values that achieve this task with a certain degree of satisfaction. With such strategies, sensor parameters values can be selected and can be purposefully changed in order to effectively perform the task at hand. Sensory systems are then able to operate more flexibly, autonomously, and reliably. This problem has recently become an active area of study with a number of researchers addressing various aspects of the problem. The focus here is on vision sensor planning for the task of robustly detecting object features. For this task, camera and illumination parameters such as position, orientation, and optical settings are determined so that object features are, for example, visible, in focus, within the sensor field of view, magnified as required, and imaged with sufficient contrast. References to, and a brief description of, representative sensing strategies for the tasks of object recognition and scene reconstruction are also presented. For these tasks, sensor configurations are sought that will prove most useful when trying to identify an object or reconstruct a scene.

I. SENSOR PLANNING

A. Motivation

SENSOR planning in computer vision is an emerging research area that tries to understand and quantify the relationship between objects to be viewed and the sensors observing them in a model-based, task directed way (see Fig. 1). The importance of the viewpoint in this object-viewer relationship is evident. It is the viewpoint, for the most part, that creates the quality of the resulting image [43], determining the feasibility of the vision task and facilitating its execution. Previous work in computer vision, however, has generally placed an emphasis on the object being observed [26] assuming that the viewpoint is given, suitable for the task, and not controlled. The issues related to determining viewpoints that will be most suitable for the vision task at hand have received considerably less attention. This latter area includes questions such as “What should the observer pose be?” or “What values should other observer attributes have?”

Even in currently employed vision systems that observe known objects in known poses (e.g., visual inspection, surveillance, or monitoring systems), appropriate sensor parameter

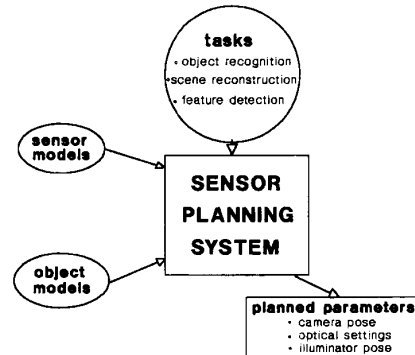


Fig. 1. Sensor planning for computer vision.

values are determined by often laborious and time-consuming techniques. Generally, a trial-and-error approach involving human interaction is taken. Sensor locations and settings are chosen and then tested in order to verify whether they meet the requirements of the task at hand. The resulting parameter values are valid for only a specific setup and can potentially become unsatisfactory when errors (e.g., robot inaccuracy) alter the environment. Such procedures constitute a major bottleneck in system installation and result in a design cost which often exceeds that of the equipment. As a result, these applications are costly and have limited intelligence and flexibility.

Sensing strategies are also required for vision systems that perform object recognition since a single sensor configuration may not always result in a sufficiently informative view. For instance, in the cases of occlusion and near-symmetrical or similar objects, more than one interpretation may be consistent with the sensory data. As a result, additional sensor configurations are needed from which observed details of the scene can help recognize the unknown objects by disambiguating among multiple interpretations.

In addition to the sensor parameters themselves, another set of key parameters in reconfigurable vision systems that can be controlled, and thus may need to be planned, are those associated with the illumination of the scene. If the interaction between lighting and the object surface to be imaged is carefully considered, the image quality can be substantially improved, leading to more informative images and making later processing of the image much easier. For instance, by planning the illumination in a visual inspection application, features of interest can be made to appear more prominent, noise due to extraneous features can be reduced or eliminated, and the accuracy with which an object is measured, can be increased. While techniques that vary the illumination

Manuscript received April 21, 1992; revised June 28, 1993.

K. Tarabanis and R. Y. Tsai are with the IBM T. J. Watson Research Center, Yorktown Heights, NY 10598.

P. K. Allen is with the Department of Computer Science, Columbia University, New York, NY 10027.

IEEE Log Number 9400866.

(e.g., illuminator location, polarization of illumination, etc.) have been employed in several vision systems, illumination planning has been overlooked to a great extent. Since image acquisition is much less costly than the subsequent image analysis, it seems advantageous to dedicate computational effort to determine appropriate values for both the viewpoint and the illumination parameters.

Sensor planning is pertinent to a number of areas of robotics and computer vision that have been studied extensively in the past. For instance, the general problem of task planning in robotics and its component areas of motion planning, grasp planning, and assembly planning can be viewed as different facets of the sensor planning problem. In addition, recent research has underlined the importance of an area very much related to sensor planning, namely, that of *active sensing*. In active sensing sensor parameters are controlled in response to the requirements of the task [7], [70]. It has been shown in [4] that active sensing can take ill-posed vision problems and render them well posed through constraints introduced by the dynamic nature of the sensor. Also, by actively reconfiguring a sensor system its effective scope is increased, and as a result, a range of sensing situations can be accommodated. For example, the field of view of a mobile observer is far more extensive than that of a stationary one, while required object detail can be resolved by either reducing the viewer-object distance or, when possible, by modifying the associated lens settings. In order to purposefully alter sensor configurations, active sensing requires a sensor planning component that determines appropriate sensor parameter values.

The goal of sensor planning is to automatically, rather than manually, generate appropriate sensor configurations based on any a priori known information that is often available. For instance, the required geometric and physical information of objects can be extracted from CAD/CAM models which are often available in today's manufacturing environment. Camera and illumination models approximating their physical and geometric properties can also provide the planning system with the required sensor and illumination characteristics. The planning algorithms can use this model information and augment it with knowledge regarding the functions the system is to perform. In this way, the sensor-based system will be able to reason about its own configuration in order to achieve the task at hand.

B. Applications

Several researchers have recognized the importance of such a planning component for a new generation of *automated visual inspection systems* [56], [78], [84]. In [78] a system is defined to automatically generate dimensional measurement applications from CAD models of parts. In [56] a concept-based inspection system is outlined in which a model-based inspection system would be augmented with a generic set of principles in order to determine appropriate and flexible behavior in new situations. The need for machine vision development tools that will assist machine vision designers and allow less specialized labor to be employed is underlined in [84]. The tool proposed would automatically generate the

configuration of a machine vision system by selecting sensors, illumination, optics, and image processing algorithms.

Vision sensor planning is also useful for *robot-controlled vision systems* in which cameras and light sources are mounted on robot manipulators (e.g., [11], [83]). In order for these systems to perform their task reliably (e.g., vision-guided remote assembly/disassembly or manipulation of objects in space or in hazardous environments), selection of the proper vision sensor parameter values is critical. The developed sensor planning techniques could be used to automatically position and orient the cameras and light sources as well as to control the camera optics (e.g., controlling the zoom, focus, and aperture settings of programmable zoom lenses).

Sensor planning techniques are also applicable to areas such as the *automatic synthesis of vision programs* from task specifications and model information [29], [40], [50]. For example, a vision program to inspect an object can be automatically generated, to a certain extent, based on sensor planning techniques that determine appropriate camera and illuminator poses, optical settings, and image processing algorithms.

It is important to note that planning techniques developed for vision sensors can prove useful in other areas of automation, such as the automated machining and dimensional inspection of mechanical parts [18], [67], [71]. For instance, planning for accessibility of a surface in order to machine or probe it by tactile means is equivalent to planning for visibility of this surface assuming orthographic projection. In these latter domains of automation, parameters such as workpiece and probe orientation, machine selection and cutter, or probe-type selection are determined by the planning systems. Vision sensor planning techniques can also be used in the area of computer graphics for the automatic generation of viewing specifications that result in scene renderings that achieve a certain intent [64].

By employing a sensor planning component in these application domains:

- 1) The development cycle and thus the cost of a sensor system is reduced since sensor configuration can be done automatically;
- 2) Sensor parameter values can be found that are robust (i.e., satisfying the task requirements at hand even in the presence of uncertainty);
- 3) Sensor parameter values can be determined that optimize the sensor output with respect to a given task criterion, (e.g., camera and illuminator poses that enhance the contrast between task relevant features and the background);
- 4) A sensor-based system can adaptively reconfigure itself to accommodate variations in the workplace.

C. Scope of this Survey

In an attempt to limit the scope of this survey, we have focused on sensor planning for vision sensors (e.g., cameras, range finders, illuminators, etc.). Furthermore, within the area of vision sensing, we have tried to limit our study to systems using higher-level model-based approaches as opposed to pixel-level active vision. We are concerned with finding a

generalized viewpoint that includes sensor parameters other than camera placement (e.g., focus, aperture, illumination placement). Active vision systems that vary their parameters by an assumed a priori schedule—randomly or heuristically rather than in a knowledge-based manner with an identifiable selection criterion—is outside the scope of this survey.

Following the approach of Maver and Bajcsy [54], we can classify the approaches to the sensor planning problem by the vision task to be achieved or similarly by the amount of a priori known information about the scene. This allows us to identify three distinct areas of research: object feature detection, model-based object recognition and localization, and scene reconstruction.

The first area, object feature detection, seeks to automatically determine vision sensor parameter values for which particular features of a known object in a known pose satisfy particular constraints when imaged [6], [8], [23], [24], [40], [45], [57]–[60], [62], [72], [74], [75], [88]. For example, the features are required to appear in the image as being visible, in-focus, and magnified to a given specification. These planning techniques draw on the considerable amount of a priori knowledge of the environment, the sensors, and the task requirements. Because the identities and poses of the viewed objects are known, the sensor parameters are usually preplanned off-line and then used on-line when the object is actually observed. Our own interest is in this first area, and this survey reflects this interest as well as our belief that it is an important emerging research area. In the remainder of this section references to, and a brief description of, the other two areas is presented.

The second area is concerned with developing sensing strategies for the tasks of model-based object recognition and localization [16], [17], [30], [37], [40], [44], [51]. Here, sensing operations are chosen that will prove most useful when trying to identify an object or determine its pose. In this work, the a priori known information about the world in the form of models of possible objects, sensor models, and information acquired to this point are compiled into recognition/localization strategies (see Fig. 1). Most approaches in this area follow a common theme. Namely, a *search* is performed in the space of object identities/poses employing the *hypothesize-and-verify* paradigm:

- 1) Hypotheses are formed regarding the object identities and poses;
- 2) These hypotheses are assessed according to certain metrics;
- 3) New sensing configurations are proposed based on a given criterion until a stopping condition is met.

Since after the first step in the above approach the identities and poses of the objects in the scene have been hypothesized, the sensor planning techniques of the first area are also applicable in this second area.

In order to limit the search of sensor parameter space in this hypothesize-and-verify paradigm, a discrete approximation of this space is commonly employed. For example in [30], sensor configurations are only chosen from a known finite set. In [15], [37] a discrete approximation of viewing space based on the

aspect graph representation of the object is used. The individual approaches following this general methodology basically differ in the hypothesis evaluation metrics, the criterion for the next sensing operation or the stopping condition.

1) The quality of the match between image and scene features is evaluated by Cameron in [16] using a cross-correlation measure. In the work by Hutchinson and Kak [37], confidence in matches between model and scene features are based on the similarity of the area and the 3D shape attributes of the surfaces.

2) In forming the hypotheses, Magee [51] takes a theorem-proving approach to establish a symbolic correspondence between scene and image points, while in Kim's work, the correspondence between object images from one view to the next is determined by employing correspondence techniques used in stereo. In the work by Grimson [30], a set of possible interpretations of the original sensory data is assumed to have been determined by some other means.

3) The criterion for the next sensing operation is based on optimally verifying the current hypotheses according to a metric. A very similar metric is employed in most of the approaches. For example, in [37] this metric involves reducing the maximum uncertainty of the hypothesis set, while in [16], [30], [44], [51] it entails viewing particular disambiguating features.

Similar to the aforementioned sensor planning work for object recognition, there has also been work [12], [13], [28], [33], [38], [40] in automatically determining strategies that guide the search for matching image and scene features in Step 2 of the hypothesize-and-verify paradigm. In this work, however, sensing configurations are not planned, but rather, for whatever the sensing configuration happens to be, a strategy is given to match image and scene features based on this initial set of sensory data without often acquiring additional sensory data.

The third area addresses the problem of determining sensing strategies for the purpose of scene reconstruction [1], [14], [19], [21], [31], [53], [66], [81], [85]. In this case, a model of the scene is incrementally built by successively sensing the unknown world from effective sensor configurations using the information acquired about the world to this point. At each step of the algorithm, new sensor configurations are chosen based on a particular criterion (e.g., ability to explore the largest area of unknown space). The sensory information acquired at each step is then integrated into a partially constructed model of the scene and new sensor configurations are generated until the entire scene has been explored. While there is no a priori known scene information that can be used in this problem, the iterative sensing is guided by the information acquired to each stage. With respect to the planning component of this work, the various approaches differ in the criterion with which a new sensor configuration is chosen. The approaches also differ in the way the multiple views are integrated into a scene model. The work described in [3], [34], [52] is related in that it addresses the problem of reconstructing a scene from multiple views, however, the viewpoints are predefined (i.e., assumed to be given) and are not planned in any manner.

II. SENSOR PLANNING FOR FEATURE DETECTABILITY

This review now focuses on vision sensor planning where the sensing task is image acquisition for robust feature detection. The techniques that are discussed are mostly applicable to vision systems that observe known objects in known poses; for example, in visual inspection systems, surveillance systems, or monitoring systems.

In the following sections, common themes are initially underlined in the various approaches in order to allow grouping and high-level contrast of methodologies. An in-depth analysis of the problem and a critical review of each approach follows.

A. Overview of Research in Sensor Planning for the Task of Feature Detectability

This problem is fairly new in the area of computer vision but has recently received considerable interest. Several vision planning systems are being developed that draw on the a priori known information regarding the observed object and the employed sensors in order to automatically determine vision sensor parameter values that satisfy certain feature detectability constraints. The developed techniques differ in their general approach to determining sensor parameter values.

Several systems [59], [60], [62], [63], [87], [88] take a *generate-and-test* approach, in which sensor configurations are generated and then evaluated with respect to the task constraints. In order to limit the number of sensor configurations that are considered, the domain of sensor configurations is discretized. The viewing space discretization in these systems is performed by surrounding the object with a tessellated viewing sphere and limiting the possible sensor locations to either groups or individual tessels on the sphere. Determination of the sensor parameter values is formulated as a search over this discretized domain of sensor configurations guided by task-related heuristics.

The sensor planning methods described in [5], [6], [22]–[24], [72], [74], [75] take a *synthesis* approach. In this approach, the task requirements are characterized analytically and the sensor parameter values that satisfy the task constraints are directly determined from these analytical relationships.

There has also been work related to sensor planning in the area of *sensor simulation* systems [36], [41], [58]. In such systems, a scene is visualized given the description of the objects, sensors, and light sources. These systems provide a framework for planning of sensor configurations. For instance, by taking a generate-and-test approach, satisfactory sensor configurations can be found by creating a simulated view of the scene and evaluating the task constraints in the simulated image.

Finally, there has been related work that follows the *expert systems* paradigm [8], [45], [57]. In such systems, expert knowledge of viewing and illumination techniques is incorporated into an expert system rule base. After acquiring information regarding the particular object to be observed, the expert system provides advice regarding the appropriate sensor configuration.

B. Defining the Problem

In this section we describe the problem of sensor planning for feature detectability in more detail. In Section I the problem was summarized as automatically determining camera and illumination parameters (e.g., position, orientation, settings) that satisfy certain feature detectability task requirements (e.g. visibility, focus, contrast) by drawing on the knowledge of the environment, the sensors, and the task requirements themselves. From this description, it can be seen that the problem is *model based* and *task driven*. The models include information of the objects in the environment and the employed sensors. This a priori knowledge, together with knowledge of the task requirements that need to be satisfied, are incorporated into the planning algorithms.

In the following sections, we discuss each of the above mentioned components of the problem. These are

- 1) The sensor parameters,
- 2) The sensor and object models, and
- 3) The feature detectability constraints.

1) *Sensor Parameters*: The parameters that need to be determined for vision sensors are basically associated with the point from which the scene is observed (i.e., viewpoint) and the point from which the scene is illuminated. Customarily, a viewpoint is considered to be the viewer location alone. However, it is useful to define the concept of a viewpoint in a broader sense that includes not only the viewer orientation, but also the optical settings associated with the viewpoint. These settings are also observer attributes affecting the resulting image of the scene.

This broader definition, then, identifies two types of sensor parameters: *geometric* and *optical*. The geometric parameters are independent of any of the lens optics. Parameters of this type the following:

- The three positional degrees of freedom of the sensor (x, y, z)—this parameter is the position vector of a point associated with the sensor. An example of a convenient point for this purpose is a point on the C-mount surface of the lens since the relationship between the C-mount and the sensor does not vary. The characteristic points of the lens itself (e.g., the principal points, the nodal points, the pupils, etc.) can also be used. However, since these points are not rigidly fixed to the sensor, the relative movement of these points as the optical settings of the lens vary must be accounted for.
- The three orientational degrees of freedom of the sensor—this parameter can be specified by the pan, tilt, and swing angles of a vector rigidly attached to the sensor. An example of a convenient vector of this type is a unit vector along the viewing direction (i.e., optical axis). While the characteristic points of the lens change as the optical settings of the lens vary, the viewing direction remains constant.

The optical parameters, on the other hand, depend on the lens settings. Parameters of this type include the following:

- The back principal point¹ to image plane distance, d —this parameter d is varied in the several focus control mechanisms that are employed in lenses in order to achieve a focused image. Also, d is the image distance that is employed in the Gaussian lens formula governing focusing of a lens.
- The entrance pupil diameter², a of the lens—the entrance pupil diameter depends on the size of the aperture (controlled by the aperture setting of the lens) and the lens optics that image the aperture. Among other things, the aperture, and hence the entrance pupil, affect the focus and the brightness of the image.
- The focal length f of the lens—the focal length is a fundamental property of a lens and it provides a measure of its refractive ability. In cases where it can be varied (e.g., zoom lenses) the focal length is a sensor parameter; otherwise, it is a sensor constant.

Other optical camera parameters that can also be controlled in some cases [65] include the following:

- 1) The exposure time of the sensor, which determines how long the sensor will be exposed to light from the scene;
- 2) The gain of the video signal by the camera amplifier;
- 3) Parameters adjusting the spectral responsivity of the imaging system (e.g., the spectral transmittance of a filter).

The illumination parameters are also of two types: *geometric* and *radiometric*. The geometric parameters are independent of any of the physical characteristics of the illumination. Parameters of this type include the following:

- The three positional degrees of freedom of the illuminator—this parameter is similarly the position vector of a point associated with the illuminator. For example, in the case of a point light source, this parameter can obviously be the location of the point light source itself.
- The three orientational degrees of freedom of the illuminator in the case of directional illumination—this parameter can be specified by the pan, tilt, and swing angles of a vector rigidly fixed to the illuminator. An example of a convenient vector of this type is the unit vector along the axis of symmetry of the illumination beam.
- The geometric characteristics of the illumination beam—for example, the illumination beam may have the shape of a conical solid angle with a constant or variable apex angle.

The radiometric parameters depend on the physical and optical characteristics of the illumination. Parameters of this type include the following:

- The radiant intensity—this parameter describes the power output of the light source.

¹The principal points of an optical system are two conjugate points on the optical axis such that planes drawn perpendicular to the optical axis through these points are planes of unit magnification. For the case of a thin lens camera model, the principal points coincide at a single point.

²The entrance pupil of a lens is the image of its aperture with respect to all elements preceding it in the lens system.

- The spatial distribution of intensity—for example, the above parameter is the most important characteristic of the illumination in the case of structured lighting.
- The spectral distribution of intensity—this parameter can be varied by employing spectral filters with a certain spectral transmittance.
- Parameters describing the polarization of the illumination—the polarization state of the illumination can be controlled by employing polarizers.

Thus, in the most general case, planning of camera, lens, and illumination parameters is to be done in a high-dimensional imaging space [65]. As will be seen in Section II-D, however, only a subset of the above set of sensor and illumination parameters is typically considered in any of the existing sensor planning systems. For example, in many systems the camera is assumed to be located at a fixed distance away from the object or the optical axis is assumed to pass through a fixed point on the object.

2) *Sensor and Object Models*: The sensor models, camera, lens, and illuminator, embody information that characterizes the operation of the sensor. Such information includes the following:

- The sensor and illumination parameters and constants—for example, the field-of-view angle of the sensor, the sensor pixel size, the lens focal length, the sensor noise, the spatial extent of illumination, its intensity distribution within this extent, etc.
- The governing relationships between the sensor parameters and the sensor constants—for example, the perspective projection imaging model, the lens formula for focusing, etc.

Object models that are commonly used (e.g., CAD/CAM models) contain geometric and topological information, which may be sufficient for some aspects of the sensor planning problem. However, the photometric properties of the object, such as color and reflectivity of its surface, are needed for illumination planning in order to accurately model the interaction between light and the object surface and then approximate the resulting image. In this case, currently used object models need to be augmented to include such information.

3) *Feature Detectability Constraints*: The feature detectability constraints discussed in this section are fairly generic to most vision tasks. An initial set of such constraints for sensor planning was introduced by Cowan *et al.* in [22], [24]. Similar to the way sensor parameters have been classified in Section II-B.1, the feature detectability constraints can also be collected into two groups depending on whether the illumination source plays a role in the constraint or not.

The first group consists of the purely *sensor constraints* for which illumination is not a factor. Such constraints include the following:

- Visibility—for a feature to be detectable by the sensor, it must first be visible to the sensor. This means that all lines of sight from the camera to each point of the feature are not obstructed (i.e., occluded) by anything in the environment.
- Field of view—while the visibility constraint requires that the rays of light from the feature reach the sensor, the

field-of-view constraint requires that these rays must then image the features onto the active area of the sensor. If this does not happen, the image of the features will either be clipped by, or will lie outside, the active sensor area and thus will not be observable.

- Focus—the goal of this constraint is to guarantee that the features are in focus. While there is only one object distance for which points are in perfect focus, there is a tolerance in position for which a feature is still considered acceptably focused based on the resolution of the image sensor. This tolerance is referred to as depth of field. Thus the focus constraint requires that the features of interest lie within the depth of field of the lens.
- Magnification or pixel resolution—this constraint governs the size of the feature images. Often, a linear feature is required to appear in the image with a certain length (e.g., 5 pixels). This constraint is referred to as the magnification constraint. A related constraint is that of resolution for which any two points of the feature must be imaged by distinct pixels on the sensor plane.
- Perspective distortion—in many applications (e.g., graphics) the images of a scene are considered undesirable when the scene is severely distorted under the projection. In such applications a constraint that minimizes the perspective distortion of features is appropriate.

The second group of feature detectability constraints depends on the illumination. These constraints are referred to as *illumination* or *radiometric constraints* and include the following:

- Illuminability—for a feature to be detectable by the sensor, it is not sufficient that it be visible to the sensor alone. It is also necessary that the feature be visible to at least some point of the light source as well. If this is not the case, the feature will not be illuminated and as a result, it cannot be detected. If a feature point is visible to only some but not all points of the light source, it will be illuminated by these points and will either be back-facing or in shadow for the remaining part of the light source. A related constraint is shadow avoidance. For example, such a constraint is important in photometric stereo (see Section II-C.1) and—in cases where shadows cast over the feature—may be mistakenly interpreted as the feature itself.
- Dynamic range of the sensor—an object point will also not be detectable in the image when the image irradiance from that object point is outside the dynamic range of the sensor. For instance, the image irradiance from an object point may be too weak to sensitize the corresponding photoreceptor cells and as a result will appear black in the image. On the other hand, the image irradiance from an object point may be too high and thus may saturate the photoreceptor area. In both cases, the object point will be considered undetectable.
- Contrast—detectability of an edge feature in a scene is often determined by the edge contrast in the image, that is, the disparity in image intensity values at points in the neighborhood of the image of the edge. For example, an

edge feature may not be detected in the image when the contrast between its adjacent faces is not sufficient for the edge operator at hand.

The above constraints can be further classified as *geometric* and *optical* constraints. Geometric constraints depend only on the geometric sensor and illumination parameters. Examples of such constraints are visibility and illuminability, where these depend only on the geometric parameter of position. Optical constraints depend on both the geometric and the optical parameters. Examples of such sensor constraints include resolution, focus, field of view, dynamic range, and contrast.

It is important to note that when the sensor and radiometric constraints are formulated in terms of the sensor and illumination parameters, the sensor constraints involve only sensor parameters, while in general the radiometric constraints include both sensor and illumination parameters. For example, the sensor parameter of aperture is involved in both the sensor constraint of focus and the radiometric constraint regarding the dynamic range of the sensor. This coupling of sensor and illumination parameters can also be seen in the case in which the reflectance model of the object surface includes a specular component or when lens collection³ is taken into account. In both cases, the sensor location is coupled into the image irradiance equation. This coupling results from the bidirectional reflectance distribution function in the former case, and in the latter from the fourth power law for the cosine of the off-axis angle [35]. There are cases in which such couplings can be neglected. For example, in the case of a diffuse object, the scene radiance is the same irrespective of the camera position. Similarly, if the field of view is narrow, the dependence of the image irradiance on the off-axis angle is negligible. In general, however, the sensor and illumination planning problems are *not separable*.

C. Review of Work Taking the Generate-and-Test Approach

1) *The HEAVEN System*: HEAVEN incorporates work by Sakane *et al.* [59], [61]–[63] in sensor and illumination planning and is the precursor to the VIO system that is discussed in the next section. HEAVEN uses a spherical representation that is both simple and efficient to model sensor configurations. A geodesic dome is created around the object, tessellating the sphere with an icosahedron that is further subdivided in a hierarchical fashion by recursively splitting each triangular face into 4 new faces. This process can be implemented at whatever level of detail is needed for the task at hand at the cost of increased processing time. The viewing sphere is centered on the object and its radius is equal to an a priori chosen distance from the vision sensor to the target object. This serves as a simple resolution constraint (although different features on the object will be imaged at different resolutions due to the variation in their distance and orientation with respect to the sensor). See Figs. 2 and 3.

Given this spherical representation, a method is needed to determine the visibility of a target point from each facet of the spherical dome. By projecting a ray from the target point of the object (located at the sphere's center) to the center of

³Lens collection [35] relates scene radiance to image irradiance.

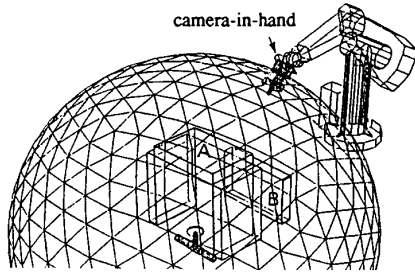


Fig. 2. A robot environment and a geodesic dome generated around a target object in the HEAVEN system (from [62]).

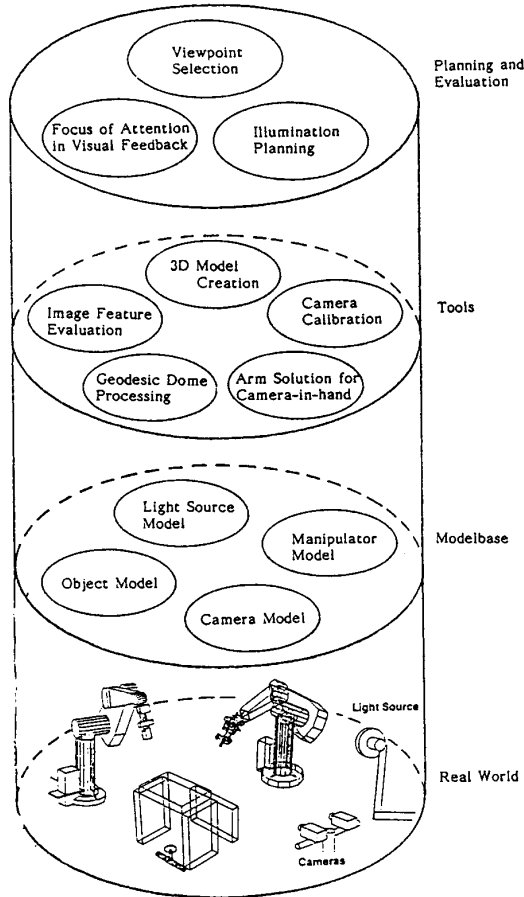


Fig. 3. The HEAVEN system (from [62]).

the facet being processed, all intersections of object faces with the ray can be computed. Depth buffering can then be used to determine if the facet is occluded or not. This approach can be quite costly, since its complexity is proportional to the product of the number of object faces and the number of facets. To alleviate this, bounding box representations of the objects to be tested are intersected with the pyramid volume consisting of the facet and the center of the dome. If these tests show an intersection, then the test is applied to the bounding boxes of the next level of the recursive facet tessellation. For linear

features, a method has been devised to take the convex hull of the two regions on the geodesic dome that correspond to the visibility of the endpoints as the occlusion-free region.

A distance method criterion has been developed to rank each facet within an occlusion-free region. The idea is that facets close to a border with an occluded region are possibly poor choices for sensor location, as noise or small errors in calibration may cause the occlusion-free facet to be occluded. The distance used in this criterion is the negated inner product of the ray from the center of the chosen facet to the center of the nearest occluded facet. This simple test allows each facet to be sorted by its distance transformation. To compute this distance, an iterative neighbor growing scheme on the tessellated sphere is used, propagating from the chosen facet via each of its three adjacent neighbors, until all paths to the occluded region are computed.

The final step in determining placement for a sensor (in this case, a camera mounted on a robot arm or "eye-in-hand" configuration) is taken as the intersection of the candidate facets (sorted by distance), any possible user-chosen locations, and the workspace of the manipulator encoded as regions on the geodesic dome also.

The occlusion-free regions for camera placement are also regions where a point light source can be placed without casting shadows on the target. This is made use of in HEAVEN when light source position planning is performed for a photometric stereo system. In order to address the case where the camera occludes the illuminator or vice-versa, the camera and illuminators are also considered as part of the scene. Since shadows caused by surrounding objects need to be avoided in a photometric stereo system, HEAVEN first computes these shadow-free facets and then ranks them based on two criteria:

- 1) Reliability for recovering the surface normal vectors, and
- 2) Detectability of the surface by the camera and the illuminators.

For the first criterion, the photometric stereo problem is posed in a least squares setting and a singular value decomposition analysis is applied. The reliability criterion of the illuminator configuration C_1 is formulated as the inverse of the condition number of the matrix related to the light-source positions:

$$C_1 = 1/\text{cond}(M_s) \quad (1)$$

where M_s is the matrix that represents the unit vectors directed from the object surface to the light sources. Such a criterion favors the most stable light position in terms of accuracy in estimating the surface normal vectors.

On the other hand, the detectability criterion C_2 is defined as the range of surface normal vectors that can be detected by the photometric stereo setup. A viewing illuminating direction is considered to detect a surface normal vector if the two make an angle less than 90° . (Similar detectability criteria are employed in the VANTAGE system that is discussed in Section II-F.1.) In addition to the three light source locations, the criterion C_2 takes into account the camera position which is also controllable. C_2 is expressed as the area of intersection of four hemispheres on the spherical viewing/illuminating

surface. Each hemisphere corresponds to the range of surface normal vectors that can be detected if a camera or illuminator is positioned at the center of the hemisphere.

A composite criterion is formulated that combines C_1 and C_2 . Initially, each criterion is required to satisfy a certain minimum value:

$$C_1 > TH_1; \quad C_2 > TH_2. \quad (2)$$

The composite criterion C is formulated as

$$C = tC_1 + (1 - t)C_2 \quad (3)$$

where t is the weight that controls the relative importance of the criteria. The composite criterion is evaluated for all combinations of occlusion-free positions of the camera and shadow-free position triplets of the illuminator. The setup with the highest value of the composite criterion is chosen by the system. HEAVEN also incorporates planning of the window layout for a visual feedback control task by again formulating the problem in a singular value decomposition framework.

2) *The VIO System:* The vision illumination object (VIO) system, developed by R. Niepold and S. Sakane [60], plans the setup of both a camera and a point light-source. Given information regarding the environment (i.e., the objects, sensors, and illuminators) and the task at hand, their system determines

- 1) The position of the camera,
- 2) The position of a point light-source,
- 3) A list of most suitable edge features to monitor on the target object, and
- 4) A quantitative assessment of each sensor setup.

The object features to be observed are not fixed in VIO, but may vary as different features are more suitable in different object-sensor-illuminator configurations. VIO also assigns to each configuration a number describing the degree of its suitability to the task at hand based on some chosen criteria.

The camera and illuminator positions are again taken to lie on the surface of a sphere of a chosen radius and centered at an object reference point. In addition, the camera optical axis is assumed to point at the object reference point. The VIO system thus considers two of the five (three positional and two rotational) degrees of freedom for camera placement and two of the three positional degrees of freedom of the illuminator. The objects are polyhedral with Lambertian reflectance. There is no spatial variation of brightness within a polyhedral face since the illumination source is taken as a point distant from the object, and interreflections and shadow effects are neglected.

Similar to the HEAVEN system discussed in the previous section, sensor illuminator configurations in VIO are generated on an adaptively tessellated spherical surface and evaluated according to the task criteria. VIO first calculates an image representation of the expected scene for each pair of camera and illuminator locations, and then evaluates chosen image feature attributes (e.g., edge visibility, edge contrast) to assess the goodness of each such pair. Threshold values are specified for each image feature attribute so that camera-illuminator pairs that do not meet this threshold condition are eliminated from further consideration.

Facets on the tessellated sphere that are associated with views containing the same edges, are grouped into regions. Each region is then evaluated based on the following criteria:

- 1) Robustness against potential object placement uncertainty: number of region facets/total number of tessellated sphere facets;
- 2) Edge visibility: number of edges visible in a region/total number of object edges;
- 3) Edge occlusion: $1 - (\text{number of occluded edges} / \text{total number of object edges})$; and
- 4) Contrast: average facet contrast over all facets in this region, where facet contrast is the normalized sum of all edge contrast values when the illuminator is placed at this facet.

The suitability value associated with a region on the viewing sphere is calculated as a normalized sum of the above attribute values. Regions that do not meet a specified threshold value are eliminated. Each camera and illuminator location is given a suitability value equal to that of the region in which it lies.

The previous evaluation of camera and illuminator locations does not consider the coupling between the two in an actual sensing situation. For this reason, the VIO system also evaluates the camera-illuminator setup as a pair. In order to limit the number of camera-illuminator combinations, only those that satisfy a set of constraints are considered (e.g., the number of visible edges common to both the camera and illuminator locations must exceed a preselected value, the camera cannot be placed in front of the illuminator and vice versa). Camera-illuminator pairs that meet these constraints are then ranked based on another set of attributes associated with image features and their combination in the scene. The image feature attributes considered were

- 1) The length of the feature edge in the image,
- 2) Its contrast,
- 3) The angle formed between an object edge corresponding to an edge in the image and an expected displacement direction in the case of location uncertainty of the target object, and
- 4) The angular disparity of object edges: this angular disparity is computed as the normalized sum of angles between all pairs of object edges corresponding to the feature edges in the image.

All this information regarding a particular sensing situation is collected in a VIO cell, which is a data frame representation containing all the characteristics of a setup proposal. The camera and illuminator suitability values, as well as the suitability value of their combination, are combined to obtain the global VIO processability value of the setup. Based on this criterion, setups are ranked and the setup with the highest processability value is taken to be the best choice.

In addition to stationary objects, VIO also plans a sensor setup for moving target objects by taking into account the previous and future sensor configurations and attempting to minimize changes in the camera or illumination setup as the target object moves.

Overall, the VIO system has the advantage of providing both sensor and illumination planning components in a unified

framework. Furthermore, VIO generates a set of candidate camera-illuminator setups along with a quantitative assessment of each. However, by taking an all-possible-combinations strategy at various stages of the algorithm, VIO may be confronted by combinatorial problems. The many thresholds that are employed to mitigate this problem may not work in some cases as they are preselected in an ad hoc manner.

3) *The ICE System:* The illumination control expert (ICE) system was developed at the University of Washington [86]–[88] to achieve automatic sensor and illumination placement planning for machine vision tasks.

Similar to both VIO and HEAVEN, ICE takes a generate-and-test approach. Camera and light positions are generated and then evaluated based on selected criteria. In ICE, as in VIO and HEAVEN, the camera points at an object reference point, and the camera and light source are placed on the surface of a sphere with its center at the origin of the object coordinate system. This spherical surface is then approximated by a discrete number of points in such a way that the distance between any two neighboring points is approximately the same.

Overall, the ICE system places emphasis on the illumination placement planning component of the problem. For the camera placement planning component, the task constraint considered in ICE is edge visibility. The optimal sensor position is considered to be that from which certain feature edges appear least occluded. This criterion is evaluated based on the aspect graph representation of the object [46], [69]. The regions of camera positions that are found using this approach are such that the set of object edges of interest is either fully or partially visible. These sensor viewpoints are then assessed based on the ratios of the unoccluded portion of these edges, as seen from a particular viewpoint, to their actual edge lengths. This metric, however, varies within an aspect graph cell. Consequently, the finer spatial subdivision (i.e., tessellation) of the spherical viewing surface, discussed earlier, is used when evaluating this metric.

The illumination planning problem is posed in an optimization setting with edge contrast as the criterion to be optimized. Illuminator locations are assessed based on

- 1) The ratio of the length of an edge for which a given contrast threshold is exceeded to the total edge length and
- 2) The amount by which the predicted edge contrast exceeds this threshold over a certain percentage of the length of the edge.

The above optimality criteria are evaluated by simulating placement of the illuminator at points on the discrete approximation of the spherical viewing surface (see Fig. 4), while the camera is placed at the one location that was independently determined. The contrast at an edge point is obtained from a prediction of the image based on the image intensity equation and is computed as the difference between intensities of the reflected light from two neighboring small patches, one on each of the faces that meet at this edge. The image intensities are computed based on the illumination and sensor parameters and the bidirectional reflectance function. Contrast for the

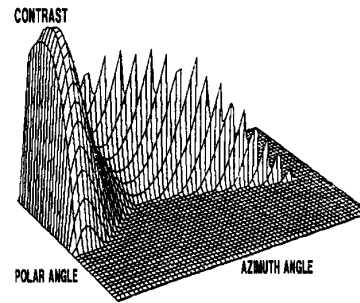


Fig. 4. Contrast criterion in ICE (from [88]).

entire edge is evaluated on a finite number of points along the edge and then fitted with piecewise continuous polynomials producing what is defined as the contrast graph for this edge. This graph describes how contrast varies spatially along an edge and is used to determine the distribution of contrast. The contrast distribution function is then used to evaluate the optimality criteria for illuminator placement that were itemized above.

On the whole, the ICE system has the advantage of providing an illumination planning system that deals with objects with both a general geometry and general reflectance characteristics. For instance, objects bounded by curved surfaces and modeled by a hybrid reflectance model can be handled. The reflectance model used in ICE is based on the Torrance-Sparrow reflectance model [77] with modifications made to the specularly reflected component so that polarization can be taken into account. However, in the actual examples given of ICE, the effect of polarization of the incident or reflected light is not taken into account and thus its importance for illumination placement planning is not clear.

In ICE the sensor placement planning problem is addressed independently of that for illuminator placement. An optimal camera location is found without considering where the light source will be placed, and the light source placement problem is then posed given the sensor location. However, as was discussed in Section II-B.3, these two problems are not generally separable. Even though specular reflection is taken into account in ICE, the camera placement planning problem is posed independently of the illuminator location and, therefore, the solution obtained may be suboptimal for the combined sensor and illuminator planning problem.

4) *Remarks on the Generate-and-Test Approach:* The generate-and-test method is appealing as an initial solution to the sensor planning problem because it draws from some well known techniques for discretizing spaces and efficiently searching them. Several positive attributes of this approach are listed below.

- 1) Generate and test uses a relatively straightforward approach to the occlusion problem, performing a search over the facets of the dome. It is simple to implement algorithmically.
- 2) The tessellation of the sphere is a hierarchical procedure that can be used to make search more efficient, since search over this sphere is at the heart of the method.
- 3) Intersecting multiple constraints posed as regions on the sphere to find a feasible region is straightforward.

4) Lighting constraints can be modeled with the same representation as visibility constraints.

There are a number of potential problems with generate and test, however, that will require ongoing research. These are listed below.

1) The computational cost of finely tessellating high dimensional parameter spaces and then searching them is high. While the hierarchical approach can guide this search, it still may require an exhaustive search (e.g., over all combinations of facets and possible occluding objects).

2) The solution space assumes a viewpoint in the center of each facet. This approximation may not be useful, for example, in characterizing an entire facet's visibility. The general question of scale (i.e., sampling rate) remains for these methods.

3) The generate-and-test methods work well for point features; however, they do not extend easily to extended features such as lines and planar faces. With more complex features, the method may become computationally burdensome as well as overly conservative in its solutions.

4) A number of generalized viewpoint parameters are either ignored or assumed fixed in the generate-and-test approach. In particular, sensor resolution is posed as a global radius on the object feature, which causes difficulties with multiple feature observation. Field-of-view constraints are also difficult since this encompasses either extended features or multiple features. In addition, camera orientation parameters are usually not solved for; they are assumed to be pointing to a designated location on the feature.

D. Review of Work Taking a Synthesis Approach

An alternative to the discrete approach is to model the constraints as analytic functions. This is somewhat intimidating because the function spaces are of high dimension (8 dimensions in the case of a generalized viewpoint that includes camera location, orientation, and the optical parameters of focal length, aperture, and focus). However, by characterizing these spaces (or subspaces formed by restricting the number of parameters) as analytic functions, a number of benefits are derived.

1) *The Automatic Sensor and Illuminator Positioning Work of SRI*: An important early work in this area is the work of Cowan and Kovesi [22], [24] in which camera locations were automatically generated satisfying several of the geometric constraints discussed in Section II-B.3; that is, chosen features of polyhedral objects are visible, in focus, within the sensor field of view, and spatially resolvable to a given specification. This work was then later extended by Cowan and Bergman [23] to include planning of illuminator placement.

The general approach is to formulate both problems as constraint satisfaction problems. Each task requirement generates an equivalent geometric constraint which in turn is satisfied in a domain of admissible locations in three-dimensional space (see Fig. 5). The admissible domains obtained for each task requirement are then intersected in order to determine locations that satisfy all constraints simultaneously.

As discussed in Section II-B.3, the general camera and illuminator planning problem involves a parametric space of

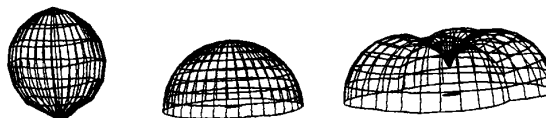


Fig. 5. From left to right, examples of the admissible regions for the resolution, field-of-view, and depth-of-field constraints from [24].

dimension higher than three. However, by considering a subset of the sensor parameters, Cowan and Kovesi were able to intersect admissible domains of individual constraints in three-dimensional space. For instance, the camera orientation is not explicitly taken as a parameter to be planned, but rather is assumed to take on an implicit value (e.g., in the field-of-view analysis, the camera viewing direction is taken from the viewpoint to the center of the sphere circumscribing all features of interest). Similarly, parameters associated with the lens optical settings are not planned for the most part. This approach has the difficulty that the chosen values for such parameters need to be at least admissible for all constraints. While a globally admissible value for the camera orientation is determined, the image plane to perspective center distance is found for each constraint separately and not in a manner that guarantees satisfaction of all constraints.

For the visibility constraint, Cowan and Kovesi initially computed the three-dimensional region from where a convex target can be viewed entirely from above a convex occluding polygon (the same approach is taken later in [49] in their work in automatic grasping). For the case where the occluding object is not a convex polygon but rather a polyhedron, the total occluded region is obtained as the union of the component occluded regions of the faces of the occluding polyhedron. The faces that are concave or contain holes must be decomposed into convex shapes for their method to be applied. Their algorithm for convex occluding and target polygons, as described, has quadratic computational complexity in the number of edges, which can be improved to linear as discussed in [80].

A general comment about the resolution, focus, and field-of-view constraints is that although they are characterized by functional relationships, iterative techniques are then used to solve for the sensor parameters, even when closed-form solutions are obtainable.

For the resolution constraint, Cowan and Kovesi defined resolution to be the minimum angle subtended by a given incremental surface length at the perspective center of the lens, rather than the minimum number of pixels per surface length which is commonly used. This definition can be shown to be more conservative than the one customarily used. In addition, it has the advantage of formulating the resolution constraint independently of the viewing orientation since the angle subtended at the viewpoint does not depend on the viewing orientation. As a result, the number of sensor parameters involved is reduced and the resolution satisfying domains of individual features can be intersected in a lower dimensional parametric space, when globally admissible resolution regions are sought.

Cowan and Kovesi also determined the permissible region for camera placement such that all points of the features to be observed are in focus. They made use of a depth-of-field formula developed by Krotkov [48] that gives the range of distances of feature points that are in focus (i.e., their blur circle diameter is less than the minimum pixel dimension). The set of viewpoints for which a feature has its farthest point at a distance equal to the upper limit of this range, and its closest point at a distance equal to the lower limit defines the boundary of the locus of admissible viewpoints for this constraint. Although these distances (as any object distances governed by the Gaussian lens law) need be measured along the optical axis, Cowan and Kovesi measured them radially, that is, from the perspective center of the lens to each feature point. With such a distance definition they developed an iterative procedure to initially obtain the domain of viewpoints that satisfy the focus constraint for a line segment. For this, it is assumed that the camera optical axis passes through the line segment midpoint. In three dimensions, a similar iterative approach is used to solve the focus constraint at small increments of azimuth and elevation. The camera optical axis is assumed to pass through the center of a circumscribing circle or sphere of the features depending, respectively, on whether they are coplanar or not. As a result, the orientational degrees of freedom of the sensor are omitted and the discretization of sensor configuration space migrates from a genuine synthesis approach. The sensor planning system has been tested on a piece of sheet metal with a number of different size slots and holes as well as on a tray of three-dimensional parts.

The analytical framework set up by the SRI group also has the advantage of including other constraints on imaging such as illumination. Cowan and Bergman [23] have extended the results of the sensor planning described above by computing the appropriate range of lens aperture settings and the three-dimensional region where a point-source light may be placed to illuminate the scene. Including illumination in the planning component is difficult, since it increases the number of controllable parameters greatly (i.e., position and orientation of the light source, kind of source—extended or point—and source energy). Cowan and Bergman have made some reasonable assumptions about the imaging and lighting in a robotic work cell to show the viability of both camera and illumination planning. In their work, the reflectance of a surface is modeled as having a diffuse component that follows Lambert's cosine law and a specular component forming a lobe within an angular tolerance of the perfect specular angle. Interreflections are ignored and viewing of the surface is assumed to be on-axis of the camera. Their analysis shows that the dynamic range constraint of the sensor (i.e., the brightness—no surface in the scene is too dark or too light) from the diffuse component of reflection lies between two closed curves that are proportional to the incident angle of the illumination. The specular region of the reflectance can also be computed and is subtracted from the region bounded by the closed curves to yield a set of admissible regions that satisfy the illumination constraints. The illumination planning system has been demonstrated with a camera on a robot arm and a point-source mounted on a second robot arm. A multicolored

planar surface with patches of different reflectance values was used as a test object. The camera was then placed in the position computed by the sensor planning system without illumination constraints. Given this location of the camera, the illuminator was moved to positions along the constraint curves of minimum and maximum brightness and the resulting images were consistent with the expected values (within roughly 5%).

This work has recently been extended to include constraints on illumination for forming edges (gradients in images) [25]. By planning regions of acceptable viewing and illumination parameters to form edges, a feature based object recognition strategy can be generated that plans new viewing locations of an unknown object. The new viewing position is selected to maximize the number of features (edges) that can be seen, given certain strong assumptions about illumination and reflectance.

2) *The MVP System:* The authors are developing a vision planning system, machine vision planner (MVP), that automatically determines vision sensor parameter values that satisfy several of the sensor constraints discussed in Section II-B.3 [72], [73], [75], [79], [80]. MVP takes as input the object geometry information from a CAD database, as well as models of the camera and lens, and determines camera poses and settings for which features of interest of polyhedral objects are visible (occlusion-free positions of the sensor), contained entirely in the sensor field of view, in focus, and resolvable by the sensor to a given specification. At this stage the MVP system does not include illumination planning.

The parameters that are determined include the three positional degrees of freedom of the sensor— $\mathbf{r}_o(x, y, z)$ —and the two orientational degrees of freedom—pan and tilt angles—described by a unit vector ν along the viewing direction. Rotation with respect to the optical axis is not considered. In addition, the three optical parameters—the back principal point to image plane distance d , the focal length f , and the aperture of the lens a —are taken into account. Thus, planning is done in eight-dimensional imaging space and a point in this space is defined as a *generalized viewpoint* $V(\mathbf{r}_o, \nu, d, f, a)$.

In MVP all task constraints are formulated analytically. Using concepts from geometry, illumination, and optics, each task requirement is modeled by an equivalent analytical relationship which in turn is satisfied in a domain of admissible values in the space of parameters to be planned. Generally speaking, for each constraint the admissible domain for sensor placement and setting is a region in eight-dimensional imaging space bounded by the hypersurfaces that are determined by these analytical relationships. The component admissible domains obtained for each task requirement are then combined in order to find parameter values that satisfy all constraints simultaneously. For this purpose, the problem is posed in an optimization setting in which a globally admissible eight-dimensional viewpoint is sought that is central to the admissible domain; that is, distant from the bounding hypersurfaces described by the constraint equations. Such a generalized viewpoint is considered desirable since it is robust in the event of inaccuracy of either sensor placement or setting. The analytical relationships for each task constraint provide the constraints of the optimization, while the objective function is

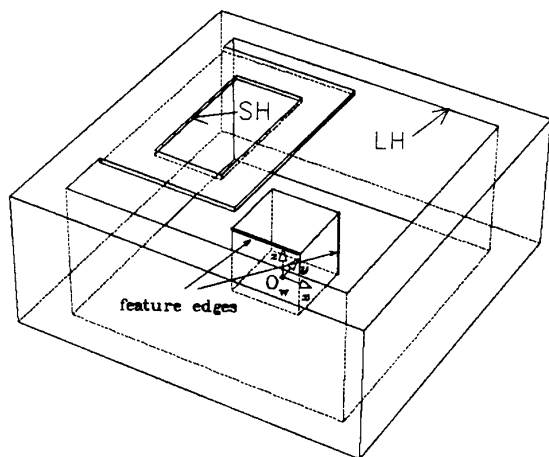


Fig. 6. CAD model of the object used in the MVP system with the feature edges to be observed (from [74]).

chosen so as to characterize the distance between a generalized viewpoint and the bounding hypersurfaces.

Once a central generalized viewpoint is determined from the optimization, it then needs to be realized in the actual sensor setup. While the task constraints are expressed in terms of the generalized viewpoint, the parameters that can be controlled in an actual sensor setup are generally different. In order to achieve these planned sensor parameter values, a mapping is established between the planned parameters (e.g., camera pose and optical settings) and the parameters that can be controlled (e.g., end effector pose, zoom, and focus settings). This mapping between the two parameter spaces is provided by calibration models that embody knowledge of the geometric relationships of the manipulator, sensor, illuminator, and optical relationships of the lenses.

Objects in MVP are modeled as general polyhedra with concave and convex faces, with or without holes. The visibility planning algorithm of MVP first considers a sufficient subset of the faces of the observed polyhedron as polygons in three-dimensional space that are potentially occluding the feature to be observed. The algorithm then determines the three-dimensional occluded regions between these occluding polygons and each target feature. The individual occluded regions of these faces are unioned in order to generate the occluded region of the target feature for the polyhedron as a whole. The complement of the occluded region is the visibility region from where the entire target feature can be viewed. The visibility region is polyhedral and defines a piecewise analytic constraint for viewpoint placement. The visibility region for the object features of Fig. 6 is shown in Fig. 7. In this case, the visibility region consists of two connected components that correspond to viewing the edge features through the small hole *SH* and the large hole *LH* of the object (see Fig 6).

The resolution, depth of field, and field of view constraints are characterized by analytic closed-form relationships. These relationships are derived using the geometry of perspective transformation and concepts from optics. The locus of generalized viewpoints that satisfies each constraint separately is

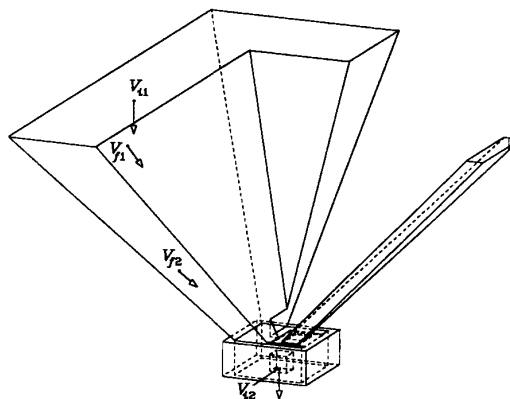


Fig. 7. Visibility region for both feature edges (from [74]).

expressed as follows:

$$g_i(\mathbf{r}_O, \mathbf{v}, d, f, a) \geq 0 \quad (4)$$

where $i = 1$ for the resolution constraint, $i = 2a, 2b$ for the near and far limits of the depth-of-field constraint and $i = 3$ for the field-of-view constraint. It should be noted that there is a g_1 relationship for each linear feature to be resolved, while for the depth-of-field and field-of-view constraints there is a unique set of relationships g_{2a} , g_{2b} , and g_3 for all features. In addition to the above inequality constraints, there is also an auxiliary equality constraint $g_s = 0$ which expresses the unit vector condition for the viewing vector.

While the constraints address the admissibility of the computed solution, the optimization function is constructed in a way that characterizes the robustness of the computed solution. The measure used to assess the robustness of a solution with respect to the resolution, field-of-view, and depth-of-field constraints is the value of each constraint relationship g_i , $i = 1, 2a, 2b, 3$. This is appropriate since a large positive value of g_i indicates that the constraint is satisfied comfortably, a small positive value indicates marginal satisfaction, and inadmissible solutions give rise to negative values. Similarly, for the visibility constraint a measure of this type is also formulated. For this purpose, the minimum distance d_v from the viewpoint to the polyhedron describing the visibility region is chosen: $g_4 = \pm d_v$, where $+d_v$ or $-d_v$ depending on whether the point is inside or outside the visibility volume, respectively. The optimization function is taken to be a weighted sum of the above component criteria, each of which characterizes the quality of the solution with respect to each associated requirement separately. Thus, the optimization function is written as

$$f = \max(\alpha_1 g_1 + \alpha_{2a} g_{2a} + \alpha_{2b} g_{2b} + \alpha_{3b} g_3 + \alpha_4 g_4) \quad (5)$$

subject to $g_i \geq 0$, $i = 1, 2a, 2b, 3, 4$ and $g_5 = 0$, where α_i , are the weights. These weights are currently chosen so that the contribution of each constraint to the objective function is of the same order of magnitude and prevents a subset of the constraints from dictating the optimization. Given the above formulation, the optimization starts with an initial point in the domain of possible generalized viewpoints and then



Fig. 8. Camera view of the features from viewpoint V_{f2} computed by the MVP system (from [74]).

generates a generalized viewpoint that is globally admissible and locally optimal, as described by the optimization function. For this computed viewpoint, all constraints are satisfied with the largest margin in the neighborhood of the initial point.

As an example, for the object and the edge features shown in Fig. 6, the two initial viewpoints that are chosen to start the optimization, V_{i1} and V_{i2} , and the corresponding camera viewpoints that are computed by the MVP system, V_{f1} and V_{f2} , are shown in Fig. 7 along with their associated viewing vectors. It can be seen from Fig. 7 that the initial guess viewpoint V_{i1} is chosen to lie inside the visibility region with a viewing vector in the $-z$ direction. For V_{i1} the field-of-view and focus constraints are violated. The initial guess viewpoint V_{i2} is chosen to lie outside the visibility region with a viewing vector again in the $-z$ direction. For V_{i2} all constraints are violated. The viewpoints V_{f1} and V_{f2} that are determined by the MVP system can be seen in Fig. 7 to lie inside the visibility region with viewing vectors in the direction of the features. Camera views taken by a robot vision system from the computed viewpoints and with the computed optical settings, verify that the constraints are indeed satisfied. The view of the object taken by the robot vision system from the computed viewpoint V_{f2} is shown in Fig. 8.

With this global optimization formulation of the problem, the MVP system provides a framework in which individual constraints can be combined and other constraints can be incorporated. However, further work is needed in order to study the convergence of the method, the dependence of the method on the quality of the initial guess, and the selection of appropriate weight settings.

The MVP system is notable for its generality of handling general viewing configurations (e.g., nonorthogonal viewing, general convex, and concave polyhedral objects) and determining a complete set of imaging sensor parameters as described by its concept of a generalized viewpoint. Another advantage of the MVP system is that it addresses the associated sensor modeling problems (e.g., in [79], models for a programmable zoom lens are described). Together with the SRI system of Cowan and Kovesi, MVP studies the broadest range of feature detectability constraints.

3) *Automatic Determination of Viewer Orientation by Anderson:* Anderson [5], [6] describes algorithms for the automatic determination of the viewing direction which will render

visible an entire object, whenever this is possible, and will minimize the maximum local distortion in the image due to perspective projection. An object point is considered visible in this analysis if and only if it lies in front of the observer; that is, occlusion is not considered and there are no field of view limits. As a result, the viewing direction that is sought is such that all object points are in front of the viewpoint and perspective distortion is minimized. The viewpoint location is assumed to be known. This is often the case in many graphics applications, where the viewpoint and the object are fixed and an orientation is sought for which the object is centered in the field of view.

Anderson proposes this viewing direction rather than viewing directions through the center of mass or the center of the smallest box containing the object, since the latter are ad hoc and are often inadequate when the viewpoint is close to the object. Anderson shows [5] that minimizing the maximum perspective distortion is equivalent to minimizing the maximum value of the off-axis angle of any object point (i.e., the angle between the viewing direction and the vector from the viewpoint to an object point). Consequently, the computed viewing direction corresponds to the axis of the narrowest circular cone emanating from the viewpoint and containing the object. By considering that the viewpoint can be taken as the origin and that the vectors to the object points can be normalized to unit length without affecting the solution, the problem can be reduced to a problem in spherical geometry. More specifically, the narrowest cone enclosing the object points is generated by finding the smallest circle on the unit sphere that encloses the object point projections onto this sphere.

Anderson provides an algorithm in [5] to find the narrowest circular cone directly and several more efficient algorithms in [6] that are based on the aforementioned reduction of the original problem to a problem in spherical geometry. These latter algorithms are based on existing algorithms for analogous plane geometry problems (e.g., finding the spherical convex hull of a set of points and the spherical Voronoi diagram of a set of points).

In summary, in this work by Anderson the viewing direction for the constraint of minimizing perspective distortion is determined. Requirements such as magnifying the object features, having them in focus, and avoiding occlusion of the features are not considered.

4) *Remarks on the Synthesis Approach:* On the whole, the synthesis approach forces an understanding of the causal relationships between the parameters to be planned and the goals to be achieved. In this way an insight is provided to the problem, rather than just blindly applying search. To this end, the task constraint are modeled analytically, bringing to bear a powerful characterization that can be utilized to answer more questions (e.g., "Using a quantitative measure, how comfortably does the computed viewpoint satisfy the depth-of-field constraint?") As discussed in Section II-D.2, with such an analytical formulation of the problem, techniques such as optimization can be used as a framework in which several task constraints can be combined. Or if globally admissible values can be generally determined for a subset of the parameters,

the method discussed in Section II-D.1 has the advantage of reducing the dimensionality of the search space and providing approximations of the entire admissible space of solutions in the dimensions that it considers. In addition, the accuracy of the techniques taking the synthesis approach can be directly computed from the analytical relationships and easily adjusted to the specifications of the task.

E. Review of Work Taking an Expert Systems Approach

1) *The Lighting Advisor*: Batchelor [8], [9] has built an expert system that suggests possible lighting and viewing configurations to the designer of a machine vision system for industrial inspection. This expert system includes the LIGHTING ADVISOR which provides advice regarding what lighting configurations are required in given circumstances. Two other advisors which will provide advice about the camera and the lens are also being developed.

The program asks a series of questions regarding the object to be inspected and the inspection task to be performed; for example, information regarding the reflectance characteristics of the object (specular, diffuse, opaque, transparent, translucent) as well as the type of feature or defect that is to be highlighted. The program then displays a line drawing which shows a sketch of the recommended lighting configuration.

Batchelor intends to incorporate more rules in the LIGHTING ADVISOR in order to actually implement the lighting arrangement that has been recommended by automatically controlling the illumination in a flexible inspection cell. This cell provides facilities such as a pick-and-place arm, a 3-degree-of-freedom (x, y, θ) table, and computer-controlled lighting with which a wide range of illumination and viewing techniques can be implemented. With these additional expert system rules, the recommended illumination will be converted into a set of control signals for the lamps.

2) *Other Expert Systems for Sensor Planning*: Along similar lines to the expert system approach of Batchelor, there has been work by Penn Video, Inc., a subsidiary of the Ball Corporation, that sells a lighting advisor program described by Novini in [57]. In [45] another expert system is presented for the design of vision algorithms. This system provides advice on the selection of effective features and the image processing operations to extract the chosen features. In addition, the Industrial Technology Institute in Dearborn, MI, has compiled a database of lighting and viewing techniques as described in [20].

3) *Remarks on the Expert Systems Approach*: The systems following an expert system approach address the high-level aspects of the problem in which a particular viewing and illumination technique is chosen from a catalogue [10], for instance, whether front or back illumination is more appropriate for the particular object and feature to be observed. However, these qualitative approaches need to be extended in order to represent and determine the exact spatial relationships between the sensor, the illuminator, and the object.

F. Review of Related Work in Sensor Simulation Systems

Several sensor simulation systems have been developed and have been used as tools to plan new robot workcell installa-

tions and test modifications to existing ones. Most of the early systems [55] modeled simple distance and tactile sensors by employing the solid modeling operation of intersection in order to detect contact between the sensor (e.g., the robot gripper surface or the light beam of a noncontact distance sensor) and objects in the robot workcell. More recently, sensor simulation systems have been built that include simulation of more complex sensors such as cameras in conjunction with light sources. In this section, we shall discuss two such sensor simulation systems: VANTAGE and ROSI.

1) *The VANTAGE System*: VANTAGE [40], [41] is a geometric and sensor modeler for model-based vision systems. In general, vision systems deal with two-dimensional appearances of an object and these appearances are determined by both the object itself (e.g., its geometric and photometric properties) and the sensor used to observe it. Typically, the object properties are part of the geometric models that are used in model-based vision systems. However, the sensory information that also determines the object appearances is not included in such modelers. VANTAGE provides sensory information such as the feature types that a sensor detects, the particular object features that are detectable at a given sensor configuration, and the reliability with which the detectable features can be observed.

In VANTAGE, a feature configuration space is first defined in order to represent the spatial relationship between the sensor and the feature. In this representation only the angular relationship between the sensor and feature coordinate systems is considered. A way to then specify sensor detectability and sensor reliability over this space is given. On the one hand, sensor detectability specifies what features can be detected for a given sensor configuration, while sensor reliability specifies the expected uncertainty in sensory measurement and the uncertainty of the scene features that are derived from this measurement.

Both illuminators and detectors are considered in a uniform manner as *generalized sources* with two properties: the illumination direction and the illumination configurations. For light sources, the illumination direction is the direction of the light source, while the illuminated configurations are the collection of features that can be illuminated for the particular illuminator pose, assuming that the illumination direction is not occluded. In the case of detectors, these terms correspond, respectively, to the viewing direction of the detector and the set of features that are visible to the detector. With these properties, feature detectability with respect to a generalized source can be characterized as follows: A feature is illuminated by a generalized source if the feature coordinate system lies in the illuminated configurations of the illuminator and the illumination direction is not occluded.

Feature detectability for a general vision sensor is decomposed into such illumination conditions of its component generalized sources and Boolean operations between them. This decomposition defines the *sensor composition tree* in VANTAGE. For example, as shown in Fig. 9, a light-stripe range finder can detect the portions of an object surface onto which the illuminator projects light directly and which the camera can observe. Thus, the detectability of such a vision

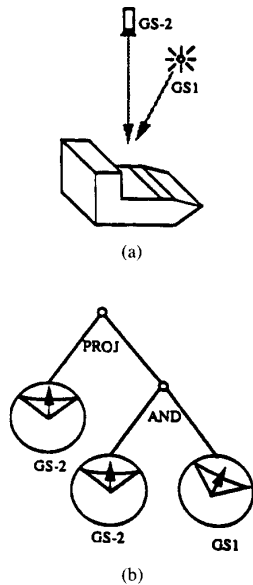


Fig. 9. Sensor detectability conditions for a camera and light stripe illuminator pair from VANTAGE [41].

sensor is decomposed into two generalized source illumination conditions and Boolean AND operation between them as shown in the sensor composition tree of Fig. 9. Faces which are both illuminated by the light source and observable by the camera are then projected (PROJ operation in Fig. 9) onto the image plane in order to generate the two-dimensional appearance of the object.

VANTAGE also takes into consideration *sensor noise* when determining feature detectability. This is required because sensor noise may interfere with detectability even if the above detectability constraints are satisfied. Since most vision sensors detect features based on brightness values, detectability of a feature will depend on the distribution of feature brightness. In VANTAGE a feature is taken to be detectable if its brightness is above a threshold. Thus, the feature brightness distribution is modeled as the probability of the feature brightness being above this threshold. Feature brightness is computed by assuming that the object surface is Lambertian and adding a zero-mean Gaussian noise term.

In addition to sensor detectability, sensor reliability is also modeled in the VANTAGE system. Sensor reliability addresses the uncertainty in sensory measurement and its propagation to uncertainty of the scene features. The sources of such uncertainty include the variance in brightness values and the variance in light source direction. These two variances are, respectively, the major causes of uncertainty for vision sensors. The uncertainty in sensory measurement for a light-stripe range finder and a photometric-stereo sensor is characterized in [40]. In addition, for the case of the light-stripe range finder the propagation of uncertainty from a two-dimensional region image feature to a three-dimensional feature area is modeled.

Appearances of the object when observed by the sensor at hand can be generated using the object geometry and

this sensor detectability information, and applying the sensor composition tree to the scene. This is done in three steps:

- 1) The three-dimensional face segments that are illuminated by each component generalized source are determined,
- 2) These segments are then combined according to the Boolean operations of the sensor composition tree, and finally,
- 3) The three-dimensional face segments that have been determined to be visible and either illuminated or in shadow are projected onto the image plane in order to generate the two-dimensional appearance of the object.

The ability of VANTAGE to predict the two-dimensional appearances of objects when viewed by a sensor can be readily employed as the kernel of a tool that plans sensor configurations. Alternatively, in [39], VANTAGE is used to automatically generate object recognition programs. The various appearances of objects are generated and then categorized into possible aspects, where each aspect is defined to have the same combination of detected two-dimensional faces. The recognition strategy is then comprised of an aspect classification component which is represented as an interpretation tree, and a linear shape change component which is generated by using correspondences between image regions and model faces. This recognition strategy is then converted into a program using an object library.

2) *The ROSI Sensor Simulation System:* Raczkowski *et al.* [36], [58] have developed ROSI, a sensor simulation system for robotic applications that includes simulation of a camera in conjunction with light sources. In the ROSI system, an industrial scene can be visualized with a simulated camera given:

- 1) The geometric description of the cell objects,
- 2) The optical properties of their surfaces,
- 3) The physical and geometric description of the light sources, and
- 4) The physical and geometric properties of the sensors.

The geometric description of an object is given by a polyhedral boundary representation based on CAD generated data. The surface optical properties include the coefficient of diffuse reflectance, the index of refraction of the surface material, the ratio of diffuse to specular reflectance, and a factor describing the rms slope of the distribution function of the reflectance. These properties are used in the bidirectional reflection model of Torrance and Sparrow [77]. Inter-reflections are not considered. The light sources are modeled as Lambertian, emitting light evenly, and are defined by the center point, the area, the normal vector, and the spectral radiation density of the emitting surface. The camera model consists of three parts: geometric information regarding the camera pose, parameters that describe the lens system, and parameters describing the electronic hardware. In particular, the lens system is modeled by a thin lens with an aperture diaphragm. The parameters characterizing the lens system are the focal length, the aperture setting, and the distance (i.e., focus setting) of the lens. This last parameter is equivalent to the back nodal point to image plane distance d , which is discussed in Section II-B.1. The electronic hardware is

described by the number of rows and columns of pixels, the dimensions of the rows and columns, and the sensitivity of the sensor.

For the given configuration of the camera and the light sources, and with the object and sensor model information, the system:

- 1) Performs a hidden-surface removal to determine the surfaces of the scene visible to the camera,
- 2) Performs a similar hidden-surface removal, with the light source as the viewpoint, in order to determine the light sources illuminating the individual visible surfaces,
- 3) Clips all visible surfaces to the view volume of each individual pixel in order to determine the list of polygons illuminating each pixel,
- 4) Determines the radiation flux illuminating the individual pixels of the sensor plane using the Torrance and Sparrow reflectance model,
- 5) Generates the corresponding gray or color levels of the pixels based on the incoming radiation flux, the minimum and maximum irradiance of a pixel, and its relative spectral sensitivity.

The ROSI sensor simulation system has been used as a tool to support the planning process of robot cells. While it does not synthetically determine robot cell configurations that are suitable for a particular task, ROSI provides the capability of varying camera and light-source parameters in order to search for a satisfactory simulated view of the scene.

3) *Remarks on the Sensor Simulation Systems:* The systems in this category have many commonalities with various approaches taken to the general graphics problem. While concurrent work in the computer graphics community has looked at the realistic rendering of scenes, the algorithms employed have a high computational cost (i.e., ray tracing) which may make them infeasible for this application. Also, sensor models like those developed in VANTAGE are outside the scope of research in computer graphics. These issues are essential, however, for sensor simulation systems since their goal is to simulate real-world sensors in order to determine and evaluate a sensor system and its operation.

Compared with the other approaches reviewed previously, the sensor simulation systems use much more illumination information. However, both these systems and the systems taking a generate-and-test approach can easily incorporate more complete reflectance and illumination models (e.g., mutual reflections, area light sources) much in the same way that work in computer graphics has done.

III. CONCLUSIONS AND FUTURE DIRECTIONS IN SENSOR PLANNING FOR FEATURE DETECTABILITY

The area of sensor planning for machine vision tasks remains a very rich research area that spans several fields of study: computer vision, robotics, optics, geometry, and computer graphics. The methods presented in this survey represent, in most cases, the initial efforts of researchers to attack a difficult problem with many degrees of freedom.

For the task of feature detectability, one family of solutions to this problem has attempted to discretize space and either

fix or assume some of the necessary sensor parameters, performing an efficient search of the remaining parameter space for feasible solutions. The other family of solutions has tried to analytically model the task constraints (using a variety of assumptions about parameters) and then find an analytical description of the feasible solutions. In the latter case, the high dimensionality of these constraint surfaces requires a nonlinear optimization approach, which is also a form of efficient search. It is not clear which of these methods is currently more viable. The simplicity of the discrete space approach is appealing, as is its hierarchical implementation. On the other hand, the inaccuracy of the tessellations, assumption of certain viewing parameters and inability to easily deal with multiple features make it less desirable. The analytical approach is elegant and extensible to multiple features and off-axis viewing, albeit at the cost of constrained nonlinear optimization.

While both methods will continue to be used by researchers, further work is needed in order to achieve robust and general purpose camera and illumination planning strategies and systems. There are common open problems that need to be addressed by future systems. We include a partial list of these areas that we believe are fruitful for future research.

A. Modeling and Incorporating Other Constraints

The sensor planning systems that have been developed have chosen a subset of the feature detectability constraints discussed in Section II-B.3. However, there are several other feature-detectability constraints as well as a broad range of constraints that arise in other sensing tasks, which could be modeled and incorporated in sensor planning systems.

An example of a constraint that arises in all sensor planning systems involving movement as part of sensor reconfiguration is *collision avoidance*. The values of the sensor parameters resulting from planning must be such that the sensor can be positioned without collision with the environment. Similarly, no point in the path to the final sensor configuration should lead to collision with the environment. One way to address this constraint would be to draw on previous work in path planning in order to model collision-free space, and then treat collision-free space in a manner similar to the way visibility regions are dealt with in the MVP system.

If the environment, the features of interest, or both, move in time along paths that in some cases are known in advance, this gives rise to a family of time-dependent constraints for *dynamic sensor planning*. With the exception of the initial work in the VIO system, the methods presented in this survey deal mostly with planning of parameters for static scenes. The sensor planning capabilities of the MVP system have recently been extended [2] to function in environments where objects in the environment move in a way that is known a priori, but the features to be observed are not moving. The problem of dynamic sensor planning is of importance because many sensing environments in practice are more often dynamic than not. Thus, the usefulness of these systems will be increased greatly if planning can be done in a dynamic fashion to include moving sensors, light sources, targets, and obstacles. If planning can be done efficiently and cheaply, then complex

tasks such as assembly can be monitored in real-time using these methods.

An example of a constraint that is generic to all sensor planning systems is the constraint associated with the *operating range* of the employed sensor. If such constraints are ignored, it is often the case that the computed sensor parameter values are not achievable by the sensor at hand. For example, the optical parameter f , d , and a discussed in Section II-B.1 often cannot simultaneously assume the planned values for a given lens. Techniques to characterize the operating range of a general lens fall in the area of sensor modeling which is discussed at the end of this section.

B. Modeling and Incorporating Other Sensors

While this survey has focused on planning using vision sensors, a host of other sensors are currently being used in many robotics and manufacturing applications. These include tactile sensors, three-dimensional range sensors, force-torque sensors, and acoustic sensors. Some of the techniques described above extend nicely for planning with these sensors (e.g., the visibility line of sight can be thought of as an approach vector for a tactile probe). However, further work needs to be done to properly integrate these sensors and their unique constraints into the overall planning system.

C. Relaxing Some of the Assumptions Made in Current Approaches

1) *Feature Uncertainty*: Underlying most of the work done in this area of sensor planning has been the assumption that the environment is known. This is often not the case in actual sensing situations. As a first step towards relaxing this assumption, sensor planning systems could investigate how to incorporate *feature uncertainty*. One approach to address feature uncertainty could attempt to combine the deterministic sensor planning techniques with approaches that use a decision theoretic framework [15], [31] in order to address the inherent uncertainty in sensing. In this respect, deterministic sensor planning systems adopting a synthesis approach provide the advantage of including analytical relationships that characterize the admissible loci. When feature uncertainty is included, these relationships can be used in order to compute the variation of the admissible loci based on known perturbations of the features and the environment.

2) *Accuracy*: The previously proposed extension leads to the related problem regarding the accuracy of the solutions for sensor parameter values. In the case of the discrete state approach, characterization of the error induced by the discretization is needed. In the analytic approach, creating tolerance constraint volumes (as in geometric tolerancing with CAD/CAM systems) may be useful in determining how robust a solution may be to sensor and calibration error.

3) *Complex Object Models*: Most of the current work has considered polyhedral objects and polyhedral type features. Objects with curved surfaces and their corresponding object features need to be included as well. As a first step in this direction, the approach of posing the problem in terms of polyhedral approximations of objects with curved surfaces

needs to be investigated. Polyhedral approximations may often be sufficient, since in many cases the resulting sensor parameter values may be robust enough to be admissible for the constraints of the original object as well. Recent work by [27], [47], [68] in analyzing the visibility of more complex shapes, such as generalized cylinders, may be useful in extending the planning systems to include curved surface objects.

4) *Constraint Integration*: Techniques that combine the admissible domains of individual constraints in order to determine globally admissible solutions need to be investigated further. Optimization, used in MVP and ICE, provides one such framework. However, further work is needed in order to study convergence properties, appropriate weight values, and the dependence of the final solution on the quality of the initial guess. Also, the optimization formulation employed currently constitutes a simple integration scheme based on the assumption that multiple and coupled objectives can be combined in an additive sense into a single global objective. Such a formulation has inherent problems with conflicting objectives. Alternative optimization schemes such as multiple objective optimization should also be explored with each task constraint constituting an individual objective.

Techniques that determine admissible regions rather than an optimal point; should also be investigated. A generalization of the method discussed in Section II-D.1 could yield the admissible space of solutions for a subset of the parameters having first determined globally admissible values for the other parameters. In order to approximate the admissible region of all constraints, interval-based techniques for solving large systems of nonlinear constraints such as those described in [32] will prove useful.

The question of weighting constraints according to the importance of each in the general planning objective is a notable one. It may be more important to have a feature visible at a poor resolution than not being visible at all. In determining solutions to parameter values, feasible solutions may be found but ranking them as to suitability is a more difficult task. While this may be somewhat task dependent, it may be important to isolate some parameters and solve for them alone. When optimization is used for integration, this amounts to determining the weights on each of the constraints that make up the objective function that needs to be optimized.

E. Illumination Planning

While some of the systems mentioned previously have either explicitly or implicitly dealt with planning of lighting parameters, current work in illumination planning is quite restricted. It should be recognized, however, that the problem of planning of general lighting for machine vision is extremely difficult. Most of the work has used point sources of light that are incident on convex Lambertian surfaces. These models, while useful, are not analogous to actual lighting conditions seen in current applications. Higher order lighting/reflectance models that include such phenomena as multiple sources (both point and extended), specularity, and interreflections from concave surfaces need to be found to properly plan lighting parameters.

F. Sensor and Illumination Modeling

A subtle, but important problem in this area is that of modeling the sensors and mapping the planned sensor parameters to real hardware systems. The coupling between the sensor planning and modeling problems becomes apparent when the planned parameters that have been determined by some means are to be achieved in an actual setup. For this, the planned parameter values need to be mapped to controllable sensor settings using the sensor models. For some parameters (camera location and orientation) this can be relatively straightforward. However, parameters such as entrance pupil diameter, focal length, and image plane to back principal point distance can be difficult to automatically set without extensive calibration effort. This is particularly true for zoom lens cameras, which provide more degrees of freedom for general imaging situations but at the cost of mapping planned parameters to their actual settings [76], [82]. A related problem is the creation of accurate sensor noise models that can be combined with the other known imaging constraints to accurately predict correct parameters.

While sensor modeling can utilize much of the work that has been done in camera and lens calibration, there has been very little previous work in the area of illumination modeling. Much in the same way that research in camera modeling has developed calibration methods in order to position and set cameras and their lenses, work in illumination modeling will also need to develop similar calibration methods with which to characterize the attributes of an illumination source and to position illuminators in light-in-hand arrangements. For this purpose, drawing on the wide body of knowledge in photometry and radiometry [42] will prove useful.

REFERENCES

- [1] A. L. Abbott and N. Ahuja, "Active surface reconstruction by integrating focus, vergence, stereo, and camera calibration," in *Int. Proc. Conf. Comput. Vision*, 1990, pp. 489-492.
- [2] S. Abrams, P. K. Allen, and K. Tarabanis, "Dynamic sensor planning," in *Proc. 1993 IEEE Int. Conf. Robotics and Automat.*, 1993.
- [3] N. Ahuja and J. Veenstra, "Generating octrees from object silhouettes in orthographic views," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 11, pp. 137-149, Feb. 1989.
- [4] J. Aloimonos, I. Weiss, and A. Bandyopadhyay, "Active vision," in *Proc. Conf. Comput. Vision*, 1987, pp. 35-54.
- [5] D. P. Anderson, "An orientation method for central projection programs," *Comput. & Graphics*, vol. 6, no. 1, pp. 35-37, 1982.
- [6] D. P. Anderson, "Efficient algorithms for automatic viewer orientation," *Comput. & Graphics*, vol. 9, no. 4, pp. 407-413, 1985.
- [7] R. Bajcsy, "Active perception," *Proc. IEEE*, vol. 76, no. 8, pp. 996-1005, Aug. 1988.
- [8] B. G. Batchelor, "A Prolog lighting advisor," in *Proc. SPIE Intell. Robots and VIII: Syst. and Applicat.*, vol. 1193, pp. 295-302, 1989.
- [9] ———, "Integrating vision and AI for industrial applications," in *Proc. SPIE Intell. Robots and Comput. Vision VIII: Syst. and Applicat.*, vol. 1193, pp. 168-173, 1989.
- [10] B. G. Batchelor, D. A. Hill, and D. C. Hodgson, *Automated Visual Inspection*. Bedford, UK: IFS Ltd., 1985.
- [11] A. K. Bejczy, W. S. Kim, and S. C. Venema, "The Phantom robot: Predictive displays for tele-operation with time delay," in *Proc. 1990 IEEE Int. Conf. Robotics and Automat.*, 1990, pp. 546-551.
- [12] R. C. Bolles and R. A. Cain, "Recognizing and locating partially visible objects: The local-feature-focus method," *Int. J. Robot. Res.*, vol. 1, no. 3, pp. 57-82, 1982.
- [13] R. C. Bolles and P. Horaud, "3DPO: A three-dimensional part orientation system," *Int. J. Robot. Res.*, vol. 5, no. 3, pp. 3-26, 1986.
- [14] A. Califano, R. Kjeldsen, and R. M. Bolle, "Data and model driven foveation," in *Proc. 10th Int. Conf. Patt. Recognition*, 1990, pp. 1-7.
- [15] A. Cameron and H. Durrant-Whyte, "A Bayesian approach to optimal sensor placement," Rept. OUEL 1759/89, Department of Engineering Science, Oxford University, Oxford, UK, 1989.
- [16] A. Cameron and H. L. Wu, "Identifying and localizing electrical components: A case study of adaptive goal-directed sensing," Tech. Note TN-90-085, Philips Laboratories, Briarcliff, NY, 1990.
- [17] O. I. Camps, L. G. Shapiro, and R. M. Haralick, "PREMIO: An overview," in *Proc. IEEE Workshop on Directions in Automat. CAD-Based Vision*, 1991, pp. 11-21.
- [18] L. L. Chin and T. C. Woo, "Computational geometry on the sphere with application to automated machining," Tech. Rept. TR 89-30, Industrial and Operations Engineering, Univ. of Michigan, Ann Arbor, MI, Aug. 1989.
- [19] J. J. Clark and N. J. Ferrier, "Modal control of an attentive vision system," in *Int. Conf. Comput. Vision*, 1988, pp. 514-523.
- [20] M. Coletta and K. Harding, "Picking the right lighting," *MVA/SME's Vision Technol. Quart.*, pp. 1-3, Jan. 1990.
- [21] C. I. Connolly, "The determination of next best views," in *Proc. 1985 IEEE Int. Conf. Robotics and Automat.*, 1985, pp. 432-435.
- [22] C. K. Cowan, "Model based synthesis of sensor location," in *Proc. 1988 IEEE Int. Conf. Robotics and Automat.*, 1988, pp. 900-905.
- [23] C. K. Cowan and A. Bergman, "Determining the camera and light source location for a visual task," in *Proc. IEEE Int. Conf. Robotics and Automat.*, pp. 509-14, 1989.
- [24] C. K. Cowan and P. D. Kovesi, "Automatic sensor placement from vision task requirements," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 10, pp. 407-16, May 1988.
- [25] C. K. Cowan and B. Modayur, "Edge-based placement of camera and light-source for object recognition and location," in *Proc. 1993 IEEE Int. Conf. Robotics and Automat.*, 1993.
- [26] R. I. D. Cowie, "The viewer's place in the theories of vision," in *Int. Joint Conf. Artificial Intell.*, 1983, pp. 952-958.
- [27] D. Eggert and K. Bowyer, "Computing the orthographic projection aspect graph of solids of revolution," in *Proc. Workshop on Interpretation of 3D Scenes*, 1990, pp. 102-108.
- [28] C. Goad, "Special purpose, automatic programming for 3D model-based vision," in *Proc. DARPA Image Understanding Workshop*, 1983, pp. 94-104.
- [29] J. L. Gordillo and A. Lux, "Synthesizing vision programs from robot task specifications," in *3rd Int. Symp. Robotics Res.*, 1985, pp. 149-154.
- [30] W. E. L. Grimson, "Sensing strategies for disambiguating among multiple objects in known poses," *IEEE J. Robot. Automat.*, vol. RA-2, no. 4, Dec. 1986.
- [31] G. D. Hager, *Task Directed Sensor Fusion and Planning*. Norwell, MA: Kluwer Academic Publishers, 1990.
- [32] ———, "Interval-based techniques for sensor data fusion," Tech. Rep., Yale University, New Haven, CT, 1990.
- [33] C. Hansen and T. Henderson, "CAGD-based computer vision," in *Proc. IEEE Workshop on Computer Vision*, 1987, pp. 100-105.
- [34] T. H. Hong and M. O. Schneier, "Describing a robots workspace using a sequence of views from a moving camera," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 7, pp. 721-726, Nov. 1985.
- [35] B. K. P. Horn, *Robot Vision*. New York: McGraw-Hill, 1986.
- [36] M. Huck, J. Raczkowski, and K. Weller, "Sensor simulation in robot applications," in *Proc. Advanced Robotics Program, Workshop on Manipulators, Sensors and Steps Towards Mobility 1987*, pp. 197-209.
- [37] S. A. Hutchinson and A. C. Kak, "Planning sensing strategies in robot work cell with multi-sensor capabilities," *IEEE Trans. Robot. Automat.*, vol. 5, no. 6, pp. 765-783, Dec. 1989.
- [38] K. Ikeuchi, "Generating an interpretation tree from a CAD model for 3D-object recognition in bin picking tasks," *Int. J. Comput. Vision*, vol. 1, no. 2, pp. 145-165, 1987.
- [39] K. Ikeuchi and T. Kanade, "Towards automatic generation of object recognition programs," Tech. Rept. CMU-CS-88-138, Carnegie Mellon University, Pittsburgh, PA, May 1988.
- [40] ———, "Modeling sensors: Toward automatic generation of object recognition program," *Comput. Vision, Graphics, Image Process.*, vol. 48, pp. 50-79, 1989.
- [41] K. Ikeuchi and J. C. Robert, "Modeling sensors detectability with the VANTAGE geometric/sensor modeler," *IEEE Trans. Robot. Automat.*, vol. 7, pp. 771-784, Dec. 1991.
- [42] H. A. E. Keitz, *Light Calculations and Measurements*. The Netherlands: Philip's Technical Library, 1955.
- [43] J. R. Kender and D. G. Freudenstein, "What is a 'degenerate' view?" in *Proc. 1987 DARPA Image Understanding Workshop*, 1987, pp. 589-598.
- [44] H. S. Kim, R. C. Jain, and R. A. Volz, "Object recognition using multiple views," in *Proc. 1985 IEEE Int. Conf. Robotics and Automat.*, 1985, pp. 28-33.

- [45] Y. Kitamura, H. Sato, and H. Tamura, "An expert system for industrial machine vision," in *Proc. 10th Int. Conf. Patt. Recognition*, 1990, pp. 771-773.
- [46] J. J. Koenderink and A. J. van Doorn, "The internal representation of solid shape with respect to vision," *Biolog. Cybern.*, no. 32, pp. 211-216, 1979.
- [47] D. J. Kriegman and J. Ponce, "Computing exact aspect graphs of curved objects: Solids of revolution," *Proc. Workshop on Interpretation of 3D Scenes*, 1990, pp. 116-122.
- [48] E. Krotkov, "Exploratory visual sensing with an agile camera," Ph.D. dissertation, University of Pennsylvania, 1987.
- [49] C. Laugier, A. Ijel, and J. Trocaz, "Combining vision based information and partial geometric models in automatic grasping," in *Proc. 1990 IEEE Conf. Robotics and Automat.*, 1990, pp. 676-682.
- [50] L. Lieberman, "Model-driven vision for industrial automation," in *Advances in Digital Image Processing*, Stucki, Ed. New York: Plenum, 1979, pp. 235-248.
- [51] M. Magee and M. Nathan, "Spatial reasoning, sensor repositioning and disambiguation in 3d model based recognition," in *Proc. Workshop on Spatial Reasoning and Multi-Sensor Fusion*, 1987.
- [52] W. N. Martin and J. K. Aggarwal, "Volumetric description of objects from multiple views," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 5, pp. 150-158, Mar. 1983.
- [53] J. Maver and R. Bajcsy, "How to decide from the first view where to look next," in *Proc. 1990 DARPA Image Understanding Workshop*, 1990, pp. 482-496.
- [54] ———, "Occlusions as a guide for planning the next view," Tech. Rep., GRASP Lab., Univ. of Pennsylvania, 1991.
- [55] J. Myer, "An emulation system for programming sensory robots," *IBM J. Res. Develop.*, vol. 25, no. 6, pp. 955-962, Nov. 1981.
- [56] J. L. Mundy, "Industrial machine vision—Is it practical?" in *Machine Vision*, H. Freeman, Ed. San Diego, CA: Academic, 1988.
- [57] A. Novini, "Lighting and optics expert system for machine vision," in *Proc. Optics, Illumination, Image Sensing*, pp. 1005-1019, 1988.
- [58] J. Raczkowski and K. H. Mittenbuehler, "Simulation of cameras in robot applications," *Comput. Graphics Applicat.*, pp. 16-25, Jan. 1989.
- [59] S. Sakane, M. Ishii, and M. Kakikura, "Occlusion avoidance of visual sensors based on a hand eye action simulator system: HEAVEN," *Adv. Robot.*, vol. 2, no. 2, pp. 149-165, 1987.
- [60] S. Sakane, R. Niepold, T. Sato, and Y. Shirai, "Illumination setup planning for a hand-eye system based on an environmental model," *Adv. Robot.*, vol. 6, no. 4, pp. 461-482, 1992.
- [61] S. Sakane and T. Sato, "Automatic planning of light source and camera placement for an active photometric stereo system," in *Proc. 1991 IEEE Int. Conf. Robotics and Automat.*, 1991, pp. 1080-1087.
- [62] S. Sakane, T. Sato, and M. Kakikura, "Model-based planning of visual sensors using a hand-eye action simulator system," in *3rd Int. Conf. Advanced Robotics (ICAR)*, 1987.
- [63] ———, "Planning focus of attentions for visual feedback control," *Trans. Soc. Instrum. Cont. Eng.*, vol. 24, no. 6, pp. 608-615, June 1988.
- [64] D. D. Seligmann and S. Feiner, "Automated generation of intent-based 3D illustrations," *Comput. Graphics*, vol. 25, no. 3, July 1991.
- [65] S. A. Shafer, "Automation and calibration for robot vision systems," Tech. Rep. CMU-CS-88-147, Carnegie Mellon University, Pittsburgh, PA, May 1988.
- [66] A. Shmuel and M. Werman, "Active Vision: 3D from an image sequence," in *Proc. 10th Int. Conf. Patt. Recognition*, 1990, pp. 48-54.
- [67] A. J. Spyridi and A. G. Requicha, "Accessibility analysis for the automatic inspection of mechanical parts by coordinate measuring machines," in *Proc. 1990 IEEE Int. Conf. Robotics and Automat.*, 1990, pp. 1284-1289.
- [68] T. Sripradisvarakul and R. Jain, "Generating aspect graphs for curved objects," in *Proc. Workshop on Interpretation of 3D Scenes*, 1990, pp. 109-115.
- [69] J. Stewman and K. Bowyer, "Creating the perspective projection aspect graph of polyhedral objects," in *Int. Conf. Comput. Vision*, 1988, pp. 494-500.
- [70] M. J. Swain and M. Stricker, "Promising directions in active vision," Tech. Rept. CS 91-27, University of Chicago, Chicago, IL, Nov. 1991.
- [71] K. Tang, T. C. Woo, and J. Gan, "Maximum intersection of spherical polygons and workpiece orientation for 4- and 5-axis machining," Industrial and Operations Engineering, Univ. of Michigan, Ann Arbor, MI, 1990.
- [72] K. Tarabanis and R. Y. Tsai, "Computing viewpoints that satisfy optical constraints," in *Proc. CVPR '91: The Comput. Soc. Conf. on Comput. Vision and Patt. Recognition*, 1991.
- [73] ———, "Computing occlusion-free viewpoints," in *Proc. CVPR '92: The Comput. Soc. Conf. on Comput. Vision and Patt. Recognition*, June 15-18, 1992.
- [74] K. Tarabanis, R. Y. Tsai, and S. Abrams, "Planning viewpoints that simultaneously satisfy several feature detectability constraints for robotic vision," in *5th Int. Conf. Advanced Robotics (ICAR)*, 1991.
- [75] K. Tarabanis, R. Y. Tsai, and P. K. Allen, "Automated sensor planning for robotic vision tasks," in *Proc. 1991 IEEE Int. Conf. Robotics and Automat.*, 1991.
- [76] K. Tarabanis, R. Y. Tsai, and D. S. Goodman, "Calibration of a computer controlled robotic vision sensor with a zoom lens," Submitted to *Comput. Vision Graphics, and Image Processing*, 1992.
- [77] K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surfaces," *J. Opt. Soc. Amer.*, vol. 57, no. 9, pp. 1105-1114, Sept. 1967.
- [78] R. Y. Tsai and M. A. Lavin, "Three-dimensional mechanical part measurement using a vision/robot system," Tech. Rept. RC 10506, IBM T.J. Watson Research Center, Yorktown Heights, NY, May 1984.
- [79] R. Y. Tsai and K. Tarabanis, "Model-based planning of sensor placement and optical settings," in *Proc. Sensor Fusion II: Human and Mach. Strategies*, 1989.
- [80] ———, "Occlusion-free sensor placement planning," in *Machine Vision for Three-Dimensional Scenes*, H. Freeman, Ed. Orlando, FL: Academic, 1990.
- [81] P. Whaite and F. P. Ferrie, "From uncertainty to visual exploration," in *Proc. Int. Conf. Comput. Vision*, 1990, pp. 690-697.
- [82] R. G. Wilson and S. A. Shafer, "Modeling and calibration of zoom lenses," in *Camera Calibration and Orientation Determination*, A. Gruen and T. S. Huang, Eds. New York: Springer-Verlag, 1993.
- [83] W. J. Wolfe, D. W. Mathis, M. Magee, and W. Hoff, "Task panel sensing with a movable camera," in *Proc. SPIE Intell. Robots and Comput. Vision VIII: Syst. and Applicat.*, vol. 1193, 1989.
- [84] K. H. Womack, "Front-end foresight," *MVA/SME's Vision Technol. Quarter.*, p. 4, Jan. 1990.
- [85] S. Xie, T. W. Calvert, and B. K. Bhattacharaya, "Planning views for the incremental construction of body models," in *Proc. 8th Int. Conf. Patt. Recognition*, 1986, pp. 154-157.
- [86] S. Yi, R. M. Haralick, and L. G. Shapiro, "Automatic sensor and light source positioning for machine vision," Tech. Report EE-ISL-89-04, University of Washington, Seattle, WA, Sept. 1989.
- [87] ———, "An illumination model for machine vision," Tech. Rept. EE-ISL-89-03, University of Washington, Seattle, WA, Sept. 1989.
- [88] ———, "Automatic sensor and light source positioning for machine vision," in *Proc. 10th Int. Conf. Patt. Recognition*, 1990, pp. 55-59.

Konstantinos Tarabanis, for a biography and photograph, please see p. 85 of this issue.

Peter K. Allen, for a biography and photograph, please see p. 85 of this issue.

Roger Y. Tsai, for a biography and photograph, please see p. 85 of this issue.