Illumination Planning for Object Recognition in Structured Environments *

Hiroshi Murase
NTT Basic Research Labs
Morinosato, Atsugi-shi
Kanagawa 243-01, Japan
murase@siva.ntt.jp

Shree K. Nayar
Department of Computer Science
Columbia University
New York, N.Y. 10027
nayar@cs.columbia.edu

Abstract

This paper addresses the problem of illumination planning for robust object recognition in structured environments. Given a set of objects, the goal is to determine the illumination for which the objects are most distinguishable in appearance from each other. For each object, a large number of images is automatically obtained by varying pose and illumination. Images of all objects, together, constitute the planning image set. The planning set is compressed using the Karhunen-Loeve transform to obtain a low-dimensional subspace. For any given illumination, objects are represented as parametrized manifolds in the subspace. The minimum distance between the manifolds of two objects represents the similarity between the objects in the correlation sense. The optimal illumination is therefore one that maximizes the shortest distance between object manifolds. Results produced by the illumination planner have been used to enhance the performance of an object recognition system.

1 Introduction

Research in the area of computer vision can be classified into two broad categories. One involves the development of passive vision systems for the analysis of unstructured environments, such as, outdoor scenes. The second is geared towards the development of vision systems for structured environments, such as, industrial assembly lines. In the case of structured environments, imaging and illumination parameters are often controllable. As a result, effective vision sensors and algorithms can be developed to recover various types of scene properties that would generally be impossible to estimate in unstructured environments. In structured environments, vision systems are used to perform a variety of tasks, such as, inspect manufactured parts, recognize objects and sort them, or aid a robot in assembly operations. In each of these cases, the illumination of the environment can be selected to enhance the reliability and accuracy of the vision system. Currently, illumination parameters are selected by human operators using the trial and error approach. The resulting illumination is seldom one that maximizes the performance of the vision system.

Lately, automatic illumination planning has emerged as a topic of research interest. Most of this work focuses on determining light source positions that maximize the detectability of image features such as edges. Cowan and Bergman [2] use CAD (geometric) models of objects to compute source positions for which all brightness values in the image lie within the sensor's dynamic range. The positions, orientations, and reflectance parameters of the objects are assumed to be known. Using the same assumptions, Cowan and Nitzan [3] compute source positions that ensure that the brightness contrast at selected edges on objects exceeds a threshold value. Recently, Yi et al. [11] used the Torrance and Sparrow reflectance model to obtain accurate predictions of the brightness of object points. Yi propagates errors due to noise in image brightness to estimate errors in the positions of edges. The planning problem then is to determine the source direction that maximizes the accuracy of edge positions.

Addressing a different problem, Sakane et al. [9] determine optimal source directions for a photometric stereo system. They use the accuracy of computed surface normals and the range of computable normals as criteria for selecting the optimal source directions. Recently, Batchelor [1] proposed an expert system that uses the knowledge of illumination experts to suggest the best illumination for a given vision application. The illumination plan proposed by an expert is based on his/her experience and not a careful theoretical analysis of the problem. Hence, the suggested illumination is not guaranteed to be optimal.

In this paper, we present a novel approach to illumination planning. There are several parameters that characterize illumination, such as, source direction, source distance, source size, and the spectral characteristics (color) of the source. We describe the planning approach using source direction and source color as the parameters of interest. In theory, as well as practice, the same approach can be extended to incorporate other illumination parameters. Object *appearance* is used as the criterion for finding the optimal illumination. Given a set of objects, our goal is to determine the illumina-

*This research was conducted at the Center for Research in Intelligent Systems, Department of Computer Science, Columbia University. This research was supported in part by the David and Lucile Packard Fellowship and in part by ARPA Contract No. DACA 76-92-C-0007.

tion that makes the objects maximally different from each other in the *correlation* [1] sense. Once the optimal illumination has been planned, its validity must be verified. For this, we have used an object recognition system that identifies 3D objects and computes their poses from brightness images [6]. Experiments using this system show that the planned illumination produces the highest recognition rate. The paper is concluded with a discussion on the merits and limitations of the proposed method.

2 Illumination Planning

In this section, we discuss the problem of finding optimal illumination for recognition and pose estimation of a set of objects. Our approach differs in two fundamental ways from previous work on illumination planning. (a) We do not use 3D geometric (CAD) models or reflectance models of objects during illumination planning. Our planning approach uses raw 2D images of objects to determine the optimal illumination for correlation-based object recognition. (b) We do not assume that the pose of each object is fixed and known a-priori. First, we describe the planning of illumination direction. Later, these results are extended to the planning of illumination color. The planning system described here is fully automated.

2.1 Illumination and Object Appearance

The appearance of an object depends on its shape, it reflectance properties, its pose, and the illumination conditions. The first two factors are intrinsic properties of the object that do not vary. On the other hand, object pose and illumination can vary substantially from one scene to the next. In most machine vision applications, the pose of the object is not within the control of the vision system; objects show up in the scene with arbitrary poses. That leaves us with illumination. In structured environments, such as industrial assembly lines, the illumination of the scene can be controlled to provide the "best" images of the objects of interest. Fig. 1 shows images of an object obtained using different illumination directions. These images illustrate that object appearance is very sensitive to the direction of illumination.

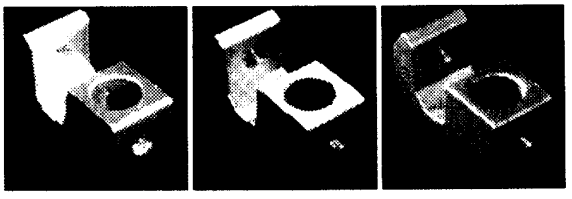


Figure 1: The effect of illumination direction on object appearance.

[1] Correlation, or template matching, remains one of the most widely used recognition strategies in the industrial arena. Finding optimal illumination for this task is therefore a problem of significant practical relevance.

Though we have posed the planning problem as one of finding the optimal source direction, the approach can also be used to determine the optimal source position. In fact, since the planning method uses 2D images and not 3D object models, other source characteristics (such as, source size and color) as well as sensor characteristics (such as, spectral response and optical settings) can be incorporated into the planning process. The only requirement is that these source and sensor characteristics be varied during the image acquisition stage of planning.

2.2 Planning Image Set

While constructing the planning image set we need to ensure that all object images are of the same size. Each digitized image is segmented into an object region and a background region. The background is assigned zero brightness value and the object region is re-sampled such that the larger of its two dimensions fits the size we have selected for image representation. The result is an image that is normalized with respect to scale and thus invariant to the magnification of the imaging system. This image is written as a vector $\hat{\mathbf{x}}$ by reading pixel values in a raster scan fashion:

$$\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, ..., \hat{x}_N]^{\mathrm{T}} \tag{1}$$

The above vector represents an unprocessed brightness image. Alternatively, processed images such as blurred images, first derivatives, and second derivatives may be used. For the purpose of developing the illumination planning method we use raw brightness images, keeping in mind that the planning methodology is directly applicable to any other image type.

We would like the illumination planning system to be unaffected by variations in the intensity of illumination or the aperture of the imaging system. This can be achieved by normalizing each image, such that, the total energy contained in the image is unity. This brightness normalization transforms each measured image $\hat{\mathbf{x}}$ to a normalized image $\mathbf{x}$, where $\mathbf{x} = \hat{\mathbf{x}} / \| \hat{\mathbf{x}} \|$.

We denote each normalized image as $\mathbf{x}_{r,l}^{(p)}$ where r is the rotation or pose parameter, l represents the illumination direction, and p is the object number. The image set obtained by varying the pose of an object for a given illumination direction l can be written as:

$$\mathbf{X}_l^{(p)} \triangleq \{ \mathbf{x}_{1,l}^{(p)}, \mathbf{x}_{2,l}^{(p)} ....., \mathbf{x}_{R,l}^{(p)} \} \tag{2}$$

where, R is the total number of discrete poses used for each object. Let P be the total number of objects and L be the total number of illumination directions. Then, the *planning image set* for the entire set of objects is:

$$\{ \mathbf{X}_1^{(1)}, ..., \mathbf{X}_L^{(1)}, \mathbf{X}_1^{(2)}, ..., \mathbf{X}_L^{(2)}, \mathbf{X}_1^{(P)}, ..., \mathbf{X}_L^{(P)} \} \tag{3}$$

In our experiments, we have used a motorized turntable to vary object pose. This gives us pose variations about a single axis. We have used several light sources positioned in a plane around the turntable. For each object, illumination

direction is automatically varied and for each illumination direction a set of images is obtained by rotating the object. Fig.2 shows some of the images in the set obtained by varying the pose of the object in Fig. 1 for a given illumination direction.

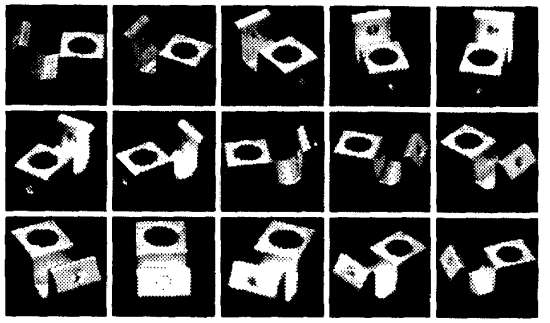


Figure 2: Image set for the object shown in Fig.1 obtained by varying pose, for a given illumination direction.

2.3 Universal Eigenspace

Consecutive images in the planning image set are correlated to a degree since pose and illumination variations between these images are not large. Our objective is to take advantage of this correlation and compress the large planning set into low-dimensional representations of each object's appearance. A suitable compression technique is the Karhunen-Loeve method [7] where the eigenvectors of an image set are computed and used as orthogonal basis functions for representing individual images. Though, in general, all eigenvectors of an image set are required for perfect reconstruction of any particular image, only a few eigenvectors are sufficient for illumination planning. These eigenvectors are the dimensions of a subspace[2] that we refer to as the *universal eigenspace*.

First, the average $\mathbf{c}$ of all images in the planning set is subtracted from each image in the set. The result is the following *image matrix*:

$$\mathbf{Y} \triangleq \{ \mathbf{x}_{1,1}{}^{(1)} - \mathbf{c}, \ldots\ldots, \mathbf{x}_{R,1}{}^{(1)} - \mathbf{c}, \ldots\ldots, \mathbf{x}_{R,L}{}^{(P)} - \mathbf{c} \} \tag{4}$$

$\mathbf{Y}$ is $N \times M$, where $M = RLP$ is the total number of images in the planning set and N is the number of pixels in each image. To compute eigenvectors of the image set we define the *covariance matrix*:

$$\mathbf{Q} \triangleq \mathbf{Y}\,\mathbf{Y}^{\mathbf{T}} \tag{5}$$

$\mathbf{Q}$ is $N \times N$, clearly a very large matrix since a large number of pixels constitute an image. The eigenvectors $\mathbf{e}_i$ and the corresponding eigenvalues λ_i of $\mathbf{Q}$ are to be computed by solving the well-known eigenvector decomposition problem:

$$\lambda_i \mathbf{e}_i = \mathbf{Q}\mathbf{e}_i \tag{6}$$

[2]This subspace has previously been used in vision to classify handwriting [4] and human faces [10], as well as for recognition and pose estimation of 3D objects [6].

Though, all N eigenvectors of the planning image set are needed to represent images exactly, only a small number ($k \ll N$) of eigenvectors are generally sufficient for capturing the primary appearance characteristics of the objects. These k eigenvectors correspond to the largest k eigenvalues of $\mathbf{Q}$ and constitute the universal eigenspace. An issue that has yet to be addressed is the selection of k. One approach is to select k such that the first k eigenvectors capture the important appearance variations in the image set, that is:

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \geq T_1 \tag{7}$$

where the threshold T_1 is close to, but less than, unity. The denominator in the above expression equals the trace of the covariance matrix $\mathbf{Q}$. For the objects we have used in our experiments, universal eigenspaces with less than 10 dimensions ($k < 10$) are found to be adequate. A property of the eigenspace that is fundamental to our planning methodology is that it is the *optimal* subspace for estimating the correlation between images (see [7], [6]).

Computing the eigenvectors of a large matrix such as $\mathbf{Q}$ can prove computationally very intensive. Efficient algorithms for this are described in [7], [5] and summarized in [6].

2.4 Parametric Eigenspace Representation

Our objective is to obtain a measure of how well the set of objects can be discriminated under illumination from each of the source directions. The image set $\mathbf{X}_l{}^{(p)}$ includes images of the object p, obtained for different object poses r, while it is illuminated by the source l. Each image $\mathbf{x}_{r,l}{}^{(p)}$ in $\mathbf{X}_l{}^{(p)}$ is projected to the universal eigenspace. This is done by subtracting the average image $\mathbf{c}$ from $\mathbf{x}_{r,l}{}^{(p)}$, then finding the dot product of the result with each of the k eigenvectors, or dimensions, of the eigenspace. The result is a single point $\mathbf{g}_{r,l}{}^{(p)}$ in eigenspace:

$$\mathbf{g}_{r,l}{}^{(p)} = [\mathbf{e}_1, \mathbf{e}_2, \ldots\ldots, \mathbf{e}_k]^{\mathbf{T}} (\mathbf{x}_{r,l}{}^{(p)} - \mathbf{c}) \tag{8}$$

By projecting all the planning samples in $\mathbf{X}_l{}^{(p)}$, we obtain a set of discrete points in the universal eigenspace. Pose variation between any two consecutive images in $\mathbf{X}_l{}^{(p)}$ is small. As a result, consecutive images are strongly correlated and their projections in eigenspace are close to one another [3]. The discrete points obtained by projecting all samples in $\mathbf{X}_l{}^{(p)}$ can be assumed to lie on a manifold:

$$\mathbf{g}_l{}^{(p)}(\theta_1, \theta_2, \theta_3) \tag{9}$$

where, θ_1, θ_2, and θ_3 are the three continuous rotation parameters needed to describe pose in three-dimensional

[3]This assumption holds well except when the object is either highly specular or has high-frequency texture. In such cases, an incremental pose variation can cause dramatic changes in image brightness.

space. The above manifold is referred to as the *parametric eigenspace representation*; it is a compact representation of the appearance of object p when illuminated by source l. In our experiments, we rotate the object about a single axis. This variation in pose is sufficient for objects that have a finite number of stable configurations when placed on a planar surface. Thus, the above manifold is reduced to a curve with a single parameter: $\mathbf{g}_l^{(p)}(\theta_1)$. Fig.3 shows the parametrized eigenspace representation of the object shown in Fig.1. The eigenspace used is 8-dimensional and is computed using a planning set that includes two object image sets. The figure shows only three of the most significant dimensions of the eigenspace since it is difficult to display and visualize higher-dimensional spaces. For illumination planning, such a curve is computed for each object, for each illumination direction.

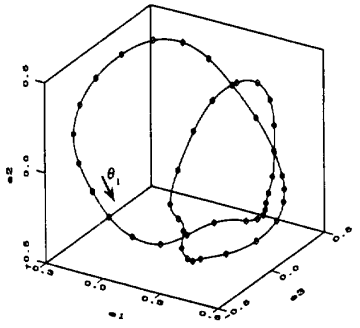


Figure 3: Curve obtained in universal eigenspace by projecting the object image set shown in Fig.2.

2.5 Optimal Illumination Direction

Consider two objects, say p and q, from the set used to compute the universal eigenspace. For each light source direction l, we compute parametric curves for the two objects: $\mathbf{g}_l^{(p)}(\theta_1^{(p)})$ and $\mathbf{g}_l^{(q)}(\theta_1^{(q)})$. Here, the parameters $\theta_1^{(p)}$ and $\theta_1^{(q)}$ represent rotations of objects p and q, respectively. The shortest euclidean distance between the two curves in eigenspace is computed as:

$$d_l^{(p,q)} = \min_{\theta_1^{(p)}, \theta_1^{(q)}} \| \mathbf{g}_l^{(p)}(\theta_1^{(p)}) - \mathbf{g}_l^{(q)}(\theta_1^{(q)}) \| \quad (10)$$

The $\theta_1^{(p)}$ and $\theta_1^{(q)}$ values that produce the minimum distance $d_l^{(p,q)}$, correspond to poses of the two objects for which they appear most similar (in correlation) when illuminated by the source l. The illumination planning problem is formulated as follows: Find the source direction $\tilde{l}$ that maximizes the minimum distance $d_l^{(p,q)}$ between the object curves. This *max-min* strategy gives us the safest illumination direction for the worst case where the two objects have poses for which they are most similar in appearance.

The above example includes only two objects. The *max-min* strategy is easily extended to a set of P objects. For a given illumination direction l, we now have P curves in the universal eigenspace. The minimum distance $d_l^{(p,q)}$ is computed for all pairs of objects in the object set, resulting in P^2 minimum distances. The minimum of all these distances, say d_l, represents the worst case for the entire object set. The source direction $\tilde{l}$ that maximizes d_l is then the *optimal source direction* for the object set. Fig. 4 shows the eigenspace curves of two objects used in the experiments, for a particular illumination direction. The solid line segment indicates the shortest distance between the two curves.

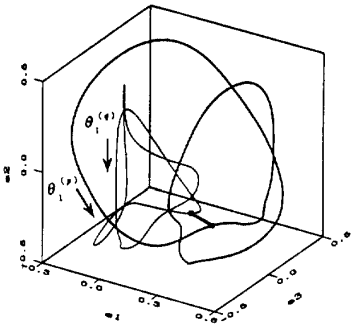


Figure 4: Parametric eigenspace curves of two different objects obtained for a given illumination direction. The shortest distance (line segment) between the two curves represents the worst case poses for which the objects appear most similar in the correlation sense.

2.6 Optimal Illumination Color

The above approach can also be used to plan other illumination parameters. Here, we show how illumination color can be computed to render a set of objects minimally correlated in appearance with each other. Many man-made objects have regions with different spectral properties. For such objects, the spectral characteristics of the illumination can be controlled to robustly identify them. The brightness at a pixel in an image can be expressed as:

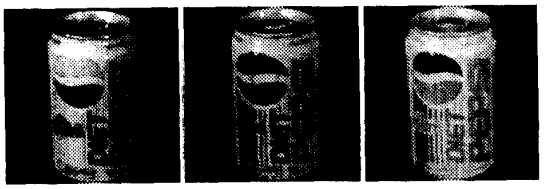
$$x = \int s(\lambda)\, h(\lambda)\, i(\lambda)\, d\lambda \quad (11)$$

where, λ is the wavelength of light, $s(\lambda)$ is the spectral response of the sensor, $h(\lambda)$ is the reflectance of the scene point corresponding to the pixel, and $i(\lambda)$ is the spectral distribution of the illumination. If the objects are illuminated using white light ($i(\lambda) = 1$), the illumination color can, in effect, be controlled using a filter with spectral response $f(\lambda)$ at the sensor end. Then, the brightness measured by a pixel can be written as:

$$x = \int s(\lambda)\, f(\lambda)\, h(\lambda)\, d\lambda \quad (12)$$

Thus, the image of an object under illumination $f(\lambda)$ can be obtained by illuminating the object with white light and using a filter with response $f(\lambda)$ in front of the sensor. Fig. 5 shows images of an object taken using three different filters

with responses $r(\lambda)$, $g(\lambda)$, and $b(\lambda)$. These response functions have their peaks close to the wavelengths that humans perceive as "red," "green," and "blue." It is interesting to note that the brightness of some of the regions on the object vary dramatically between the three images.



(a) (b) (c)

Figure 5: Images of an object obtained using three color filters under white light illumination: (a) red; (b) green; and (c) blue. The use of filters at the sensor end is equivalent to varying the color of the illumination.

The planning set in this case is obtained by varying the pose of each object, for each filter f. Each image in the planning set can be denoted as $\mathbf{x}_{r,f}^{(p)}$ where r is the rotation or pose parameter, f represents the filter or illumination color, and p is the object number. In this case, the direction of illumination is held constant. Once again, the universal eigenspace is computed, and each image in the planning set is projected to a point $\mathbf{g}_{r,f}^{(p)}$ in eigenspace. Again, consider two objects, say p and q, from the set used to compute the universal eigenspace. For each filter f, we compute parametric curves, $\mathbf{g}_f^{(p)}(\theta_1^{(p)})$ and $\mathbf{g}_f^{(q)}(\theta_1^{(q)})$, for the two objects, where $\theta_1^{(p)}$ and $\theta_1^{(q)}$ are the poses of the objects. The shortest distance between the two curves is computed as:

$$d_f^{(p,q)} = \min_{\theta_1^{(p)},\theta_1^{(q)}} \| \mathbf{g}_f^{(p)}(\theta_1^{(p)}) - \mathbf{g}_f^{(q)}(\theta_1^{(q)}) \| \quad (13)$$

The $\theta_1^{(p)}$ and $\theta_1^{(q)}$ values that produce the minimum distance $d_f^{(p,q)}$, correspond to poses of the two objects for which they appear most similar when imaged through filter f. The optimal illumination color is determined by finding the filter that maximizes the minimum distance $d_f^{(p,q)}$ between the object curves. As in Section 2.5, the above optimization is extensible to a set of P objects.

Note that we have treated the planning of source direction and source color as two separate problems. It is possible to simultaneously determine the optimal direction and color. This is done by computing manifolds of objects in eigenspace for each direction-color pair and finding the pair that maximizes the minimum distance between object manifolds.

3 Object Recognition

In this section, we describe an object recognition system that is based on the parametric eigenspace representation. We first presented this system in [6] where it was successfully demonstrated as a robust and efficient approach for recognizing a variety of complex objects. In the experimentation section, it is used to evaluate the performance of the illumination planning method described above.

Consider an image of a scene that includes one or more of the objects we have used to compute the universal eigenspace. We assume that the objects are not occluded by other objects in the scene when viewed from the sensor direction, and that image regions corresponding to objects have been segmented away from the scene image. These assumptions are valid for a variety of industrial applications, for instance, when manufactured parts pass by on an assembly line and need to be recognized.

The first step is to normalize the segmented image regions with respect to scale and brightness as described in Section 2.2. The normalization renders the recognition system invariant to imaging optics (magnification[4] and aperture) and the intensity of the illumination. An image region, after normalization, is referred to as *input image* $\mathbf{y}$.

For recognition, the average $\mathbf{c}$ of the planning set used to compute the universal eigenspace is subtracted from the input image $\mathbf{y}$. The resulting image is projected (as described in Section 2.4) to the universal eigenspace to obtain a point $\mathbf{z}$. The recognition problem then is to find the object p whose eigenspace representation (manifold in general, and curve in our case) the point $\mathbf{z}$ lies on. Here, the source direction l is known a-priori, and so are the object curves in the eigenspace for the direction l. Due to factors such as image noise, aberrations in the imaging system, and digitization effects, the point $\mathbf{z}$ may not lie exactly on an object curve. Therefore, we find the object p that gives the minimum distance $h^{(p)}$ between its curve $\mathbf{g}_l^{(p)}(\theta_1)$ and the point $\mathbf{z}$:

$$h^{(p)} = \min_{\theta_1} \| \mathbf{z} - \mathbf{g}^{(p)}(\theta_1) \| \quad (14)$$

If $h^{(p)}$ is less than a small threshold value, we conclude that the input image is of object p. The value of θ_1 that corresponds to $h^{(p)}$ represents the pose of the object in the scene. Fig. 6(a) shows an input image of the object whose parametric curve was shown in Fig. 3. In Fig. 6(b), the input image is mapped to eigenspace and is seen to lie close to the parametric curve of the object.

4 Experiments

If an object's geometry and reflectance are known a-priori, its images under different poses and illumination conditions can be synthesized using image rendering techniques such as radiosity or ray tracing. Here, we have not assumed that object models are available. Therefore, we need a mechanism that automatically varies object pose and illumination and generates image sets. Fig. 7 shows the setup we have developed for illumination planning. The object is placed on a motorized turntable and its pose is varied about a single axis, namely, the axis of rotation of the turntable. The turntable position is controlled via software and can be varied with an accuracy of about 0.1 degrees. Most objects

[4]The image projection model is assumed to be weak-perspective; orthographic projection followed by scaling.

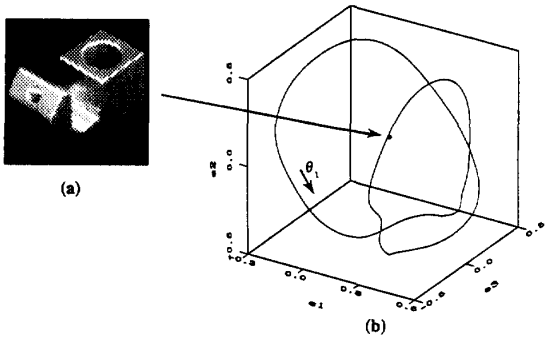


Figure 6: (a) An input image. (b) The input image is mapped to a point in eigenspace. The location of the point determines the object and its pose in the input image.

have a finite number of stable configurations when placed on a planar surface. For such objects, the turntable is adequate as it can be used to vary pose for each of the object's stable configurations.

The objects are illuminated by the ambient lighting conditions of the environment that do not vary during the acquisition of image sets. This ambient illumination is of relatively low intensity. In addition, 8 incandescent light bulbs (100 Watts each) are used to illuminate the objects from different directions. Of these only 6 sources were used since sources 1 and 8 generate strongly self-shadowed images of the objects. The light bulbs are uniformly distributed in a plane around the turntable and the angle between adjacent light bulbs is 30 degrees. These light sources are activated via software. The planning problem is to find the optimal light source among the six. We have also conducted experiments on planning illumination color. In this case, the illumination direction is fixed while three filters (red, green, and blue) are sequentially used at the sensor. For each filter, images sets are obtained for each object by varying pose. Images are sensed using a 512×480 pixel CCD camera, and are digitized using an Analogics frame-grabber board.

The experiments were conducted using three pairs of objects. These objects are shown in Fig. 8. Object pairs **A** and **B** were used for experiments on illumination direction planning, while pair **C** was used for illumination color planning. In the case of direction planning, for each of the 6 light sources, each object was placed on the turntable and images were obtained for 45 different poses (6 degree rotations of the turntable). For each of the object pairs, **A** and **B**, therefore, a planning set with 720 images was obtained. Images are automatically segmented and normalized in scale and brightness as described in Section 2.2. Each normalized image is 128×128 pixels in size. 8-dimensional universal eigenspaces were computed from the planning image sets.

The 45 pose images of each object, taken for each of the light sources, are projected onto the universal eigenspace to get a set of discrete points. These points are interpolated using a standard cubic spline interpolation algorithm [8] to obtain a parametric curve. Fig. 9(a) shows the minimum distance $d_l^{(p,q)}$ between the eigenspace curves of the two objects in pair **A**, plotted as a function of source number. Note that poses of the two objects corresponding to the minimum distance $d_l^{(p,q)}$ need not be among the ones present in the planning image set. Since the curves are obtained by interpolation, the worst case poses may lie in between the discrete poses used for planning. We see that source 6 (at 45 degrees) is found to be optimal for object pair **A**. Fig. 9(b) shows results for an object set that includes all objects in pairs **A** and **B**. Here, all 720 images of the 4 objects were used to compute the universal eigenspace and for each illumination direction 4 appearances curves were computed. As seen from Fig. 9(b), source 6 is optimal in this case also.

The optimal source direction determined by the illumination planner is meaningful only if it can be used to accomplish a vision task. We have used the correlation-based recognition system presented in Section 3 to verify the above results. For each light source, 45 test images of each object are used as inputs to the recognition system. All of these test images are different from the ones used for illumination planning; they correspond to object poses that lie in between the poses used for planning.

We define *recognition rate* as the percentage of test images for which the object in the image is correctly recognized *and* the computed pose is within 6 degrees[5] of the actual pose. Fig. 10(a) compares recognition rates produced by the optimal source 6 and the sub-optimal source 2, for object pair **A**. To test the sensitivity of the optimal source, we added white noise to the test images. The noise level is expressed in decibels of signal to noise ratio; i.e. 10 $log_{10}(S/N)$. Hence, a noise level of -10 dB corresponds to noise that is 10 times the signal. Note that the noise levels added to the test images are substantial. As noise increases, recognition rates naturally deteriorate but the optimal source 6 consistently produces higher recognition rates than source 2 (used as a non-optimal source). Fig. 10(b) shows the validity of optimal source 6 for the set including all 4 objects. These results demonstrate the robustness of the source selected by the illumination planning method to image noise.

In Fig. 10(c), the effects of segmentation error on the planning result are explored. Here, the object region is first segmented from each of the 720 test images and scale normalized to fit a 128x128 pixel image, as described in Section 2.2. Then segmentation errors are introduced in each normalized image by shifting the object region in a randomly selected direction ($+x$, $-x$, $+y$, or $-y$) by some percentage of the image dimension (128 pixels). The resulting image emulates one with segmentation error. In Fig. 10(d), the segmentation error, or percentage shift, is plotted along the horizontal axis. As the segmentation error increases, recognition rate deceases. However, the optimal source is seen

[5]This pose tolerance was selected arbitrarily. It is used to ensure that the optimal source yields the highest accuracy not only in object identification but also in pose estimation.

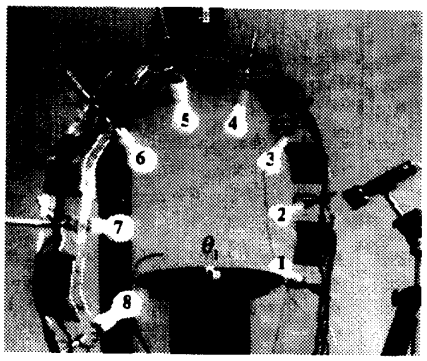


Figure 7: Device developed to obtain planning image sets. Each object is placed on the motorized turntable and illuminated from eight different directions.

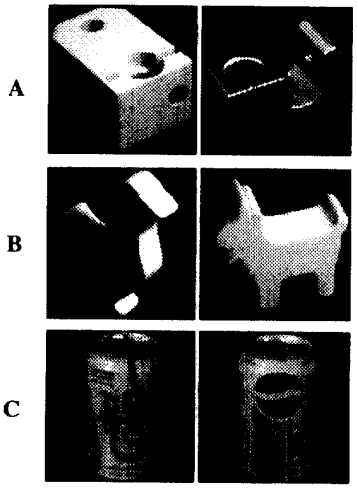


Figure 8: Three pairs of objects used to test the illumination planning method. Pairs A and B are used for source direction planning, while pair C is used for planning illumination color.

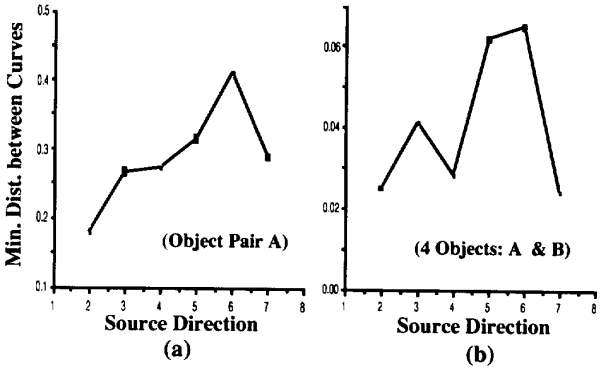


Figure 9: The minimum distance between parametric curves plotted as a function of source number for (a) object pair A, and (b) for the set including pairs A and B.

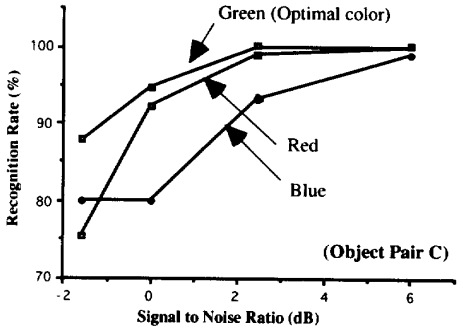


Figure 11: Recognition rates obtained using optimal (green) and sub-optimal (red and blue) illumination colors for object pair C. Even in the presence of noise, the optimal color always produces the highest recognition rate.

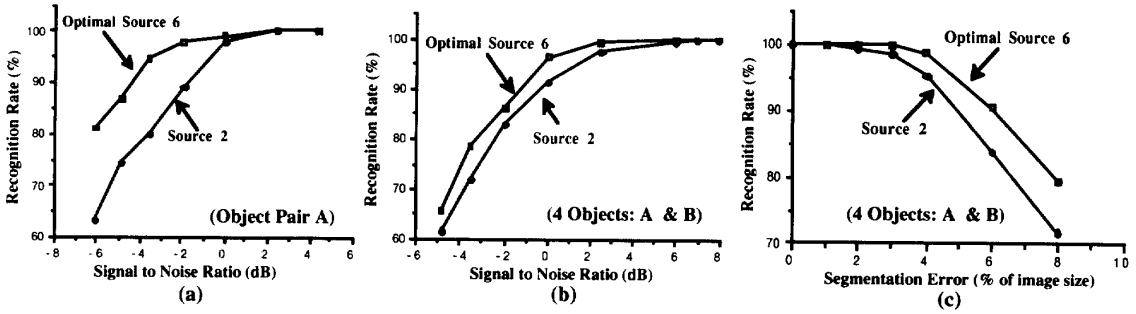


Figure 10: Recognition rates obtained using optimal and sub-optimal source directions for (a) object pair A, and (b) the set including pairs A and B. For each object a total of 45 test images were used. As the noise level in the input images is increased, recognition rate decreases but the optimal source always produces the highest recognition rate. (c) Sensitivity of recognition rate to segmentation error.

to always produce higher recognition performance than the sub-optimal source.

Figure 11 shows results on determining the optimal illumination color for object pair **C**. We have used three filters (red, green, and blue) to obtain the object image sets. The planning system determined the green filter to be optimal. Note that the green filter consistently produces the highest recognition rate when image noise is increased.

5 Discussion

In structured environments, the performance of machine vision systems can be enhanced by controlling illumination. In this paper, we have presented a method for determining illumination parameters that make a set of objects maximally different from each other in the correlation sense. The proposed approach was shown to be effective in improving the performance of a correlation-based recognition system. Such recognition systems are widely used in the industrial domain for object identification and classification. The planning methodology presented here is not geared towards the optimization of image features, a problem that has been investigated by other researchers [2], [11].

The planning technique uses samples of the objects of interest and does not require that the geometry or reflectance of the objects be known. An object could have complex geometric features, varying reflectance properties, produce specular reflections, or even interreflections. Since illumination planning is based on object appearance, none of the above effects need be analyzed in isolation. Further, the parametric eigenspace representation enables us to determine an illumination that is optimal when the poses of the objects are unknown.

Several experiments were conducted for planning both source direction as well as source color. In these experiments, a single parameter was used to represent object pose (rotation) for a given stable configuration. For certain applications, 3 degrees of freedom (DOF) may be needed to describe object pose. In such cases, for any given illumination, object appearance is represented in eigenspace as a 3 DOF manifold. This, of course, involves the acquisition of a larger number of object images for each illumination. Further, illumination itself can be described using additional parameters, including, source size, source distance, and the number of sources. The proposed method can be used to simultaneously optimize multiple parameters. The only requirement is that these parameters be varied during the acquisition of the planning image set. Clearly, for multiple parameters, acquiring image sets, computing parametric eigenspaces, and determining the optimal parameter values can be very time consuming. Therefore, the planning method may prove impractical when more than three illumination parameters need to be jointly optimized. A lesser number of parameters, however, can be easily accommodated since illumination planning is typically done off-line and only once.

References

[1] B. G. Batchelor, "A Prolog Lighting Advisor," *SPIE Proc. of Intelligent Robots and Computer Vision VIII: Systems and Applications*, B. G. Batchelor, editor, vol 1193, 1993.

[2] C. Cowan, and A. Bergman, "Determining the Camera and Light Source Location for a Visual Task," *Proc. of IEEE International Conference on Robotics and Automation*, pp. 509-514, 1989.

[3] C. Cowan, and D. Nitzan, "Automatic Placement of Vision Sensors" *18th Annual NSF Conference on Design and Mfg. Systems*, 6a. Tech, Atlanta, GA, January, 1992.

[4] H. Murase, F. Kimura, M. Yoshimura, and Y. Miyake, "An Improvement of the Auto-Correlation Matrix in Pattern Matching Method and Its Application to Handprinted 'HIRAGANA'," *Trans. IECE*, Vol. J64-D, No. 3, 1981.

[5] H. Murase and M. Lindenbaum, "Spatial Temporal Adaptive Method for Partial Eigenstructure Decomposition of Large Images," *NTT Technical Report No. 6527*, March 1992. Also *IEEE Trans. Image Processing*, (in press).

[6] H. Murase and S. K. Nayar, "Visual Learning and Recognition of 3D Objects from Appearance," *International Journal of Computer Vision*, (in press).

[7] E. Oja, *Subspace methods of Pattern Recognition*, Research Studies Press, Hertfordshire, 1983.

[8] W. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1988.

[9] S. Sakane, M. Sato and M. Kakikura, "Automatic Planning of Light Source Placement for an Active Photometric Stereo System," *Proc. of IEEE International Workshop on Intelligent Robots and Systems*, pp. 559-566, 1990.

[10] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586-591, June 1991.

[11] S. Yi, R. M. Haralick, and L. G. Shapiro, "Optimal Sensor and Light Source Positioning for Machine Vision," *Proc. of International Conference on Pattern Recognition*, 1990.