Massachusetts Institute of Technology

Artificial Intelligence Laboratory

A.I. Memo No. 746 October, 1983

Picking Parts out of a Bin

Berthold K. P. Horn and Katsushi Ikeuchi

Abstract

One of the remaining obstacles to the widespread application of industrial robots is their inability to deal with parts that are not precisely positioned. In the case of manual assembly, components are often presented in bins. Current automated systems, on the other hand, require separate feeders which present the parts with carefully controlled position and attitude. Here we show how results in machine vision provide techniques for automatically directing a mechanical manipulator to pick one object at a time out of a pile. The attitude of the object to be picked up is determined using a histogram of the orientations of visible surface patches. Surface orientation, in turn, is determined using photometric stereo applied to multiple images. These images are taken with the same camera but differing lighting. The resulting needle map, giving the orientations of surface patches, is used to create an orientation histogram which is a discrete approximation to the extended Gaussian image. This can be matched against a synthetic orientation histogram obtained from protoypical models of the objects to be manipulated. Such models may be obtained from computer aided design (CAD) databases. The method thus requires that the shape of the objects be described, but it is not restricted to particular types of objects.

Key Words: Machine Vision, Bin Picking, Shape Representation, Object Recognition, Hand Eye System, Attitude, Extended Gaussian Image.

Picking Parts out of a Bin

Berthold K. P. Horn and Katsushi Ikeuchi

Massachusetts Institute of Technology

Artificial Intelligence Laboratory

Abstract

One of the remaining obstacles to the widespread application of industrial robots is their inability to deal with parts that are not precisely positioned. In the case of manual assembly, components are often presented in bins. Current automated systems, on the other hand, require separate feeders which present the parts with carefully controlled position and attitude. Some of the methods developed recently in machine vision allow one to automatically direct a mechanical manipulator to pick one object at a time out of a pile. The attitude of the object to be picked up is determined using a histogram of the orientations of visible surface patches. Surface orientation, in turn, is determined using multiple images. These images are taken with the same camera but differing lighting. The resulting needle diagram, giving the orientations of surface patches, is used to create an orientation histogram which is characteristic for a particular object. This can be matched against an orientation histogram computed from a geometric model of the object to be manipulated. Such models may be obtained from computer aided design (CAD) databases. The method thus requires that the shape of the objects be known, but it is not restricted to objects with particular shapes. Similarly, the way in which the surface of the object reflects light must be known, but the method is not restricted to materials with particular reflecting properties.

1. Introduction

We have developed a system that will determine the position and attitude of a part in a pile of parts, using a few images taken by an electronic camera. The results can be used to direct a mechanical arm to pick up the part. The system uses stored models of the objects and can identify which of several different parts is seen. The method is not restricted to cylindrical parts or even solids of revolution. Extended light sources can be used in essentially arbitrary positions and the objects need not be ones having very special reflective properties. The system adapts to these variables by means of a calibration step involving an object of known shape. Another, different, calibration process is used to determine the transformation between the coordinate system tied to the manipulator and that of the camera. The type of sensing system described here will extend the range of application of today's industrial robots.

Mechanical manipulators are being used more and more for spot welding, machine loading, painting, deburring, seam welding and sealing. They have, however, not been utilized extensively for many other application, like assembly. One of the reasons is that today's industrial robot typically just plays back a fixed sequence of motions taught by an operator. The blind robot cannot deal with uncertainty in the positions of the parts. Feeding mechanisms and fixtures are needed to present the parts in precisely the place in which the industrial robot expects to find them.

2. The Problem

Some means of sensing the position and attitude of the objects is desirable. This information may be obtained using a system which forms an image of the objects. Electronic cameras provide a ready means of feeding a digitized image into a computer. The image plane, inside the camera, is covered by sensing elements arranged in a regular pattern . The area corresponding to a sensing element is called a picture cell. The quantized measurement of brightness in one of these elemental areas is called a grey level. The grey levels taken together form an array of numbers, which is the discrete approximation of the continuous image. Image brightness, by the way, is hard to measure accurately, so grey levels are usually quantized to only 64, 128 or perhaps 256 levels.

The problem, of course, is not how to digitize the image, or how to store it, but what to do with the information once it has been read into the computer. How can one recognize an object and determine its attitude in space using the array of grey levels produced by the camera?

Means for solving such problems, in special cases, were developed in research laboratories 10 to 15 years ago. These methods, to be described next, work well when the environment is controlled in a suitable way. In particular, there are situations in which it is possible to distinguish those points in the image which correspond to the object of interest, from those which do not. Such a segmentation, into object and "background," is usually based on differences in brightness. The result is called a binary image, since at each point it is either one (object present) or zero (object absent).

3. Binary Image Processing (*)

A few properties of the binary image, such as the area of the object region and its perimeter, are calculated readily. There may be more than one connected region in the binary image and some of these regions may have one or more holes in them. It makes sense then to calculate the Euler number, the difference between the number of objects and the number of holes. The Euler number of the capital letter "B," for example, is minus one, while it is two for the lower case letter "i." Measures such as area, perimeter and Euler number can be computed rapidly, in
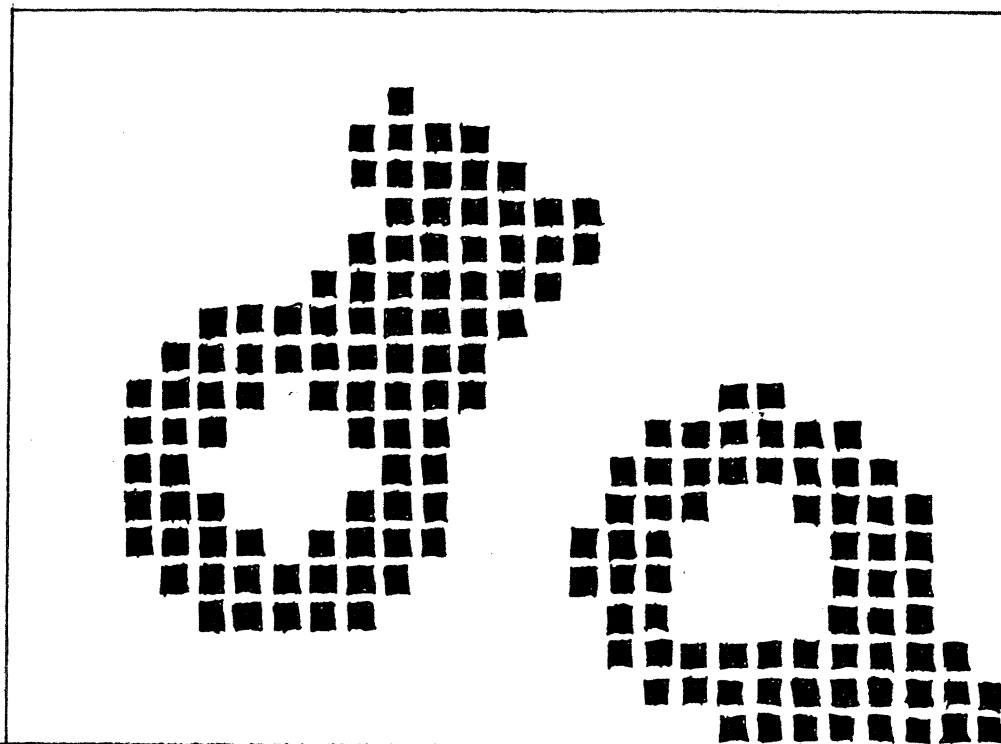
Figure 1. A binary image can be obtained by thresholding brightness values. Picture cells, arranged on a regular raster, are assigned one of the two possible values, 0 or 1, depending on whether the brightness is above or below some threshold value. The example shown is of rather low resolution. In practice one might work with perhaps 256 rows and 256 columns. Binary images are easy to digitize, store, transmit and process, but are limited in their usefulness.

---

parallel, and can be used to distinguish amongst a small number of different objects that may appear in the image.

Secondly, the position and rotation of the objects can be readily calculated using the first and second moments of the regions. The position of the object is considered to be given by the location of the center of area, while the rotation of the object in the image plane is defined by the axis of least inertia. If there is more than one region of ones in the images, the above mentioned calculations can be applied to each region separately. Naturally, the individual regions have to be labeled first. Methods for doing this in one pass over the image have been invented too.

Finally, it is possible to "grow" a binary image region, that is, add to it picture cells within a specified distance from its margin. Similarly it can be "shrunk" by growing the background. Such iterative modification techniques have proved useful in inspection, in recognizing characters and in the automatic digitization of line drawings.

The three classes of methods mentioned above are easily implemented in high speed hardware of relatively modest cost. Various clever techniques are used, such as run length coding and one dimensional projections of the image taken in a number of directions. Several vendors offer devices based on this approach. Binary image processing systems suffer from limitations however, resulting in part from the fact that all the information in a binary image is in the silhouette:

1. There must be strong contrast between the object of interest and its background (Otherwise it is hard to separate the object from the background using a simple threshold on the grey levels).

2. There should be only one object in the field of view, or, if there are several, they may not overlap or touch.

3. The object may only rotate in a plane parallel to the image plane (Otherwise the silhouette of the object changes in a complicated fashion).

As a result of these limitations many applications cannot be handled directly using binary image processing methods.

4. The Bin of Parts

In manual assembly, it is common to find components arranged in trays or bins surrounding the work station. All three conditions for the successful application of binary image processing are violated in this case. An obvious solution is to avoid jumbling the parts together in the first place, keeping them carefully oriented right from the time they are made. There is a trend to do this now, partly because of the shortcomings of present-day automation techniques. Parts may be organized on carriers or attached to pallets, so that they can be mechanically positioned without the need for sensing.

There are costs associated with this solution. The carriers and pallets must be designed and manufactured, often to tight tolerances. Pallets also typically are heavy, take up a lot of space, and may have to be redesigned when the part is modified. Often the design of the part itself must be altered to allow automatic feeding. In any case, there are still plenty of situations where limited production volume has not presented the incentive to depart from the more traditional, manual methods.

A number of attempts have been made to find mechanical solutions to this problem. In many cases, for example, it is possible to throw the parts into a vibratory bowl with carefully designed selectors, and have them emerge oriented at a feeder station. Screws and objects with cylindrical geometry are subject to this approach. Not all parts can be handled this way, however. Large or heavy parts, as well as parts with complex shapes, do not succumb to this methodology.

Attempts to equip robot arms with electromagnets or vacuum suction cups have met only with limited success. It is hard to be certain that such a device picks up exactly one object, and it is *still* necessary to reorient the object after it is

picked up. John Birk at the University of Rhode Island, developed a system using machine vision methods to pick up ground cylindrical parts. Grinding produces circumferential striations in metal, which catch the light in such a way that a bright highlight appears along the length of the object, when it is illuminated by a point source. Thresholding of image brightness values allows one to locate these lines in the image. A robot arm can then be directed to pick up a part with its gripper aligned perpendicular to the direction of the highlight. A slanted mechanical chute can be used to complete the re-orientation of the part once it is picked up. This approach, however, is limited to objects with particular shapes and surface properties.

5. Machine Vision

There has been considerable progress in machine vision since the time that the first binary image processing systems were demonstrated. The overall task of a machine vision system is the generation of a symbolic description of the three dimensional world which gave rise to the image. The form of the description will depend on the application. In our case it can be concise: the identity, position and attitude in space of the objects. In other cases it may need to be more elaborate.

In some sense, machine vision represents an inversion problem. When an image of a surface is formed, information about the distance to that surface is lost. The image is a two dimensional representation of a three dimensional world. There are a about a dozen depth cues which permit one to recover the missing third dimension from the image. If asked, most of us would think of stereo first as a method for recovering the distances to objects. We can see in depth partly because we have two eyes and so obtain images obtained from two slightly different viewpoints. This is a very effective depth cue, as long as there are contrasting features on the surface that can be matched. Also, for accuracy, the distance of the objects should not be too large compared to the separation between the two image forming systems. We know that this method works well, given the right circumstances, since almost all topographic maps are made by (manual) interpretation of pairs of aerial photographs.

At this time, there are a number of systems which automatically match points in one image to corresponding points in the other. Existing systems are however complex, expensive, slow and typically able to deal only with certain restricted types of images. Application to robotics may still be some time away.

6. Shape from Shading (*)

Another important depth cue is shading, the variation in apparent brightness with surface orientation. When we look at the picture of somebody's face in a newspaper, we cannot use stereo as a cue, yet we get a clear impression of the shape of the features. Enough in any case to help us recognize the person. The
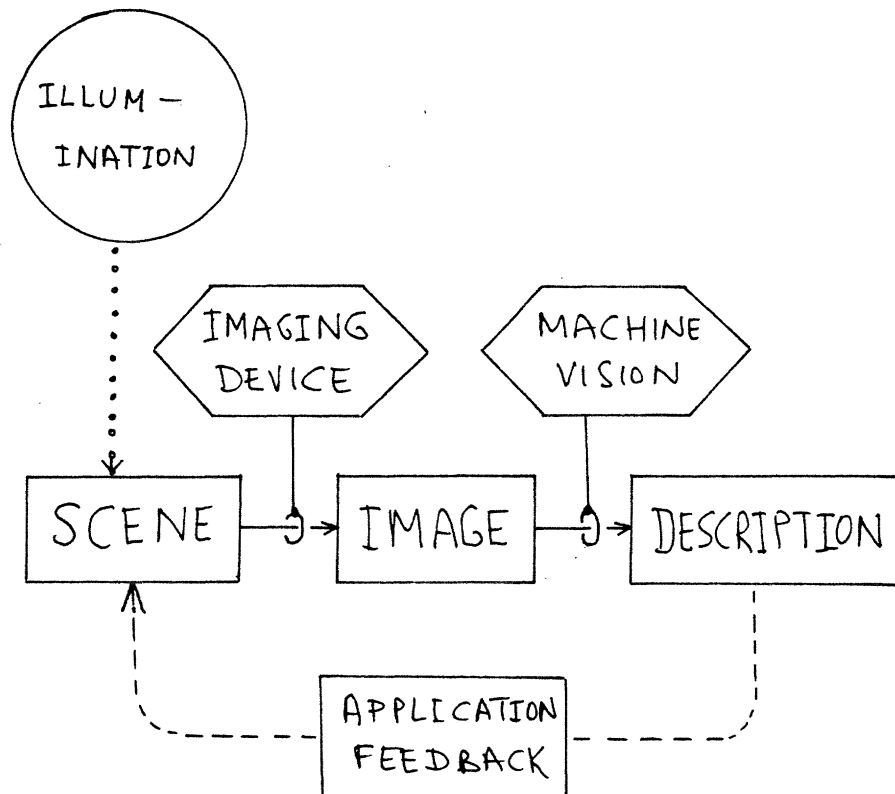
Figure 2. An imaging forming system produces a pattern of brightness in a two dimensional projection of the three-dimensional world. The task of the machine vision system is to derive a symbolic description of the scene viewed from the image. The result may be used in the intelligent interaction of the machine with its environment. If the overall system works, one may conclude that the machine vision system is performing its task. Note that it may be helpful to understand the physics of image formation when designing the machine vision system, since it performs a kind of inversion of the transformation performed by the image formation system. Also, lighting plays an important role. In an industrial setting, for example, lighting may be controlled to simplify the task of the machine vision system.

region of the picture corresponding to the face is not uniform in brightness, even though skin has essentially the same optical properties everywhere. Different parts of the face appear to have different brightness because they are oriented differently with respect to the light sources and the camera. We use this cue all the time in interpreting images, particularly those of smoothly curved objects. It has been possible to analyze this effect and develop automated methods based on the solution of a non-linear first-order partial differential equation. This so-called shape from shading method is however too complex and too slow to form the basis of a useful industrial robot sensing system.

In practical applications of machine vision, we do not necessarily have to emulate the admirable capabilities of biological vision systems. We can exploit special properties of the materials or arrange the lighting to simplify the interpretation of
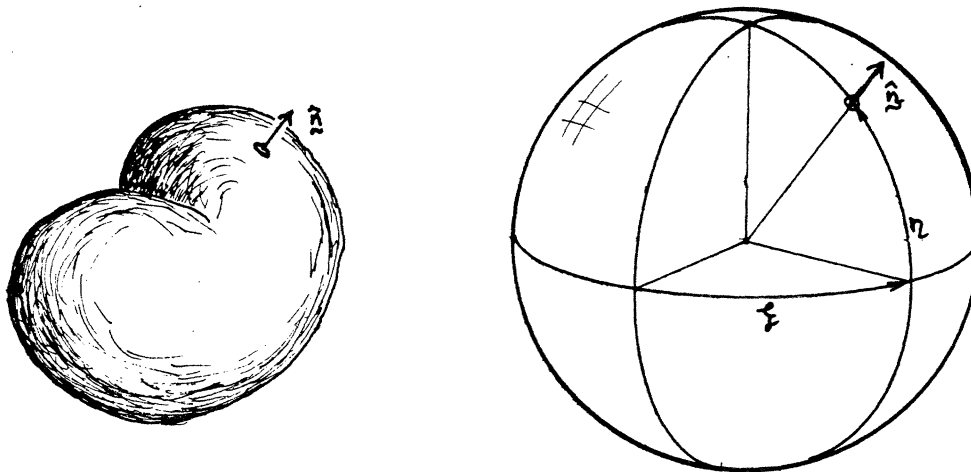
Figure 3. The orientation of a surface patch can be represented by a point on a unit sphere. One simply finds the place on the sphere which has the same surface orientation. A normal to the surface patch will be parallel to a normal of the sphere at that point. The point on the sphere can be identified using two parameters, like latitude and longitude. A sphere used in this fashion is called a Gaussian sphere. The mapping of points on the surface of an object onto a unit sphere is called the Gauss mapping.

---

the images. We will describe one such method, after considering the problem of the representation of the shape of a surface.

7. Surface Orientation

Surface orientation has two degrees of freedom. That is, it takes exactly two numbers to specify it fully. This can be seen as follows: Consider a plane surface. Now imagine a line perpendicular to this surface. To specify the orientation of the plane, we need only give the direction of this line, also called a normal to the surface. Now construct a line parallel to the normal, passing through the center of a unit sphere. The direction of this line is fully specified if we are told where it intersects the sphere. So, to each orientation of a planar surface corresponds a point on the unit sphere. We see that surface orientation has two degrees of freedom, since points on the sphere can be identified using two quantities, longitude and latitude, say.

· A unit sphere used as a means of specifying surface orientation is called a Gaussian sphere. If we are dealing with a curved surface, instead of a planar one, then surface orientation varies from point to point. We may consider the orientation at a particular point on the surface to be that of a plane tangent to the surface at that point.
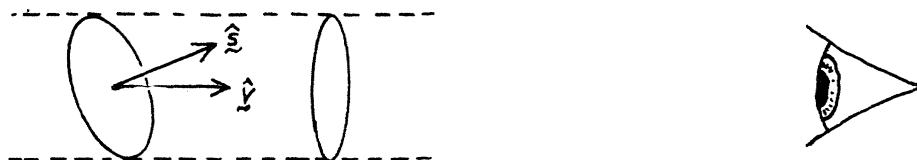
Figure 4. A surface patch viewed from a direction that is not perpendicular to the surface appears foreshortened. The apparent area is its true area times the cosine of the angle between the surface normal and the direction towards the viewer. A surface patch will intercept an amount of light proportional to its apparent area as seen from the light source. In the case of an ideal Lambertian reflector, all of this light is re-emitted. So the brightness is proportional to the cosine of the angle between the surface normal and the direction towards the light source.

8. Photometric Stereo

How can we determine the orientation of a patch of the surface of an object? We use a method here which depends only on local information and makes no assumptions about the overall shape of the object. Consider, at first, that we deal with objects which are Lambertian reflectors. An ideal Lambertian surface satisfies two conditions which fully determine is reflective properties:

1. All incident light is reflected, none is absorbed.

2. The surface appears equally bright from all viewing directions.

The amount of light which a surface patch captures depends on its apparent area as seen from the light source. A surface viewed from a direction other than along its surface normal appears foreshortened. The apparent area is the true surface area multiplied by the cosine of the angle between the viewing direction and the surface normal. Thus the amount of light falling on the surface is proportional to this quantity. We note, from the first condition stated above, that the brightness of an ideal Lambertian surface must be proportional to the cosine of the angle, usually called the incident angle. So we obtain the familiar cosine law of reflection for diffuse surfaces.

From the second condition, we note that the brightness of such a surface does *not* depend on the angle between the surface normal and the direction towards the viewer, usually called the emittance angle (This is need not be the case for surface materials which are not ideal Lambertian reflectors).

Imagine a planar patch of the ideal material illuminated by a single distant point source. Suppose the orientation of the patch is to be determined. The brightness of the surface will be proportional to the cosine of the angle between the surface normal and the incident rays. So we get a constraint on the possible surface orientations if we measure this brightness. But a single measurement is not sufficient to determine the orientation *uniquely*, because many lines make the same
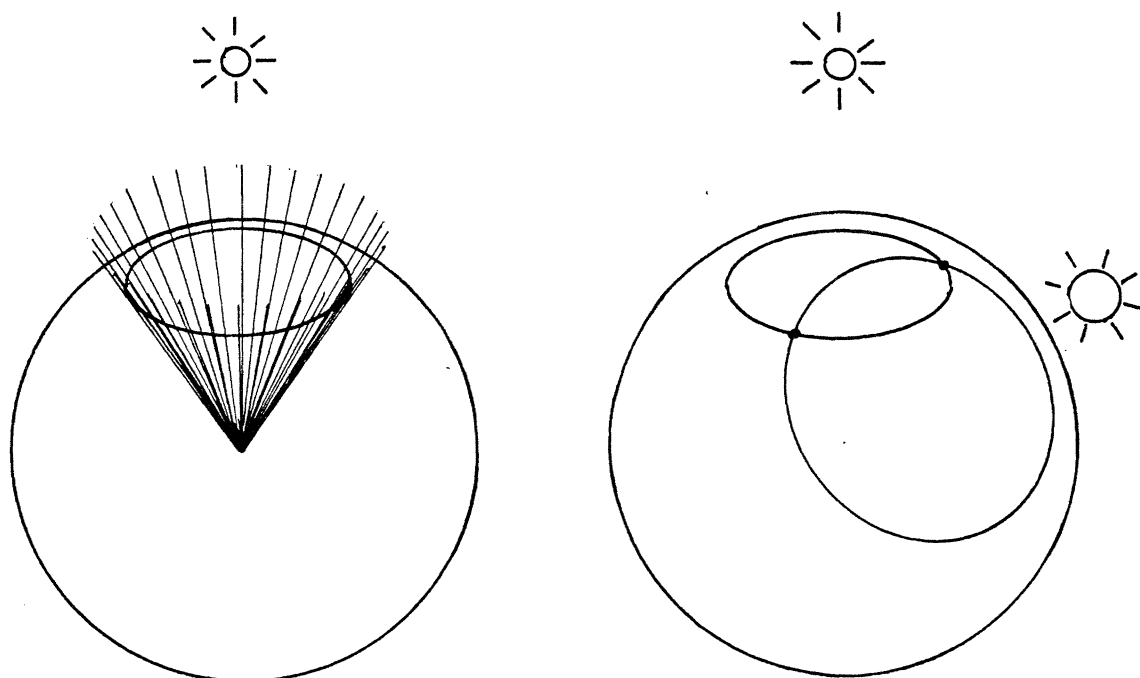
Figure 5. A single measurement of brightness constrains the surface normal to lie at a fixed angle from the direction towards the light source. The locus of all directions satisfying this constraint is a cone, whose axis points towards the light source. The intersection of this cone with the surface of the unit sphere is a small circle. The orientation of the surface patch must correspond to one of the points on this small circle. It is clear, however, that a single measurement does not provide enough information to uniquely determine the actual surface orientation. A second measurement, using a different light source, produces additional constraint. The surface orientation must correspond to one of the points where the two small circles intersect.

---

angle with the direction of the incident rays. The locus of all these lines is a cone, with axis pointing towards the point source. The normal of the surface must lie on this cone. We note that in terms of the Gaussian sphere, the possible orientations lie on a small circle, which is the intersection of the cone and the sphere. The point at the center of this circle corresponds to the orientation of a surface patch which lies perpendicular to the incident light rays.

If we now repeat the experiment with a second distant point source, we get a second constraint on possible surface orientations. The orientation has to lie on a second, different, small circle. Again, we find that the size of the circle depends on the observed brightness and the center of this circle corresponds to the direction of the second light source. The actual surface orientation must satisfy both constraints and thus lies at the intersection of these circles.

This all makes eminent sense if we remember that surface orientation has two
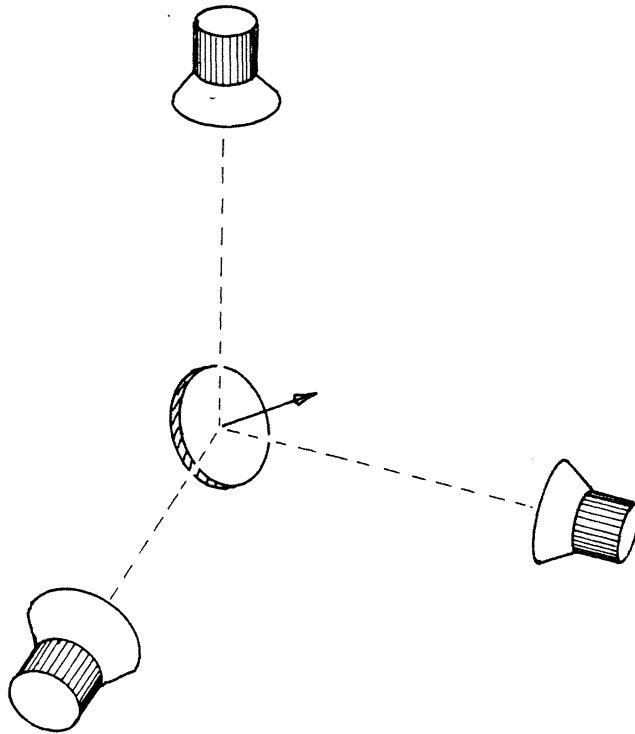
Figure 6. Three measurements of surface brightness can be obtained using three light sources. For each of the three image measurements, a different light source is turned on. Equivalently, three colored lights and a color camera can be used. In the case of a grey Lambertian surface, each measurement provides the product of the albedo and the cosine of the angle between the surface normal and the direction towards one of the light sources. The surface orientation and the albedo can be recovered easily from the three measurements. In practice, of course, one does not usually encounter surfaces with simple reflecting properties. It is also better to use extended sources instead of point sources in order to extend the range of measurement. Under these circumstances a closed form solution is no longer feasible.

---

degrees of freedom. We expect it would take two measurements to provide enough constraint to pin these down. A final difficulty is that the two circles typically intersect in *two* points instead of just one. Thus there is a remaining ambiguity in the determination of the surface orientation. We could use a third point source as a probe to obtain a third brightness measurement. This solves the problem, but constitutes overkill, since all we really need is *one bit* of information more.

If we are going to make a third measurement, we may as well use it to determine another parameter of interest. To illustrate this idea, consider a "grey" Lambertian surface. This is a surface which absorbs some of the incident light, re-emitting only a fraction, which we will call the albedo. In other respects it behaves just like the ideal Lambertian surface.

In this case, brightness is the product of the albedo and the cosine of the incident angle. Each of the three brightness measurements provides us with one equation. We have three unknowns, the albedo and the two parameters of orientation. The problem can be recast in the form of three linear equations in three unknowns. It is well known that such a system of equations has a unique solution, provided that the equations are *linearly independent*. The system of equations is dependent if, and only if, the three light sources and the object lie in a plane. In this case, one of the three measurements is just a linear combination of the other two.

Here we have exploited the redundancy provided by a third measurement to derive information about surface properties, such as albedo. If we happen to know that the surface has uniform albedo, we can instead use the extra information as a check.

Note that the brightness of a surface patch depends on its orientation, *not* its position (Provided that the light sources and the viewer are far away). A smoothly curved surface can be thought of as divided into many small patches, each of which is approximately planar. The three measurements are made for each patch. Conveniently, these measurements can be made for all surface patches *at once* by taking three images. A different light source is powered up for each image. Alternatively, one can use three colored light sources and extract the three images from the signals produced by a color camera. This is faster, but requires a more expensive camera. Also note that this last approach will not work if the surface consists of patches of different *colors*.

What we have just described is a simple example of the photometric stereo method. Note that we cannot expect to determine the surface orientation with high precision, since the individual grey levels are noisy. In practice we may be able to determine the direction of the surface normals to within about 5° or 10°. This is not a serious problem, however, since estimation of the attitude of an object is based on information about *many* surface normals.

9. Generalizations (*)

Note that there is a problem when the surface is inclined so far that it is not visible from one of the light sources. Basically, one measurement is missing when a surface is self-shadowed, and so the method only works for the range of orientations which correspond to surface patches visible form all three sources. This range can be made large by moving the light sources close together. It should be obvious though that accuracy is compromised this way. In the extreme case, for example, when the light sources have all been moved to the same place, all three measurements are the same. There is thus a trade-off between accuracy and range.

The problem can be ameliorated by using extended sources instead of point sources. Extended light sources have other desirable features. Many surfaces, for example, in addition to a diffuse component of reflection, have a glossy reflection component. When illuminated by a point source, disturbing highlights appear,

which are smeared out virtual images of the source. In the extreme case of a perfectly specular surface one cannot use a point source at all, since it creates only an isolated virtual image. These disturbing highlights can be spread out by means of an extended source.

Real surfaces generally do not behave like ideal Lambertian reflectors. In practice then the photometric stereo method has to be able to deal with extended light sources and arbitrary surface reflectance properties. The above departures from our ideal model make it unreasonable now to look for a closed form solution to the three equations for brightness corresponding to the three lighting conditions.

10. Calibration Object

It is much more convenient to use a numerical solution, based on a lookup table. The idea is to employ a calibration object of known shape, as for example, a sphere. Images of the sphere are taken under the same lighting conditions to be used later in finding the position and attitude of the objects. In the case of the sphere, the surface normals are particularly easy to calculate: At a particular point on the sphere, the normal is parallel to the radius. The position and size of the disc which is the image of the sphere is easily determined from the brightness pattern in the image. It is then possible to calculate which point on the sphere each picture cell corresponds to and what the normal is there. The grey levels at this picture cell in the three images are then determined. This experiment thus provides us with a mapping from surface orientation to brightness triples (or color).

What we need, however, is just the inverse: A mapping from brightness triples to surface orientation. The experimental data can be numerically "inverted" and a three dimensional lookup table developed which allows one to efficiently determine surface orientation. To use the table, the three brightness measurements from a point in the image of an unknown object are quantized. That is, each one is allocated to an interval corresponding to an incremental range in the table. The three numbers obtained are used as indices into the array. The entry located in this fashion contains the sought after surface orientation. The lookup table need not be especially large, in practice, 16 by 16 by 16 may be quite adequate, for example.

Note that the calculation of surface orientation is always very fast, involving nothing more than looking up the result in a table. This is quite independent of how complicated the surface reflectance properties are, and how strange the arrangement of light sources.

Large parts of the lookup table are blank, corresponding to "impossible" combinations of brightness measurements. This follows from the fact that surface orientation has only two degrees of freedom, and the table has three dimensions. If we find the brightness triples for all possible orientations, we only explore a two dimensional surface in the three dimensional space of possible brightness triples. We could fill the table completely by introducing another parameter, like albedo as suggested above. Alternatively, we may exploit the redundancy
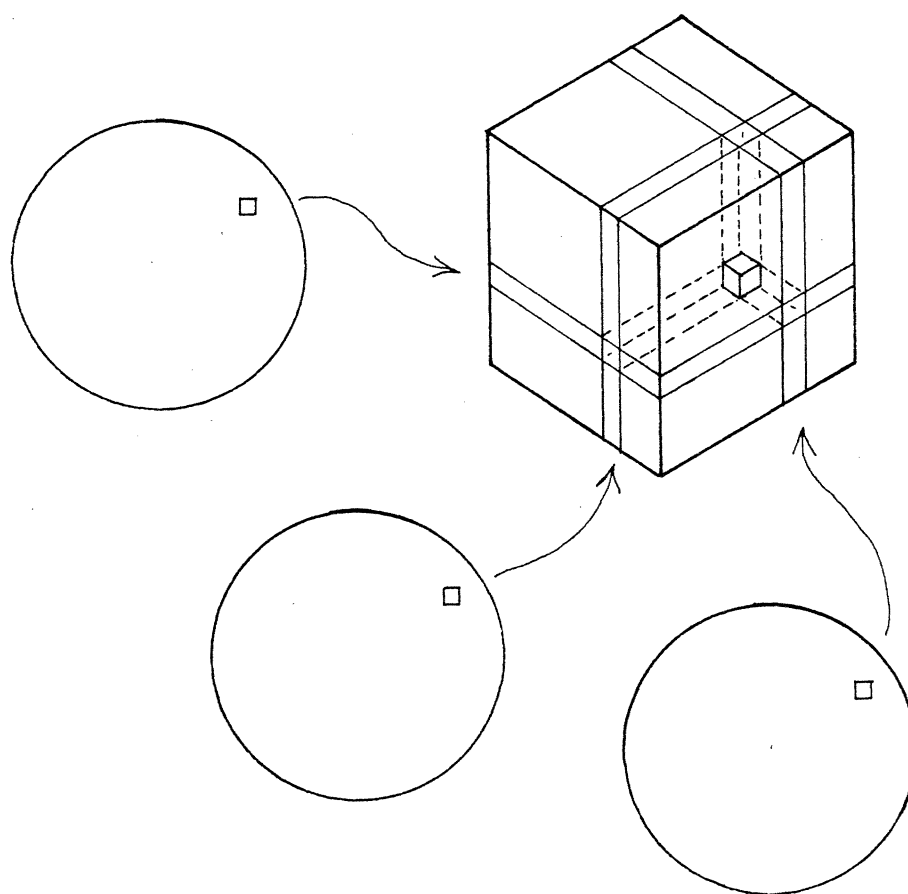
Figure 7. Images taken of the calibration object provide the transformation from surface orientation to brightness triples. For each picture cell, brightness is measured in three images taken under three different lighting conditions. The surface orientation at a patch corresponding to a particular picture cell can be computed from the known shape of the calibration object. The lookup table employed by the photometric stereo method is built by inverting the relationship between orientation and brightness: This three dimensional table is addressed using quantized values of brightness and contains the corresponding surface orientation.

---

provided by three images in another way. The blank areas of the table can help us detect shadowing and mutual illumination, since these effects produce "impossible" brightness combinations.

11. Segmentation

One of the hardest problems in machine vision is the segmentation of an image into regions corresponding to different objects. Only when this is done can one apply the techniques used to recognize an object and to determine its attitude in
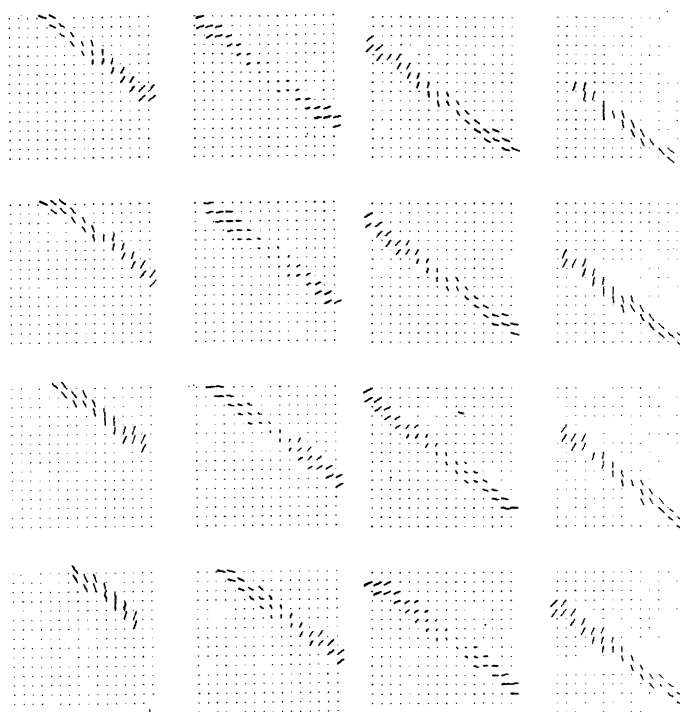
Figure 8. The lookup table can be dissected into layers, and each layer displayed in the form of a needle diagram. The short lines indicate surface orientations. The direction of each line corresponds to the direction of steepest descent on the surface. The length of the line corresponds to the inclination of the surface. Dots indicate "impossible" brightness combinations, triples which do not correspond to any surface orientation. These typically are found only when there is shadowing or mutual illumination.

---

space. One can employ several methods to help ensure accurate segmentation of the image.

First of all, objects cast shadows on one another. The result is that some points on the shadowed object have brightness readings different from what they would have been if there was no shadowing. One must detect this condition lest it lead to incorrect estimates of surface orientation. A crude way of detecting shadows is to use thresholds on each of the three brightness measurements. Note, by the way, that objects near the top of the pile, those of most interest to us, will typically not suffer from shadowing.

Secondly, mutual illumination, or interflection, occurs where objects of high albedo face each other. Amplification of illumination occurs as light is reflected back and forth. Again, we find brightness combinations that would not occur if the object was only illuminated directly by the source. Mutual illumination should be detected as well, in order to avoid incorrect estimates of surface orientation. Fortunately, this problem tends to occur near the edges of objects and the boundaries where
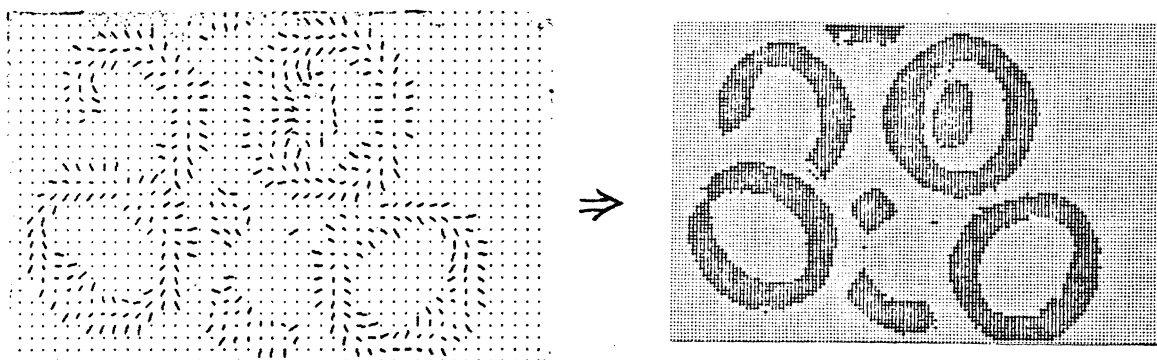
Figure 9. The image must be segmented before properties of an image region corresponding to a particular object can be computed. Photometric stereo is used to obtain a needle diagram of the whole image. A binary image is developed from this using several heuristics. First of all, picture cells at which illegal brightness combinations where found are marked as belonging to the background. This removes points which were shadowed or subject to mutual illumination. A number of heuristics can be employed to improve the robustness of this procedure. In the case of objects with smoothly curved surfaces, for example, one can reject points where the surface inclination is high and points where there is discontinuity in surface orientation. The binary image shows the remaining regions, which are now labeled and analyzed further.

---

objects obscure one another.

To obtain robust segmentation we mark image points based on four notions:

1. Low grey levels in at least one image suggest shadowing of one object by another.

2. Combinations of grey levels not found in the look up table are usually due to the effects of mutual illumination.

3. Discontinuities in surface orientation most often occur where one object obscures another.

4. High surface inclination occurs near the occluding boundaries where one object obscures another.

The points so marked form "moats" around the images of the objects, isolating them from each other. The remaining connected regions in the image can then be analyzed further. This segmentation method is robust, but depends to some extend on the properties of smoothly curved objects. Somewhat different criteria would be appropriate, for example, for objects with planar faces, like children's toy blocks.

The segmentation method we use is quite aggressive, in order to be robust. So, for example, regions of the object where the surface normal is inclined more than 45° with respect to the direction to the camera are allocated to the background. If only what remains was used in further processing, the position and attitude of the object would not be found accurately. For this reason, the region allocated to an

objects is "grown" again, once segmentation has been accomplished, to encompass as much useful data as possible.

In some cases an object which is highly inclined with respect to the viewer may get broken up because of this approach. In our case this not a serious problem, since objects which are highly inclined are difficult to pick up in any case. It is better to concentrate on the others.

12. The Needle Diagram

The normals are found at points on the surface corresponding to the picture cells into which the image is been divided. Consider placing a short surface normal at each of these points on the object. If we take a picture of the result we obtain lines in the image corresponding to the projections of the normals. These lines appear short in areas where the normals point more or less towards the viewer. They appear longer where the surface is tilted. The direction of these lines gives us the direction in which the surface is tilted: The lines point in the direction of steepest descent. The resulting figure is called a needle diagram. It is one way of representing the information obtained using photometric stereo.

The needle diagram describes the shape of the surface. How can it be used in recognizing an object and determining its attitude in space? Curiously, we can discard the information on *where* a surface normal occurs, retaining only its direction. Essentially, we build a histogram of surface patch orientations. This is a quantized version of the so-called extended Gaussian image (EGI), which will be introduced next. In effect, one decouples the problem of determining the *attitude* of the object from that of determining its *position*.

13. The Gaussian Image

First, consider a particular mapping from points on a smoothly curved, convex object onto a unit sphere. In the so called Gaussian image, a point on the object is associated with that point on the sphere which has the same surface orientation. We have already seen this mapping earlier, when we used latitude and longitude on a sphere to specify the direction of a normal to a surface patch. If the object has positive curvature everywhere, like a football for example, then there is only one point which has a given surface orientation. In this case, the mapping from the object to the sphere is invertible, that is, we can find a *unique* point on the object corresponding to a particular point on the Gaussian sphere. The Gaussian image can be used to transfer information given on the surface of an object onto the surface of a sphere.

The earth, for example, is not perfectly spherical, having a shorter "diameter" measured between the poles than between opposite points on the equator. The surface of the earth can be mapped onto the surface of a perfect sphere using the Gaussian image. Cartographers may then project the surface of this *ideal* sphere in one of several ways onto a plane to provide us with maps that can be printed on flat sheets of paper.
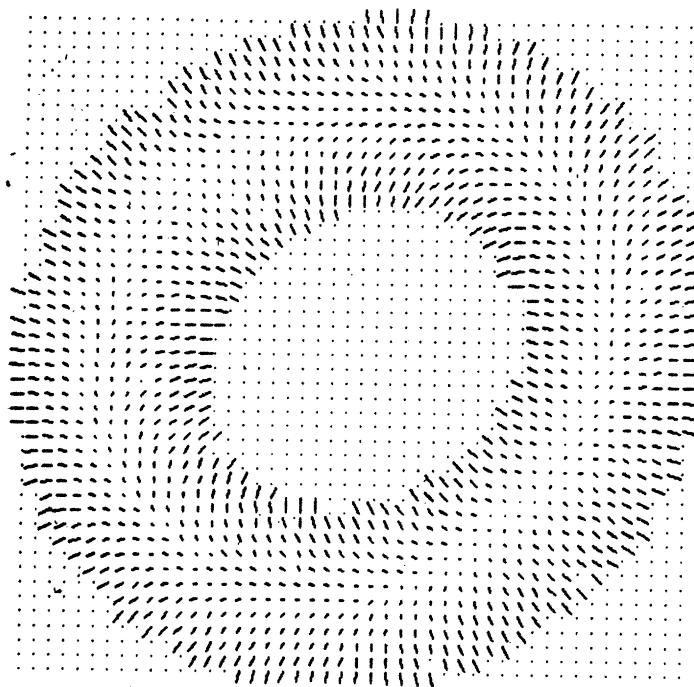
Figure 10. The shape of a surface can be represented by a needle diagram. It gives the orientation of surface patches on a regular raster. The result can be illustrated by imagining the surface covered with needles sticking out perpendicularly to the surface. The projection in the image plane of a surface normal is a line. The longer the line, the more steeply the surface is inclined. Also, the direction of the line indicates the direction of steepest descent. Shown is the needle diagram of a toroidal object.

14. Gaussian Curvature

So far we have considered the image of a particular point on the surface. If we consider the images of all points in a *patch*, we will get a corresponding patch on the surface of the Gaussian sphere. The surface normals of the points in the patch will point in widely differing directions if the surface is highly curved. Correspondingly, the patch on the sphere will be large. Conversely, if the surface is almost planar, neighboring normals will point in almost the same direction and the patch on the sphere will be small.

This suggests an intuitively satisfying definition of curvature. The Gaussian curvature is defined as the ratio of the areas of the patch on the sphere to that on the object. The reader can easily verify that the Gaussian curvature of a sphere is everywhere the same, namely one over its radius squared. The Gaussian curvature of a cylinder on the other hand is zero, since any patch on it maps into a portion of a great circle on the sphere. This is because all points along a line parallel to the axis of the cylinder have the same surface orientation.
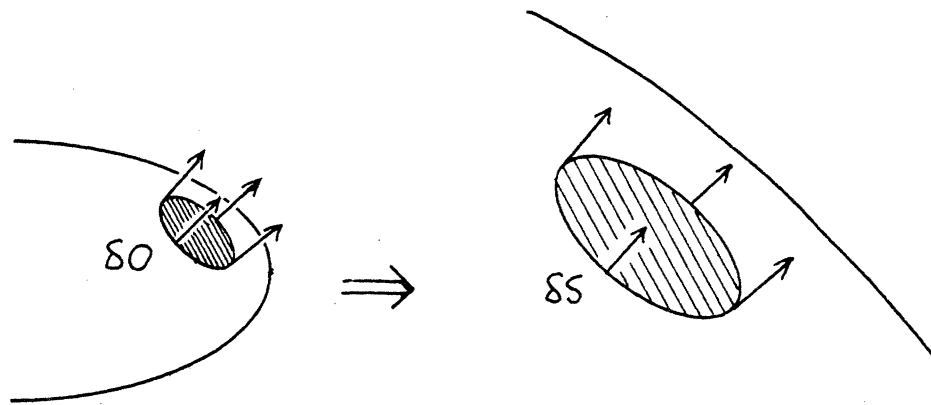
Figure 11. A patch on the object maps into a patch on the Gaussian sphere. The patch on the sphere will be large when the corresponding part of the surface of the object is strongly curved. Conversely, it will be small it the surface is almost flat. The ratio of the area of the patch on the sphere to that of the patch on the object becomes the Gaussian curvature, as the patches are shrunk.

15. The Extended Gaussian Image

The Gaussian image can be used to map *any* information which is given on the original surface onto the unit sphere. We now introduce a particular mapping called the extended Gaussian Image (EGI). It is convenient to think of the EGI in terms of a mass distribution on the surface of the Gaussian sphere. Imagine first that the surface of the original object is covered with a material which has unit density (mass per unit area). The material from a patch on the object is then spread onto the corresponding patch on the sphere. The density on the sphere will be low in areas which correspond to parts of the object which have high curvature. Conversely, the density will be high in areas which correspond to parts which are nearly planar.

In fact, the density is just equal to the inverse of the Gaussian curvature. The EGI, in the case of a convex object, is the Gaussian image of the inverse of the Gaussian curvature. The reason we choose to define it this way, is that it allows us to estimate a discrete approximation of the EGI just by counting how many surface normals point into cells on the Gaussian sphere, as will be shown.

The shape of a surface can be given by means of parametric formulae. The Gaussian curvature can be computed in terms of the first and second partial derivatives of these formulae. We completely side-step the need to estimate these derivatives by using the *inverse* of Gaussian curvature and the definition of curvature in terms of areas of corresponding patches. This is important, because it is unlikely that derivatives of the somewhat uncertain surface orientation information would be very reliable.
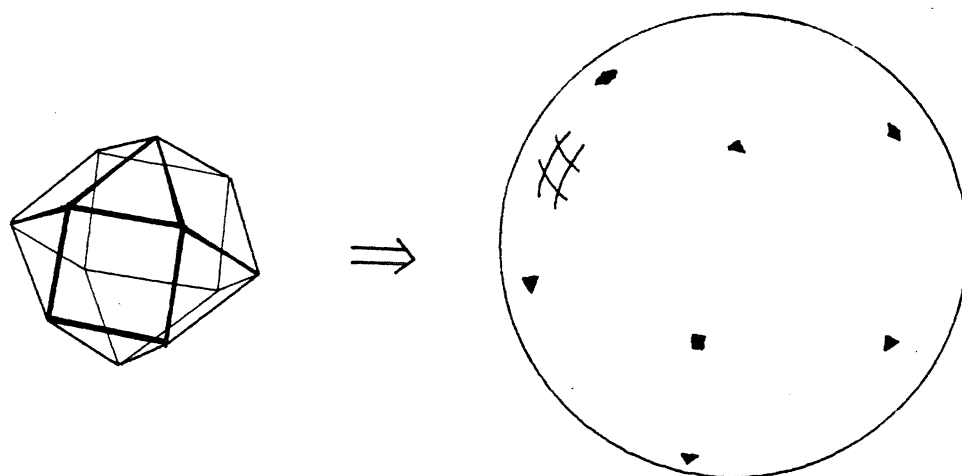
Figure 12. The extended Gaussian image of a polyhedral object is a distribution of point masses on the sphere. The position of the points is determined by the orientation of the faces of the polyhedron, while the masses are equal to the corresponding areas. For clarity only points lying on the visible hemisphere of the Gaussian sphere are shown.

Polyhedral objects have planar faces of zero Gaussian curvature. What then is the EGI of such an object? Using our idea of spreading mass from a given surface patch onto the corresponding patch on the Gaussian sphere, we see that the EGI of a polyhedral object is just a collection of point masses. Corresponding to each face, there is a mass equal to the area of the face at the point where a line parallel to the normal of that face intersects the sphere.

16. The Orientation Histogram

We can estimate the EGI numerically from the experimental data contained in a needle diagram. First of all, we divide up the surface of the object into patches corresponding to picture cells. We know the surface orientation of each of these patches and so can place a mass at the appropriate place on the sphere. The mass is equal to the surface area of the patch. We just have to remember that, because of foreshortening, the areas of these patches on the surface are not all equal. That is, patches which are inclined a lot with respect to the direction towards the imaging system are larger than those which are perpendicular to that direction.

To tally up the result, we divide the surface of the Gaussian sphere into cells. This is called a tesselation of the sphere. One can associate a mass with each cell of the tesselation, equal to the total area of the surface patches which have orientations falling within the range of orientations belonging to the cell. We call the result an orientation histogram, because it tells us how much of the surface is oriented in various directions. In the limit, as we make the sizes of the cells
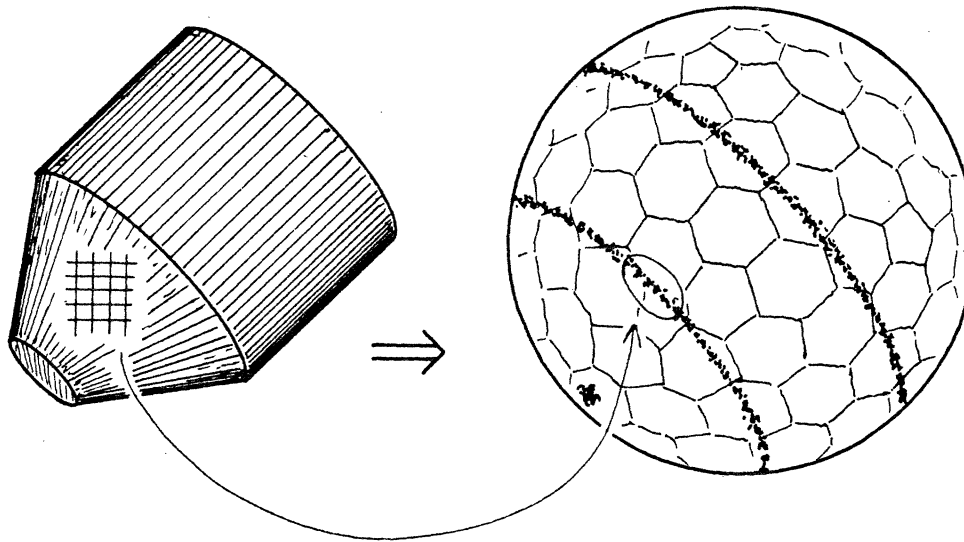
Figure 13. The extended Gaussian image (EGI) of an object can be estimated using the known orientation of surface patches corresponding to picture cells. A point mass is placed on the Gaussian sphere corresponding to every surface patch. The position on the sphere is determined by the orientation, while the mass equals the actual area of the surface patch. In order to represent this information conveniently in the computer, the sphere is divided up into cells, and the total mass determined for each cell. The discrete approximation of the EGI is called the orientation histogram.

---

smaller and smaller, at the same time also dividing the image more and more finely, the orientation histogram becomes the extended Gaussian image. It should now be clear why we chose to define the EGI the way we did.

The orientation histogram can be represented graphically in a number of ways. One can show a sphere with a normal vector for each cell of length proportional to the mass in that cell. This is called a spike model. Another way, if a grey level display is available, is to show a sphere with brightness in each cell proportional to the mass in that cell. The sphere can be projected onto the display surface in a number of ways, as, for example, orthographically. A slightly better display is obtained if the sphere is projected stereographically, since the angles between cell edges are preserved in this projection and it is possible to show more than one hemisphere at once.

17. Properties of the Extended Gaussian Image

At this point we may take note of some of the properties of the EGI. First of all, the mass of the whole EGI just equals the surface area of the whole object. This follows directly from the definition of the EGI.
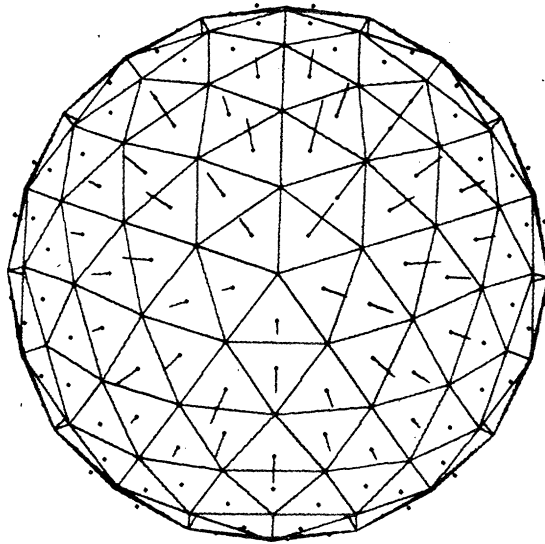
Figure 14. An orientation histogram can be shown in the form of a tesselated sphere with perpendicular spikes drawn on each cell of length proportional to the total mass in that cell. The result is called a spike model.

Next, consider the apparent cross-sectional area of the object when viewed from a particular direction. As noted before, a surface patch will appear foreshortened if viewed from a direction other than one parallel to its normal. The apparent area is the actual area times the cosine of the angle between the surface normal and the direction towards the viewer. The cross-sectional area is just the sum of all of these foreshortened patch areas. Now imagine looking at the object from the opposite direction. The silhouette of the object is mirror reversed, but the apparent cross-sectional area should be the same. This must hold for all possible directions.

Suppose now that we cut the Gaussian sphere into two using a plane at right angles to the given viewing direction. All visible surface patches correspond to points in one hemisphere. These are the patches with surface normals which make an angle of less than 90° with the direction towards the viewer. Let us call this the visible hemisphere. Surface patches corresponding to points in the other hemisphere are turned away from the viewer.

The first moment of a mass on the surface of the sphere, relative to the dividing plane, is just the product of that mass and the perpendicular distance of the mass from the plane. This distance, in turn, is equal to the cosine of the angle between the radius and the direction towards the viewer. It follows that the first moment of the mass distribution in the visible hemisphere is just equal in magnitude to the cross-sectional area of the object! Since the cross-sectional area is the same when the object is viewed from the opposite direction, we conclude that the first moments of two complementary hemispheres are equal in magnitude.

They have opposite signs, however, since the masses are on opposite sides of
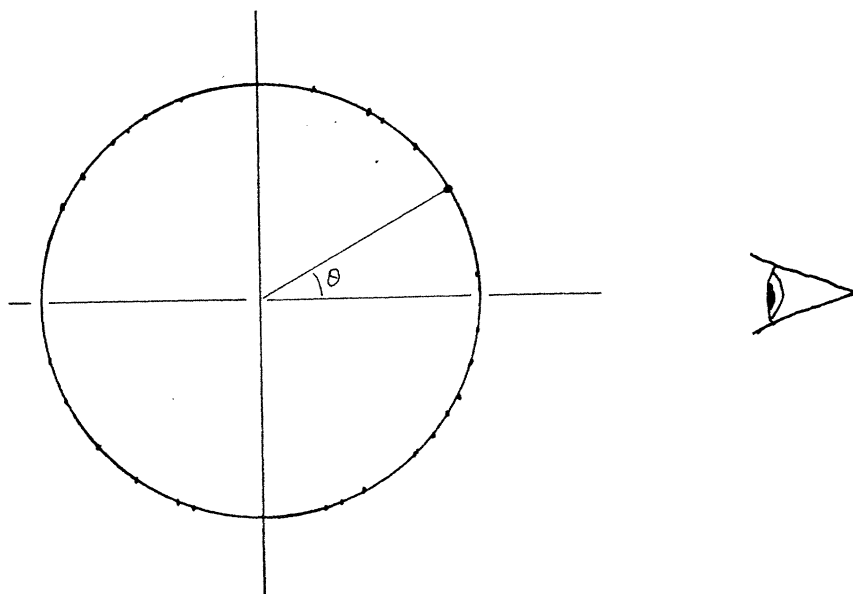
Figure 15. The cross-sectional area of an object can be obtained by adding up the apparent areas of all visible surface patches. The apparent area is the product of the actual area and the cosine of the angle between the surface normal and the direction towards the viewer. Now, the moment about a plane through the center of the sphere can be found by summing the product of the masses on the surface and their perpendicular distance from the plane. This distance equals the cosine of the angle between a line perpendicular to the plane and a radius to the mass. Thus the moment of the visible hemisphere is equal to the apparent cross-sectional area of the object! Since the object has the same apparent area when viewed from the opposite direction, the moments of the opposite hemispheres must be equal in magnitude. The moment of the mass distribution on the whole sphere then is zero. If this is to be true for *all* choices of viewing direction, the center of mass of the extended Gaussian image must be at the origin.

---

the dividing plane. The first moment of the *whole* EGI is the sum of the first moments of the two complementary hemispheres. This sum is zero. It follows from the above, that the center of mass of the EGI is on the dividing plane. Since this has to be true for *all* dividing planes, we conclude that the center of mass of the EGI is at the center of the sphere.

An even more powerful result was derived by Minkowski in 1897. He first showed that the areas and orientations of the faces of a closed polyhedron have to satisfy the condition given above. But then he went on to prove that there is only *one* convex polyhedron which has faces with the given areas and orientations. In our terminology, no two convex polyhedra have the same EGI. He showed this in an indirect way, by noting that the convex object minimizes the integral of the product of surface patch area with distance of the patch from the origin, subject to the constraint that the volume is fixed. The object is uniquely determined since there is only one global minimum. While Minkowski's proof is not constructive, it

has been used recently, by James Little of the University of British Columbia, in deriving an iterative reconstruction method for the polyhedral case.

The result was extended later to convex, smoothly curved objects. It was shown that there is a unique convex object corresponding to an EGI with center of mass at the center of the sphere. It may be thought that this result restricts our method to convex objects, since a given EGI is shared by many, an infinite number in fact, of non-convex objects. This is not a problem, however, since it is very unlikely that two objects found in a typical application have the same EGI. There are, however, other problems with non-convex objects, which will be addressed later.

18. Tesselation of the Gaussian Sphere

How do we divide the Gaussian sphere into cells to be used in accumulating the orientation histogram? Ideally the cells should satisfy the following criteria:

1. They should all have the same area.

2. They should be well "rounded."

3. They should all have the same shape.

4. Each cell should map onto another cell for some set of rotations of the sphere.

It is possible to satisfy these criteria if the sphere is to be covered with only a few cells. We can use the tesselations produced by projecting the regular solids onto the sphere. These give us six cells for the cube and twelve cells for the dodecahedron, for example (The tetrahedron, octahedron and icosahedron are less suitable, since they do not lead to rounded cells). The cells in each case have the same shape and area. The projection of the dodecahedron even leads to well rounded cells. Also, the cells map into one another for a finite number of rotations. In the case of the dodecahedron and the icosahedron this group of rotations has 60 elements.

Before we go any further, let us see how one might calculate which cell a particular surface normal belongs to. It turns out that the edges between cells are portions of great circles of equal distance from the centers of the cells. The centers of the cells in turn are the vertices of the dual of the given polyhedron. Thus all we need is a list of unit vectors pointing in the direction of the vertices of the dual. We assign the unknown unit vector to the cell for which the dot-product is largest.

Unfortunately, even 20 cells is not good enough, particularly if we keep in mind that the visible hemisphere is covered by only 10 of these! It helps then to look at *semi-regular* solids. Semi-regular polyhedra differ from regular ones in that more than one type of regular polygon may be used to construct the surface. The edges are still all the same length however. Combining pentagons and hexagons, for example, we obtain the truncated icosahedron. It has 12 pentagonal cells and 20 hexagonal ones. This is the tesselation of the soccer ball. It has the advantage over the icosahedron that its cells are fairly rounded.
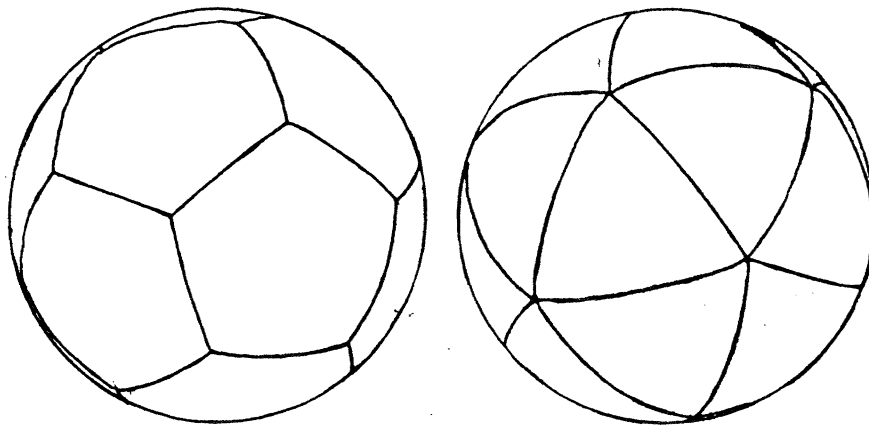
Figure 16. One way to tesselate the sphere is to project a regular polyhedron placed at the center of the sphere onto its surface. A regular dodecahedron leads to a tesselation into twelve pentagonal cells, while a regular icosahedron leads to twenty triangular cells. The resulting cells are curvilinear polygons whose sides are portions of great circles of the sphere.
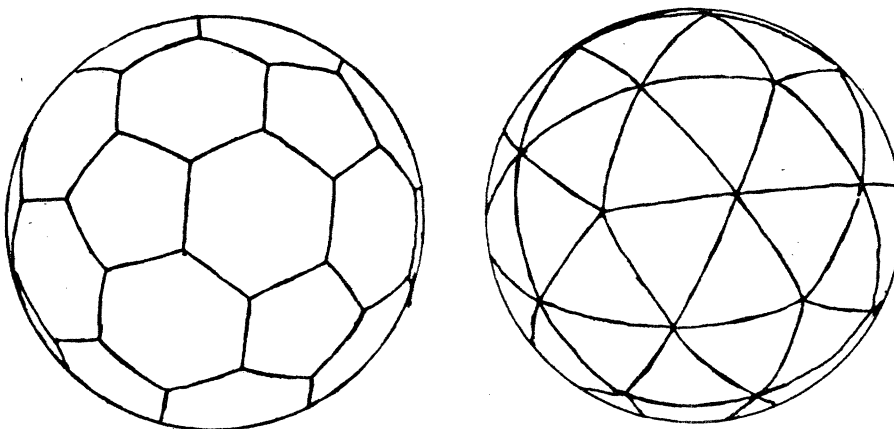

Figure 17. The tesselation used in the construction of the soccer ball is obtained by projecting a semiregular polyhedron, the truncated icosahedron, onto the sphere. It has 32 cells. Another useful tesselation is obtained by projecting the Pentakis dodecahedron which is made by dividing each pentagon of the dodecahedron into five equal triangles. It has 60 equal (but not regular) faces. If each of these triangular faces is further divided into four smaller triangles, one obtains a frequency two geodesic dome with 240 cells. This tesselation was used for the figure of the spike model of the orientation histogram.

19. Geodesic Domes

To get still finer tesselations, we may use geodesic domes. To construct such a dome, one starts with a regular polyhedron and divides its faces into triangles

(unless, of course, they are already triangular). In this way, for example, we get the Pentakis dodecahedron, with 60 faces, from the dodecahedron. Assigning unit vectors to cells is particularly easy in this case. We just need to know which cell of the dodecahedron had the *second* nearest center to the unknown in order to assign it to one of the five triangular cells into which a particular cell of the dodecahedron has been divided. Little extra work is involved since we had to compute the required dot-products already to determine the cell with the *nearest* center.

In a still finer geodesic dome, assignment of an unknown to a cell can be done efficiently using stepwise refinement. This is possible because the cells at successive levels can be arranged in a hierarchy. Only three new dot-products are needed for each level of refinement. If even this is considered too slow, a lookup table can be constructed indexed by quantized values of two of the components of the unit vector.

Triangular cells have corners which are further away from the center than those of a more rounded cell of equal area. So tesselations with triangular cells are not as desirable as others. Thus we ought to actually use the *duals* of geodesic domes which have many (irregular) hexagonal cells plus twelve pentagonal cells. Unfortunately, it appears that it is now more expensive to compute which cell an unknown normal belongs to, since there is no longer a nice hierarchical arrangement.

Geodesic domes can be made with very large numbers of cells. How many cells are enough? It is clear that if we have too few cells, angular resolution will be low and the orientation histogram a poor approximation to the EGI. Conversely, when we have too many cells, only a few normals will fall in any given cell. That means that the total in a given cell is a very noisy estimate of the average of the inverse of the Gaussian curvature. We have found that a few hundred cells provide a reasonable compromise. The answer depends, of course, on several details, such as how many picture cells fall on the region corresponding to the object of interest. We typically used 256×256 images with a couple of thousand picture cells on the object of interest.

20. Prototypical Object Models

In order to recognize an unknown object and determine its attitude in space, data derived from its image is compared against that obtained from a stored model. The approach outlined earlier works well for determining an orientation histogram of an object given as a prototype. The surface can be divided up into patches and the orientation of each one determined. The patches do not necessarily all have the same area. This is easily taken into account by weighting their contribution to the orientation histogram according to their area. Note that the prototypical orientation histogram is known over the *whole* sphere, unlike the one obtained from the needle diagram. In that case we only have information for the visible hemisphere.

A stored prototypical orientation histogram is to be compared with one obtained from a needle diagram. The picture cells in the image all have the same area. The

areas of the corresponding patches on the surface of the object are not all the same, however, because of foreshortening. We could correct for this, when constructing the orientation histogram from the needle diagram, by dividing by the cosine of the angle between the direction towards the viewer and the surface normal. Applying the correction this way has the unfortunate effect of amplifying errors associated with measurements of surface patches whose normal is nearly at right angles to the direction towards the viewer. It is better, therefore, to instead multiply the prototypical orientation histogram by the cosine factor, when matching the two.

Also, note that we can only calculate the actual area if we know the properties of the camera and the distance to the object. Photometric stereo does not provide us with the latter information. We may not be able to tell the absolute size of the object in this case. The EGI can be normalized by dividing by its integral over the sphere. The result can be used in matching. Naturally, we lose the ability to distinguish objects with the same shape but differing sizes if we do this.

A further complication in the case of an orientation histogram derived from images is that we only get information on the visible hemisphere. Surfaces whose surface normal is turned more than 90° from the direction towards the viewer cannot be seen. In fact, because of limitations of the photometric stereo method, we typically have information about the surface over an even smaller area, perhaps up to 60° from the direction towards the viewer. Some obvious methods for matching extended Gaussian images work only if the whole sphere is known.

21. Moment Calculations (*)

It is not difficult, for example, to calculate the inertia matrix of a mass distribution on the sphere. David Smith at MIT developed a method based on this matrix of second moments. This matrix is useful in that it contains all the information needed to compute the inertia of the mass distribution about an *arbitrary* axis through the center of mass. In particular, using straight-forward calculus methods, one can locate three special axis corresponding to stationary values of the inertia (maximum, minimum and saddle point). These directions, called principal axes, are at right angles to each other (unless the mass distribution happens to be especially symmetrical).

The principal axes are fixed relative to the mass distribution. That is, if the mass distribution is rotated, so are the principal axes. The relative rotation between two extended Gaussian images of the same object can be found simply by calculating the rotation needed to align their principal axes. This provides us with an explicit algorithm for directly computing the attitude of an object relative to its prototype. Nothing more involved than the determination of the eigenvectors of a 3×3 matrix is needed and that, in turn, just requires the solution of a cubic polynomial.

We cannot use such an elegant method here, unfortunately, since the experimentally obtained orientation histogram is known only over some part of the
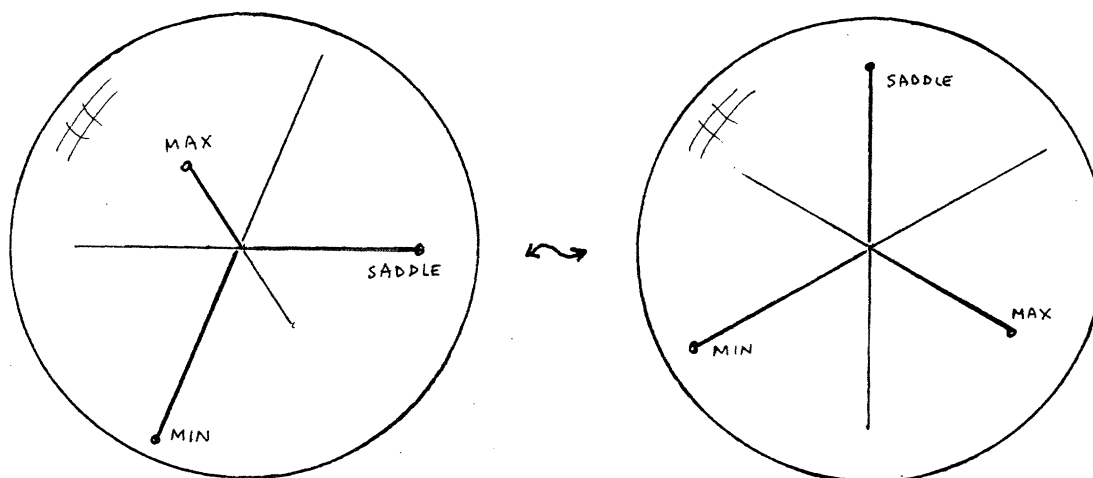
Figure 18. The moment of inertia of a mass distribution about an axis passing through its center of mass depends on the orientation of the axis. The moment of inertia is maximal for one orientation, minimal for another, and has a saddle point for a third. These three special orientations for the axis are called principal axes and lie at right angles to one another. There direction can be conveniently shown as dots on the unit sphere. One mass distribution on the sphere could be lined up with another, just by lining up the principal axes. This represent a straightforward technique for determining the attitude of an object if the *whole* EGI is known.

---

sphere. Also, the match must take into account the foreshortening effect. We do not however have to throw out methods based on moment calculations altogether.

We can, for example, make use of the center of mass of the visible hemisphere. We saw that the center of mass of the *complete* EGI is always at the origin. It is therefore of no use in matching. The center of mass of the visible hemisphere, however, will be at a position which depends on the attitude of the object. We have shown that the first moment of the mass distribution on the visible hemisphere is equal to the apparent cross-sectional area of the object. Now the mass in the visible hemisphere is equal to the *actual* area of the portion of the surface which is visible. Consider again the plane cutting the sphere into visible and invisible hemispheres. The perpendicular distance of the center of mass from this plane is just equal to the ratio of the *apparent* to the *actual* surface area. This will typically vary with the attitude of the object. If we view a football end on, for example, we see half of its surface, but the apparent area is relatively small. Conversely, when we view it from the side, we also see half of its surface, but now the apparent area is relatively large. The ratio is determined easily from the orientation histogram, or, for that matter, directly from the needle diagram.

While the center of mass of the visible hemisphere does not uniquely define the attitude of the object, it can be used to save computation. To speed the matching process one can precompute the expected center of mass given the prototypical orientation histogram and a set of viewing direction for which the match is to be
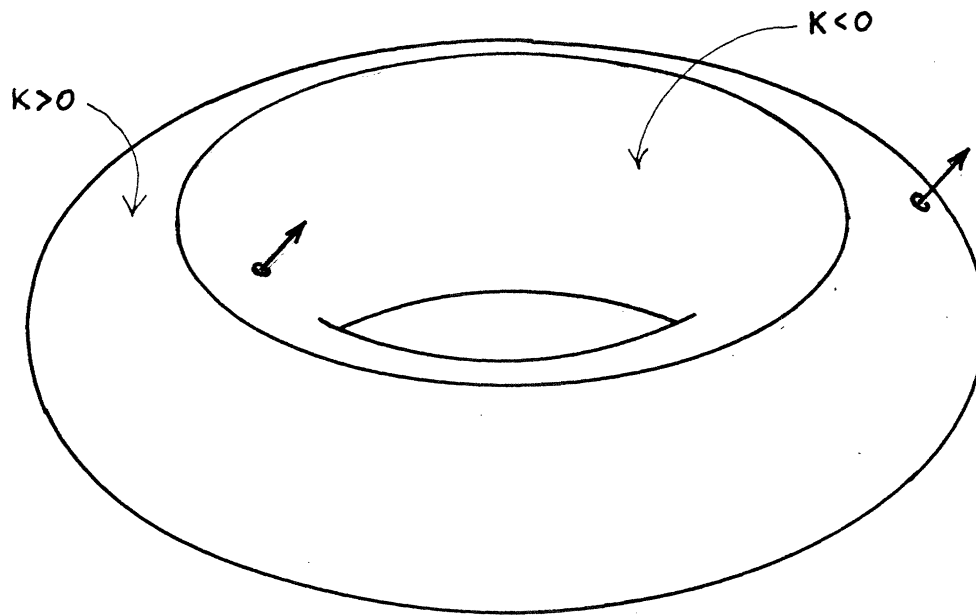
Figure 19. In the case of an object which is not convex, like a torus, the Gaussian curvature will be negative for some points on the surface, and more than one point may have a particular orientation. In this particular case, two points on the surface contribute to a single point on the Gaussian sphere. Furthermore, some parts of the surface may be obscured even if the surface normal there makes an angle of less than 90° with the viewing direction. The definition of the EGI has to be modified to take these effects into account.

---

attempted. Any viewing direction for which the center of mass is not at least in approximately the right position need not be scrutinized further. The discrete set of directions to the viewer for which this calculation is performed may be chosen to be the directions to the cells of the Gaussian sphere for convenience. It may also be advantageous to eliminate potential matches for which the second moments do not agree, although we did not do so.

22. Objects that are not Convex

There are three problems with objects that are not convex

1. The Gaussian curvature is negative for some points on the surface.

2. More than one point on the surafce may map onto the same point on the Gaussian sphere.

3. One part of the object may obscure another.

The precise definition of Gaussian curvature takes into account the direction in which the boundary of corresponding patches on the object and the Gaussian sphere

are traversed. At a convex (or concave) point, the Gaussian curvature is positive, and the boundaries are traversed in the same direction. If they are traversed in opposite directions, as happens at a saddle point, the Gaussian curvature is considered to be negative. Analysis of our simple local process for computing the orientation histogram suggests that we extend our definition to be the inverse of the *absolute value* of the Gaussian curvature, since no account is taken of this.

Also, consideration of the local process for computing the orientation histogram suggests how one can deal with the fact that more than one point on the surface will contribute to a given point on the sphere. We simply add up the inverses of the absolute values of the Gaussian curvature at the corresponding points on the object. This idea can be further developed to deal with cases where all points along a curve or even in a region have the same orientation. We obtain impulse functions on the Gaussian sphere in these cases. We have already seen this in extended Gaussian images of polyhedral objects.

The mapping from the object to the Gaussian sphere is not invertible, unless the object is convex. The only consequence of concern to us here is that there are an infinite number of non-convex objects corresponding to a particular EGI. We do not, however, expect to encounter two different objects with the same EGI in a typical application.

Obscuration is a more difficult issue. In many cases it will be a small effect except for certain directions of viewing, where parts of the object appear to be lined up. One solution is to take obscuration into account by building a view-point dependent EGI, adding in only the contributions from surface patches that are actually visible. The discrete set of directions to the viewer for which this calculation is performed may be once again chosen to be the directions to the cells of the Gaussian sphere for convenience. There is a considerable increase in storage required, but the matching is now no longer disturbed by the effects of obscuration.

It is interesting to determine the EGI of some non-convex object. We can do this for a torus, a good model of the object we used in one of our experiments. The torus is a solid of revolution obtained by rotating a circle about an axis which does not pass through the circle. Consider a plane containing the axis of the torus. It intersects the torus in two circles. It should be clear that points on either one of these circles correspond to points on a particular great circle on the Gaussian sphere. This great circle is obtained by cutting the sphere with a plane parallel to that used to cut the torus. Consider the diameters of these circles which lie parallel to the axis of the torus. The relationship between one of the circles on the torus and the circle on the Gaussian sphere is very simple, one is just a dilation of the other, and points at equal angles to the relevant diameters correspond to each other. Note, however, that to each point on the Gaussian sphere correspond *two* points on the torus, one on each of two circles.

Now add a second plane containing the axis of the torus, but rotated slightly relative to the first. Two narrow slices of the torus lie between these planes. Repeat the construction for the Gaussian sphere. Two pieces shaped like slices of an orange are cut out. These so-called lunes of the sphere are delimited by meridians (lines
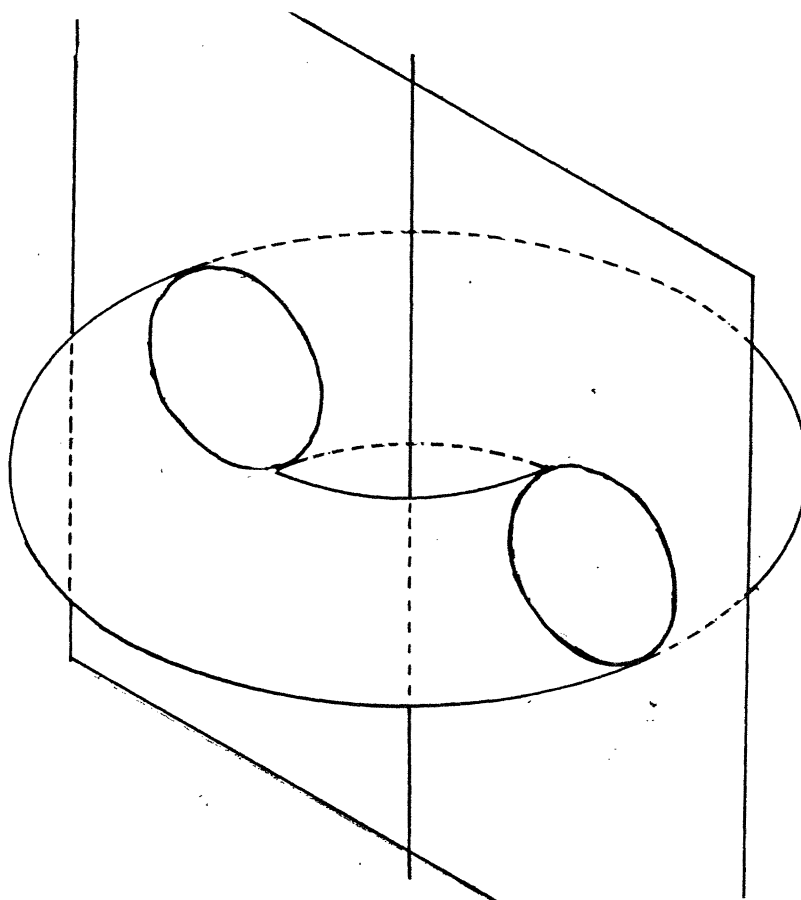
Figure 20. A plane passing through the axis of a torus cuts its surface in two circles. A parallel plane passing through the axis of the Gaussian sphere cuts it in a great circle. Points on the two circles of the torus map onto this great circle. Thus two points on the surface of the torus correspond to every point on the Gaussian sphere.

---

of longitude). Points on one of the slices of the torus map into points on one of the lunes of the Gaussian sphere.

Each of the slices of the torus is narrower where it comes closer to the axis of the torus than where it is further away. The width varies linearly with distance form the axis. This makes it difficult to project one slice onto the Gaussian sphere. It is much easier to consider the two slices together. To obtain the mass density projected onto the Gaussian sphere we have to add up contributions from *both* slices of the torus. Assume now that the slices are very narrow. If one adjoins the two slices one obtains a ring of constant width. The mass from this uniform ring is now projected onto the two lunes on the Gaussian sphere.

Consider encircling the sphere with evenly spaced parallels (lines of latitude). These lines cut the lunes into quadrilaterals. The quadrilaterals are widest near
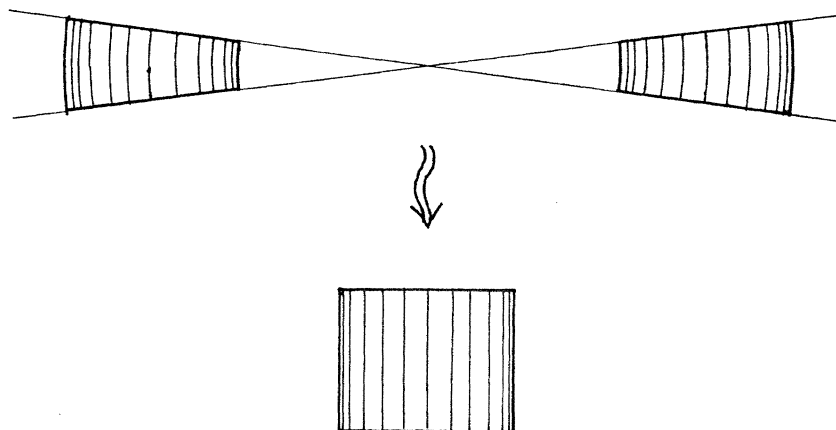
Figure 21. If two planes are used, two slices are cut from the torus. Planes parallel to these cut two lunes from the sphere. The two slices are not of constant width, but can be abutted to form a ring of constant width, provided that the slices are very narrow.
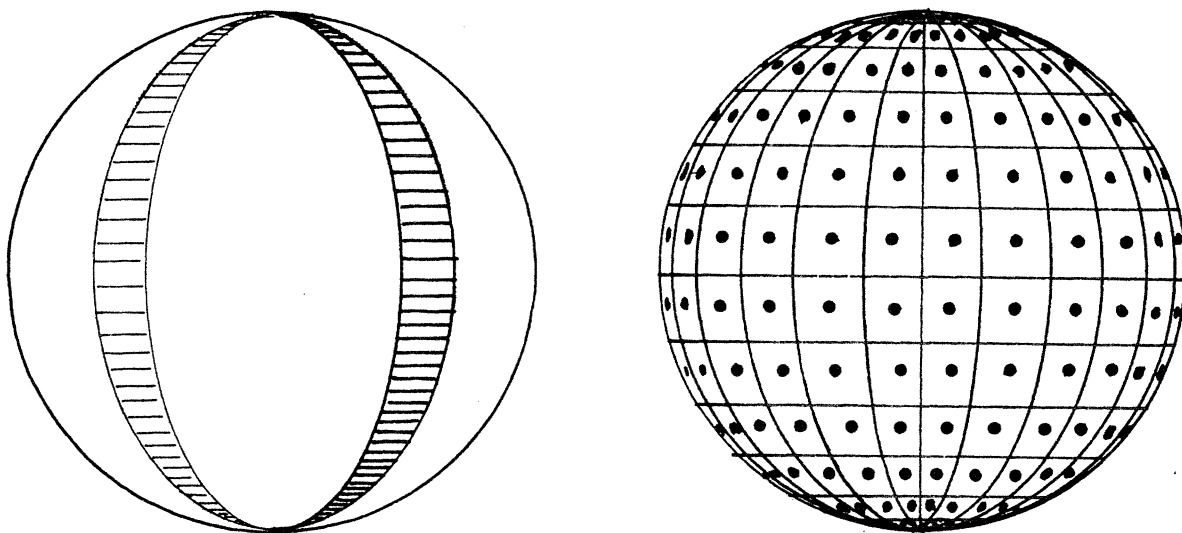

Figure 22. The ring constructed form the torus has to be mapped onto the lunes of the sphere. We can divide the ring into equal strips along its circumference. Each of these strips corresponds to a cell, namely a piece of one of the lunes lying between two curves of constant latitude. The mass in each of these cells is equal to the area of one of the strips of the ring. Therefore the mass in each of the cells is the same. Tese masses are shown here concentrated at the centers of the cells. Clearly the mass *density* varies inversely with the cosine of latitude, since the area of the cells is proportional to the cosine of the latitude.

the equator and become progressively narrower as one approaches one or the other of the poles. They correspond to square areas of *fixed* size on the ring we just constructed. Thus the mass projected into each of the quadrilaterals is the same. But the area of the quadrilaterals varies as the cosine of latitude. The mass *density* on the Gaussian sphere thus varies inversely with the cosine of latitude.

The area of one of the slices of the torus equals the area of the whole torus times the angle between the two cutting planes divided by 2π. The EGI then ends up being equal to the area of the torus divided by 2π times the cosine of latitude. The EGI has singularities at the poles, where the density increases without bound. The poles correspond to the two circles on which the torus would rests if it were dropped on a planar surface. The singularity at a pole arises because all of the points on the corresponding circle have the same surface orientation.

Note that all torii with the same surface area have the same EGI. To find the area of a torus, consider it to be generated by rotating a circle about an axis. The surface area then is equal to 4 times π squared times the product of the radius of the circle and the distance of the center of the circle from the axis of revolution. Thus all torii for which this product is the same, have the same surface area and thus the same EGI. Some will be large and skinny, while others will be small and fat.

While there are many non-convex objects which have the same EGI as the torus, there is only one *convex* object which has this property. It can be shown that this object is a solid of revolution obtained by spinning the curve of least energy about an axis through its endpoints. The curve of least energy is the curve which minimizes the integral of the square of the curvature along the curve.

23. Attitude in Space

The attitude in space of an object is its rotation relative to some reference. To determine the attitude of an object, its EGI is matched with the prototypical EGI. It is easier to first explain how this can be done in the case of solids of revolution.

A solid of revolution is symmetrical about its axis. The attitude of a solid of revolution is fully specified by the direction of its axis. The direction of the axis in turn can be specified by the point were a line parallel to the axis intersects the surface of the Gaussian sphere. Alternatively, it can also be given in terms of the angle it makes with the image plane (elevation) and the angle between its projection in the image and some reference axis (azimuth).

The image of a solid of revolution is symmetrical about the projection of its axis. We could therefore simply find the axis of least inertia of the image region corresponding to the projection of the object. That would pin down one degree of freedom with very little work. This would however mean resorting to binary image processing methods discussed earlier. Their accuracy depends on how well we can find the silhouette of the object. It is better to work with the surface orientation information.
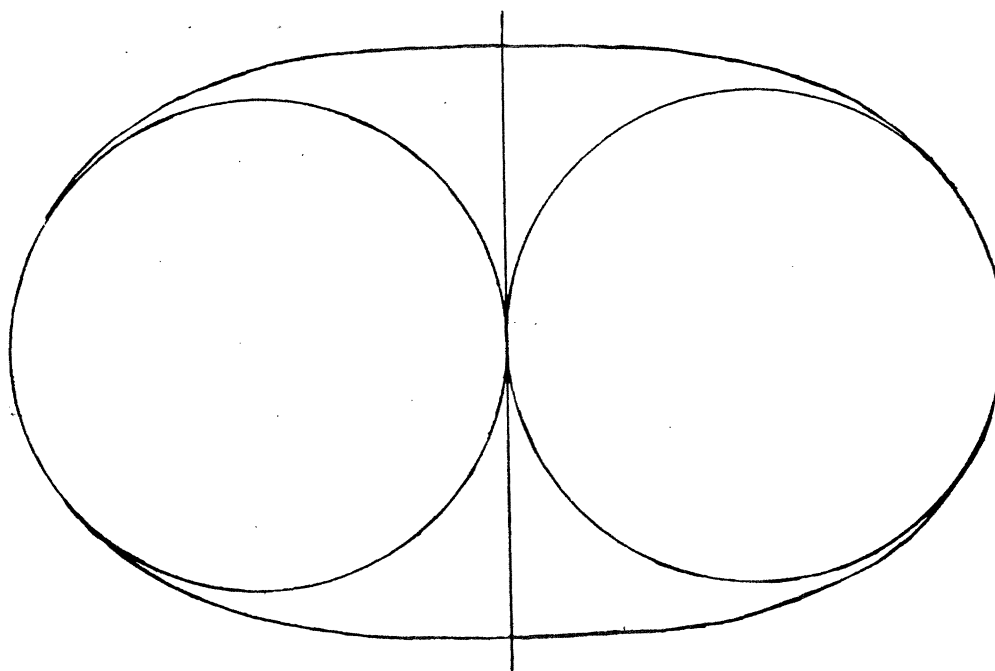
Figure 23. There is only one convex object which has the same EGI as a torus. It is a solid of revolution obtained by spinning the least energy curve about an axis through its endpoints. The least energy curve is the shape adopted by a uniform strip or beam when bent so as to pass through two points with direction at right angles to the line connecting the two points. Such a curve can be used to obtain smooth interpolation between given points and its shape can be given in terms of elliptic integrals.

We can sample the space of possible directions for the axis, trying to match the EGI for each one. It is desirable to sample the space of possible directions evenly. The reason is that one ought to search the space efficiently and avoid sampling one area more finely than another. This leads us to the problem of placing a given number of points "uniformly" on the surface of a hemisphere. We are looking for placements which maximize the minimum distance between points.

This is a problem which has received some attention. It is known, for example, that the best placements for four, six, and twenty points are obtained by projecting the regular tetrahedron, octahedron and icosahedron onto the sphere (The other two regular solids, the cube and dodecahedron, do not lead to optimal placements). It turns out also, that for 32 points, the combination of the dodecahedron and its dual works well. There is no general rule for the optimum. Fortunately, however, the centers of the triangles of geodesic domes appear to provide near optimal placements.

We need not perform a detailed match for each of the chosen directions for the axis of the object. Only directions for which the center of mass matches reasonably well need to be further explored. This means that very few full matches of EGIs
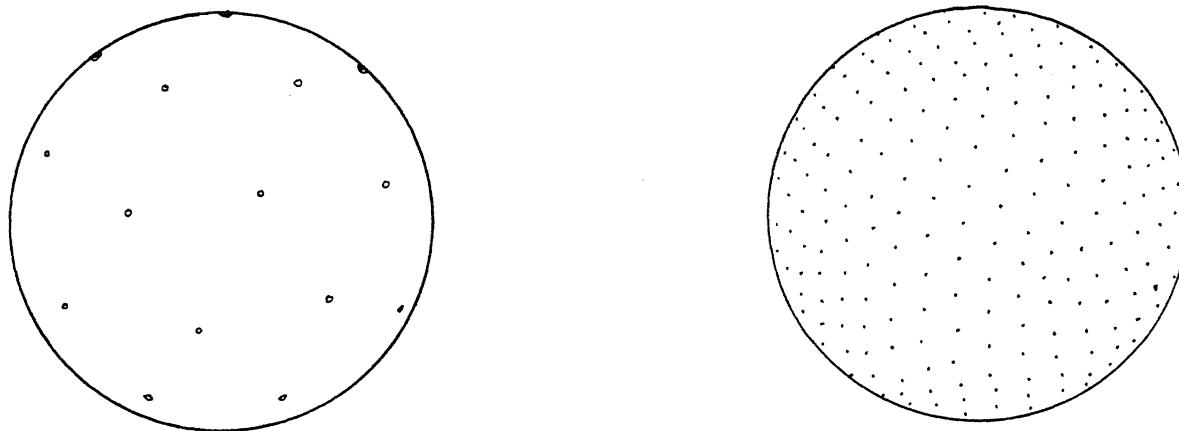
Figure 24. To evenly sample the space of possible attitudes in which a solid of revolution can appear, we need to place points on a sphere, so that they evenly sample the surface of the sphere. Ideally, each point should have the same distance to its closest neighbor. This can be done only if the number of points is small. The optimal placement of 32 points, for example, can be found by combining a dodecahedron and its dual, the icosahedron. For a larger number of points, one searches for a placement which maximizes the minimum separation between points on the sphere. There is no general method known for solving this problem, although geodesic domes combined with their duals are reputed to be good.

---

actually have to be performed. The axis direction which gives the best match is considered to be the correct direction of the axis of the solid of revolution. The match is repeated for several different prototypes if one is to distinguish between several different objects. The unknown is considered to be the object whose prototype it matches best.

Another approach, is to first determine the axis of least inertia of the mass distribution on the visible hemisphere of the EGI. The projection of this axis into the image plane gives us the axis of symmetry of the image of the object. This pins down one degree of freedom (azimuth) with very little computation. It only remains for us to find the inclination of the axis of the solid of revolution (elevation). Thus the search space is reduced from two degrees of freedom to one. Significantly, the axis of least inertia can actually be computed easily from the needle diagram *before* projection of the normals onto the Gaussian sphere, since it is easy to add up the required products to compute the first and second moments. This approach has the advantage that the tesselation of the sphere can be lined up with the axis of least inertia *before* projection of the surface normals onto the Gaussian sphere.

24. Matching Orientation Histograms

Two orientation histograms with their cells aligned can be matched in several ways. One can, for example, take the sum of the squares of the differences of the

totals in corresponding cells as a measure of how different they are. The best match of a given orientation histogram with a set of prototypical ones is the one for which this sum is smallest. Alternatively, one can compute the sum of the products of the totals in corresponding cells. In this case the best match is the one which produces the largest correlation. An advantage of the first method is that a poor match can be rejected without completing the computation whenever the accumulated sum of the squares of the differences becomes large. More complicated, but also more *ad hoc*, comparison functions are easy to dream up.

There are some problems with this approach. This is best illustrated using a polyhedron as an example. Suppose that one of the faces has a normal which points in a direction which just happens to correspond to the edge between two cells on the tesselation of the sphere. Then, a tiny change in attitude can move the full contribution of this particular face from one cell to a neighboring cell. Thus the EGI is changed rather dramatically and the match will be upset. The problem is much reduced for smoothly curved surfaces, but cannot be ignored. One approach to this problem entails storing a *vector* in each cell, which is the sum of the weighted surface normals.

Another approach, is to perform the projection several times, for each attitude, with slightly different alignment of the cells. This would have to be done for both the prototypical and the experimental data. The total amount of work would be multiplied, in this case, by the number of shifted tesselations that are to be used.

In practice there are always small errors in the determination of surface orientation, due to noise in the grey level measurements. Noise in estimating surface orientation tends to smooth the distribution on the sphere, since it displaces some surface normals to the cell next to the one they ought to have been assigned to. The fineness of the tesselation obviously affects how the effects of noise will manifest themselves. If we make the cells large, few surface normals will be placed into the wrong cell. Each cell will have a large total which, statistically speaking, is likely to be a more accurate estimate of the average of the inverse of the Gaussian curvature. At the same time, large cells mean poor accuracy in the determination of attitude. Conversely, if the cells are very small, many will have a zero total, or perhaps just from one patch. Such noisy distributions are hard to match. The problem is entirely analogous to that of picking the "right" histogram bin size for estimating two dimensional probability distributions from a finite number of random samples.

We do not know of an elegant solution to this problem. Inspired by the smoothing effect of noise, however, we decided to deliberately smooth the orientation histogram before matching. This is equivalent to matching a given cell on one histogram with a weighted average of the corresponding cell and its neighbors on the other. It is also possible, when building the orientation histogram, to distribute the contribution of one surface patch to several cells according to how close their normals are to that of the given surface patch.

How many directions should one try for the axis of the object? On the one hand, one need not try too many, since surface normals are not known perfectly in any case. One cannot expect to find the direction of the axis with much more

accuracy than that with which the surface normals can be found. On the other hand, one has to try a large enough number of directions to make sure that the cells on the sphere are brought close to their correct position. An axis direction must be tried which is close enough to the correct one, so that most of the cells line up with each other. In a typical case, we found that about a hundred represent a suitable compromise. Remember though that EGIs will have to be matched in *detail* only for a few of these axis directions. The rest will be rejected on the basis of a gross mismatch in the center of mass of the visible hemisphere.

In practice, we find that the direction of the axis of an object can be determined with an accuracy of about 5° to 10°. This is good enough to permit a robot arm to pick the object up. If better accuracy is required in attitude, a mechanical alignment method may be used *after* the object has been lifted free of the others.

25. Reprojection of the Needle Diagram

If we wish to compare the experimental orientation histogram obtained from the needle diagram, with the synthetic one obtained from the object model, we can arrange for the cells of the two to line up. When the experimental orientation histogram is now rotated however, its cells will generally no longer line up with those of the synthetic orientation histogram. This means that one has to rotate the normals in one of them, before projecting them onto a tesselated sphere in the standard attitude. Reprojection of the normals is perhaps most conveniently performed with the synthetic data, since it can be done once, ahead of time, and the results stored. Fortunately, as mentioned before, we can greatly reduce the effort if the chosen tesselation has the property that the cells will line up again, at least for some special rotations. A tesselation with this property simplifies matching, since some rotations of the orientation histogram merely *permute* the totals in the cells. This is why we were interested in choosing tesselations which have this property.

The faces of the regular solids will line up for the rotations belonging to the finite subgroup of the continuous group of rotations corresponding to that solid. These subgroups have size 12, 24, and 60 for the tetrahedron, octahedron and the icosahedron respectively. Tesselations based on the icosahedron and its dual, as for example, the soccer ball and the Pentakis dodecahedron, have the same rotation group. In the case of the soccer ball, we can easily list the rotations by considering three classes of rotation axes.

1. First, we have a five-fold symmetry about any axis passing through the center of one of the pentagonal cells. This gives us $(12/2) \times 4 = 24$ rotations.

2. Secondly, we have a three-fold symmetry about any axis passing through the center of one of the hexagonal cells. This gives us $(20/2) \times 2 = 20$ rotations.

3. Finally we have a two-fold axis of symmetry about the center of any edge between hexagonal cells. This gives us another $(30/2) = 15$ rotations.

If we add the identity to the above, we end up with 60 altogether. Unfortunately, there are no finite subgroups of the group of rotations with a larger number of
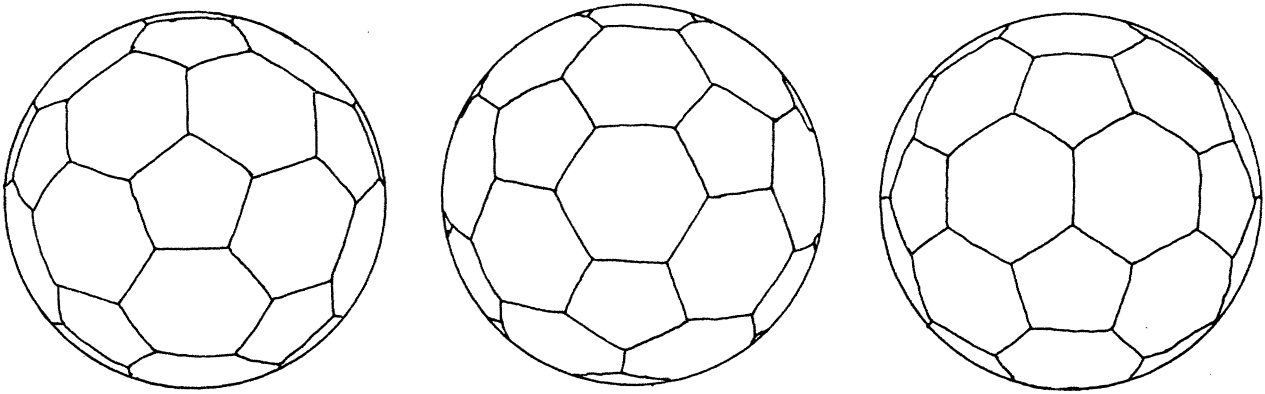
Figure 25. The soccer ball can be used to illustrate the group of rotations of the dodecahedron and the icosahedron. There are six five-fold axes of symmetry passing through the centers of each of the pentagonal cells. There are ten three-fold axes of symmetry passing through the centers of the hexagonal cells. Finally, there are fifteen two-fold axes of symmetry passing through the centers of the edges. Together with the identity rotation one obtains sixty ways of rotating the soccer ball in such a way as to bring pentagonal cells back into alignment with pentagonal cells and hexagonal cells back into alignment with hexagonal cells. Unfortunately, there is no finite subgroup of the group of rotations in three dimensions with a larger number of elements.

---

elements than this (If we ignore groups which contain only rotations about a single axis). To deal with more than 60 rotations then, reprojection is required.

26. Corrections for Departure from Ideal Conditions (*)

Several of the implicit assumptions in the above analysis are violated in practice. It is assumed, for example, that the brightness of a surface depends *only* on its orientation, not on its position. This is the case when the light-sources are infinitely far away. In practice, light sources are close enough to the surface on which the objects are placed so that the inverse square law comes into play. This can be taken account of by a normalization of the brightness values read. One first takes images of a uniform white surface using each of the three sources in turn. We found that a linear approximation to the resulting brightness distribution is accurate enough. All images are then corrected for the non-uniformity in illumination by means of a linear function of the position in the image.

There is another problem which is harder to deal with. Since the light sources are nearby, the *direction* of the incident rays is not the same for all points. This means that the computed surface normals will be off. We found the error due to this effect to be smaller than that due to non-uniform illumination and harder to correct for. So we ignored it.

No image sensing device is perfect. Fortunately, charge coupled device (CCD) cameras have very good geometric accuracy and are linear in their response to

brightness. The sensor cells do not, however, all have the same sensitivity to light. Some, due to defects in the silicon, are "weaker" than others. One could take this into account by taking a picture of a point source on the optical axis of the camera when the lens is removed. This would provide uniform illumination of the image plane. The result could be used for correcting all future brightness measurements.

Instead, we normalize the three brightness measurements at each picture cell by dividing by their sum. This eliminates the effect of non-uniform sensor response and also accounts for fluctuations in illumination. Furthermore, it makes the system insensitive to differences in surface albedo from point to point on the object. Objects typically do not have perfectly uniform surface reflectance properties. In our experiments, for example, the debugging effort entailed episodes of rather rough handling of the parts by the manipulator. The normalization method used to deal with non-uniform sensitivity of the image sensor automatically also provides for fluctuations in surface reflectance. This approach does however make it harder to detect shadowing and mutual illumination, which we saw were helpful in segmentation of the image.

At times, because of severe noise, an imaging device defect, or a surface mark, an isolated point in the image will not be assigned a surface orientation by the photometric stereo method. We search for these isolated points and enter a normal which is equal to the average of the neighboring values. The main reason for doing this, is that such blemishes would count as holes in the computation of the Euler number.

We also have developed a method which will deal with noise using a constraint based on the assumption that surface orientation varies smoothly almost everywhere (So far, we have only assumed that the surface is continuous almost everywhere). This iterative method, based on the solution of a calculus of variation problem, can deal with severe noise, but is slow. Fortunately, we did not have to use it.

27. Picking the Object to Pick Up

Once the image has been segmented into regions which appear to be parts of objects, a decision can be made about which one of these is to be analyzed further. The region chosen should correspond to an object near the top of the pile. As little as possible of this object should be obscured. This is so that the manipulator can easily pick it up, but also, so that matching with the prototype will work well. Furthermore, there may be reasons to prefer objects with certain attitudes, either because they are easier to pick up or because it is known that the system is more accurate in determining their attitude. No absolute depth information is available from photometric stereo, so that it is not trivial to pick a suitable object.

Several heuristics can be used to select a "good" object for the manipulator to pick up. First of all, the region in the image should have a relatively large area if the object is unobscured. Also, the ratio of perimeter squared to area can be used to estimate the elongation of the region in the image. A highly elongated

region may be a cue that the object lies in an attitude that the manipulator will have difficulty with. Finally, the Euler number may be relevant. In the case of an unobscured toroidal object, the Euler number will be zero, unless the axis of the torus is highly inclined relative to the direction towards the viewer.

Another task for the system is to decide how to pick up the object, once its attitude in space is known. The system has to be told which points on the surface of the object are suitable for grasping. Also, the gripper should be placed so that it will not interfere with neighboring objects. It is helpful, in this regard, to pick a point which is relatively high on the object. Such a point can be found since the object's shape and attitude are known. It would also be reasonable to avoid places on the object which correspond to places in the image where neighboring regions come close to the region analyzed.

It may not always be possible to guarantee that the object can be picked up as calculated, particularly if absolute depth information is not available. In this case, tactile sensors help to detect problems such as collisions with neighboring objects and loss of grip on the part being picked up. It is best then to remove the arm from the field of view and start over. An obvious problem is that the rate at which parts are picked up is not constant if this happens. Some mechanical buffering scheme can be used to solve this problem.

When there are no more objects to pick up the needle diagram will be uniform. The image will then not be broken into separate regions and processing can stop.

28. Moving the Arm

Control of the mechanical manipulator is relatively straightforward compared to the vision part. We have used photometric stereo and matching of orientation histograms to determine the attitude of the object we wish to pick up. The *position* of the region of interest can be estimated by finding its center of area. This binary image processing technique is to be avoided, however, since the silhouette of this region may be quite rough. It is better to obtain the position more accurately by matching the needle diagram with one computed using the object prototype and the now known attitude of the object.

The position in the image of the region corresponding to the object of interest defines a *ray* from the camera. Since photometric stereo does not provide absolute depth information, we cannot tell how far along this ray the object is. The arm is therefore commanded to move along the ray, starting at some safe height above the surface on which the objects lie. A proximity sensor is used to detect when the arm comes near an object. In our case, a modulated infrared light beam from one finger of the gripper to the other is interrupted by the object. At this point the hand can be re-oriented so that its attitude matches that of the object. The gripper is then closed and the object lifted free.
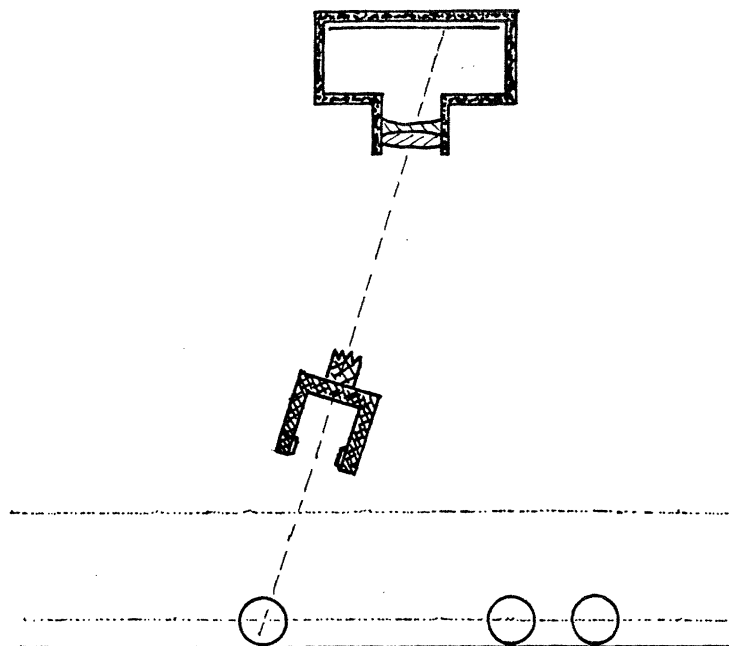
Figure 26. From a single camera position we cannot determine the actual three dimensional coordinates of an object. From where the object appears in the image, however, we can tell what ray in space must be followed to find it. The computer controlled arm can then be sent along this ray until it detects the object by means of a proximity sensor. To avoid this relatively slow search, another method, like binocular stereo, can be used to determine the absolute distance to the object.

29. Calibration of the Hand-Eye Coordinate Transform

In order to command the arm to trace along a particular ray from the camera, it is necessary to transform coordinates measured relative to the camera to coordinates measured relative to the arm. This transformation has six degrees of freedom and can be represented by a translation and a rotation. It is hard to determine it with sufficient accuracy using direct measurements of the camera's position and attitude. It is much more convenient to have the arm move through a series of known positions in front of the camera. The position of the image of the arm in the camera is then determined and used to solve for the parameters of the transformation. To make for high accuracy, more than the minimum number of measurements are used, and a least squares adjustment carried out.

It is very hard to develop a program which can recognize and track the arm. For this reason we actually have the arm hold a so-called surveyor's mark which is easy to locate in the image. It is essentially a 2 × 2 sub-block of a checker-board. The intersection of the two lines separating dark from light areas can be located with high precision.

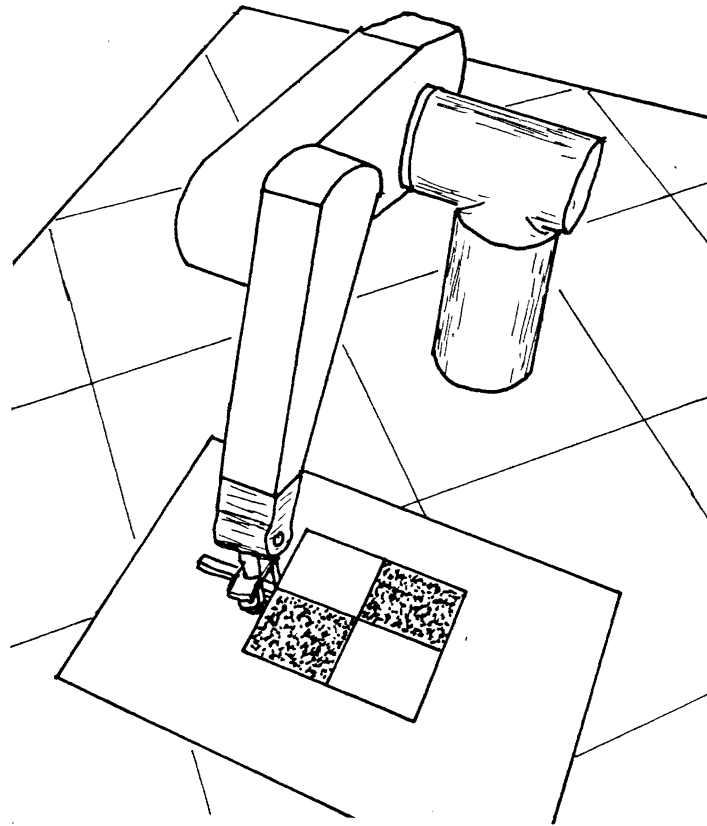In our experiments, the camera is mounted high above the arm in such a way

Figure 27. The relationship between the coordinate system of the robot arm and the camera eye is determined by a calibration process. An object which is easy to locate in the image is carried by the arm to a series of positions while the corresponding image coordinates are measured.

---

that it effectively looks straight down (Actually, a mirror is used to prolong the optical path). The image plane is nearly parallel to the plane containing the two horizontal axes of the arm's coordinate system. This means that for this plane, or one parallel to it, one can approximate the perspective projection by an affine transformation having six parameters. So, in order to simplify matters, we have the arm move through a number of points in one plane to determine one such affine transform. This process is then repeated in a plane closer to the camera. Thus each point in the image can be mapped into one point in each of the two planes. These two points define a ray in arm space.

30. Objects of Arbitrary Shape

The methods described above made use of the fact that the objects were solids of revolution. We only had to recover the two degrees of freedom of the axis of the object. In the general case, the EGI certainly can still be used, but attitude now

has *three* degrees of freedom. One way to see this is to note that an object can be rotated about an arbitrary axis by an arbitrary angle. It takes two parameters to specify the axis and one for the angle. What this means is that matching becomes more tedious. A larger number of potential matches have to be tried. Still, the same filtering operations can be employed to eliminate most of them.

A simple extension of what we described above allows us to deal with objects that are not solids of revolution. We once again use the axis of least inertia of the mass distribution on the visible hemisphere to pin down one degree of freedom. The remaining problem is to determine the direction from which the object is viewed. The possible directions can be specified by points on a sphere. We generate a discrete sampling of the surface of the sphere which is as near to being uniform as possible. One can use the same tesselation of the sphere used for the orientation histogram.

One way to represent rotations of a rigid object is by means of unit quaternions. These can be thought of as vectors having *four* components or a "hyper-complex" numbers with a real part and three imaginary parts. Amongst all of the ways commonly used to deal with the rotation of a rigid body, this one has the advantage that it allows one to define a metric on the space of rotations. That, in turn, permits one to consider averages over all rotations, for example. Recently, Phillipe Brou at MIT, has develped methods for evenly sampling the space of rotations using specially designed polytopes in four dimensional space. His approach allows one to attempt matches for large sets of rotations without storing a large number of prototypical EGIs. Essentially, one obtains 60 attitudes from each stored EGI. Precomputing six EGIs allows one to sample the space of rotations (nearly) uniformly with 360 points.

The brute-force matching of orientation histograms described can become expensive if the attitude is to be determined with high precision. This is because the space of rotations is three dimensional and so the number of attitudes we have to try goes up with the *cube* of the precision. Hill-climbing methods for searching the space of rotations may appear attractive in view of this. One could imagine, for example, first finding a rough estimate of the attitude, by considering the 60 rotations of the icosahedron. The attitude which produces the best match is then used as an initial value for an iteration which at each step seeks to improve the match further by making small adjustments. It is unfortunate that such methods do not seem to work. We found that the match does not become good until one is really close to the correct attitude.

31. Experimental Results

We chose plastic torii of about 120 mm outer diameter as the test objects. Their geometry is simple to model and they can be easily picked up using a crude parallel jaw gripper. We used torii as test objects because they have a shape that is easy to model, while not being polyhedral or convex. The system looks at a
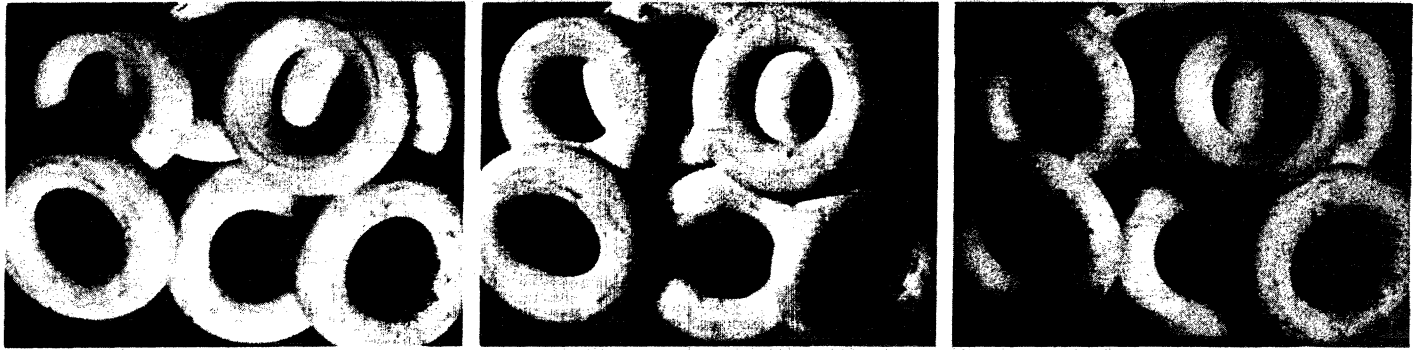
Figure 28. Three images of a pile of torus shaped objects. The images are taken with three different light sources turned on. At first glance the images may look very similar. This is because we interpret the shading in terms of object shapes. Close inspection shows, however, that the grey values at corresponding points of the three images are typically very different. Photometric stereo is used to obtain a needle diagram from these images.

---

pile of these objects using a Hitachi (TM) charge coupled device (CCD) camera. Three images are digitized with each of three banks of four 40 watt fluorescent lights powered on in turn. The grey level images are digitized to about 256×256 picture cells and read into a single user computer called a Lisp machine (TM). We used a frequency two geodesic dome based on the Pentakis dodecahedron for the orientation histogram. It has 240 cells. The attitude of one of the objects is then determined by matching the experimental orientation histogram against a prototypical orientation histogram. We make use of the axis of least inertia of the orientation histogram to reduce the search space. A Unimation Puma (TM) arm is employed to pick up the object chosen.

We found, by the way, that inexpensive vidicon cameras suffer from significant geometric distortion. An even more important problem with these devices is that the digitized grey levels do not bear a reproducible relationship to image brightness, even with the automatic gain control (AGC) disabled. This is why we prefer CCD cameras. It should also be said that industrial robots today typically have very good repeatability, but poor absolute accuracy. That is, they will go back to a position taught in terms of joint angles with great precision, but can be several millimeters off when asked to go to a position specified in Cartesian coordinates. This is a significant problem when sensors are used to locate parts.

Our system takes about a minute to read in the images, switch lights on and off, perform the matching and send commands to the manipulator over a serial line. There is no inherent reason why the cycle time could not be much shorter. We were interested in demonstrating the feasibility of this approach, not in the maximum speed possible with our particular arrangement of system modules. Most of the time the system successfully picks up one of the objects in the pile. Occasionally it fails, usually because the fingers bump into another object before picking up the
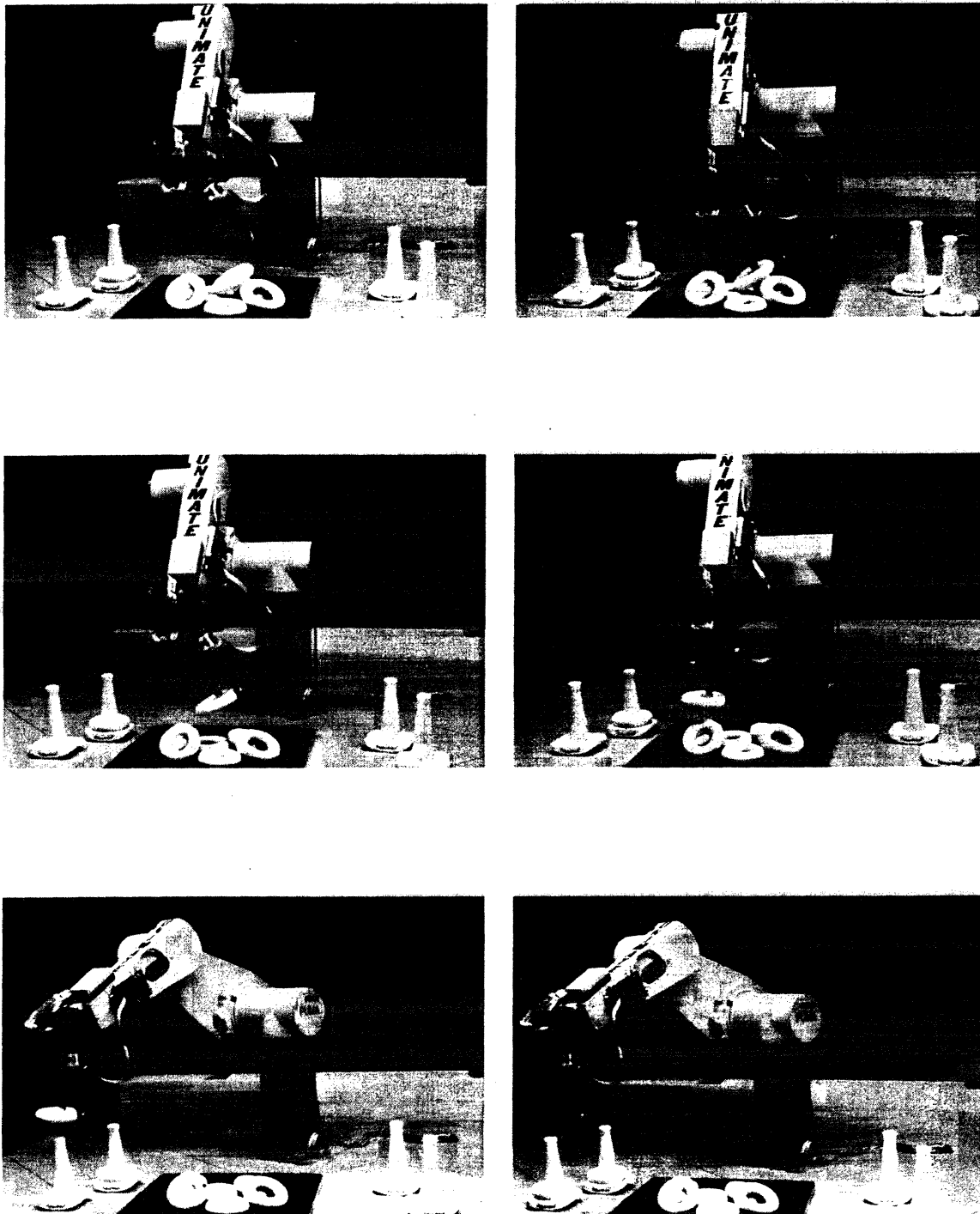
Figure 29. A picture sequence showing the arm picking up a few of the objects from the pile using the image information to tell it where the objects are and how they lie in space.

desired one. In this case it just removes the arm from the field of view and starts over. Better algorithms for picking a good grasping position would help to improve the performance even further. These would make use of depth information which is not available from photometric stereo.

We did just that recently, using a robust, low resolution but high speed, binocular stereo system developed by Keith Nishihara. In order to use the depth information we had to solve the spatial reasoning problems involved in determing a suitable grasping position on the object; one which would hold the object stably and not cause the gripper to collide with the other objects.

32. Conclusions

We have demonstrated the feasibility of a machine vision system for picking objects out of a pile of objects. Our system uses multiple images obtained with one camera under changing lighting conditions. From these images a needle diagram is computed, which gives estimates of the orientation of surface patches of the objects. This in turn is used to compute the orientation histogram which is a discrete approximation of the EGI. The experimental orientation histogram is matched against an orientation histograms determined using computer models of the objects. In this way the attitude of the object in space is obtained. A manipulator can then be sent along a ray in space to pick up the object.

While our system is not particularly fast, there is no reason why a faster one could not be built, since all of the computations are simple, mostly involving table lookup. Special purpose hardware could also be build to speed up the matching process. It would not have to be very complicated since it performs a kind of correlation process.

We believe that what we have described provides a robust approach to the recognition of objects and the determination of their attitude in space. It will work better than an approach based on recognizing some special feature of the object given that only a few thousand picture cells are scanned per object region. In the case of an approach based on recognition of special features a few thousand points would be needed for *that feature*, so that the number of picture points for the whole object would be much larger.

The needle diagram can be computed from a depth map by taking first differences. The method we described is therefore also applicable to other input, such as depth maps obtained using laser range finders. We did not use one in our experiments since they still appear to be quite expensive and slow. We did, however, experiment with depth maps obtained using automated stereo.

The above is representative of a new approach to problems in machine vision. It is based on careful analyses of the physics of image formation and views machine vision as an inversion problem.

33. Further Reading

Bolles, R. C., Horaud, P. and Hannah, M. J. (1983) "3DPO: A Three-Dimensional Part Orientation System," Proceedings, International Joint Conference in Artificial Intelligence, Karlsruhe, August, pp. 1116–1120.

Brown, C. M. (1977) "Representing the orientation of dendritic fields with geodesic tesselations," Internal Report TR-13, Computer Science Department, University of Rochester.

Coxeter, H. S. M. (1973) *Regular Polytopes*, Dover Publications, N. Y.

Fejes Toth, L. (1964) *Regular Figures*, Pergamon Press, N. Y.

Hamilton, W. R. (1899) *Elements of Quaternions*, Chelsea Publishing, N. Y.

Hilbert, D. and Cohn-Vossen, S. (1952) *Geometry and the Imagination*, Chelsea Publishing, N. Y.

Horn, B. K. P. (1970) "Shape from Shading: A method for determining the shape of a smooth opaque object from one view," Internal Report TR-79, M.I.T. Project MAC, November.

Horn, B. K. P., Woodham, R. J. and Silver, W. M. (1978) "Determining Shape and Reflectance using Multiple Images," M.I.T. A.I. Laboratory Memo 490, August.

Horn, B. K. P. (1979) "Sequins and Quills — representations for surface topography," M.I.T. A.I. Laboratory Memo 536, May.

Horn, B. K. P. (1983) "Extended Gaussian Images," M.I.T. A.I. Laboratory Memo 740, August.

Kelley, R. B., Birk, J. R., Martins, H. A. S. and Tella, R. (1982) "A Robot System which Acquires Cylindrical Workpieces from Bins," I.E.E.E. Trans. on S.M.C., Volume SMC-12, No. 2, March/April.

Lysternik, L. A. (1963) *Convex Figures and Polyhedra*, Dover, N. Y.

Pearce, P. and Pearce, S. (1978) *Polyhedra Primer*, Van Nostrand Reinhold, N. Y.

Pogorelov, A. V. (1956) *Differential Geometry*, Noordhoff–Groningen, The Netherlands.

Smith, D. A. (1979) "Using Enhanced Spherical Images," Memo No. 530, M.I.T. A.I. Laboratory, May.

Woodham, R. J. (1978) "Photometric Stereo: A Reflectance Map Technique for determining Surface Orientation from a Single View," *Image Understanding Systems and Industrial Applications*, Proceedings S. P. I. E. 22nd Annual Technical Symposium, Vol. 155, August, pp. 136–143.

Woodham, R. J. (1980) "Photometric Method for Determining Surface Orientation from Multiple Images," *Optical Engineering*, February, pp. 139–144.