Combining Shape and Color Information For 3D Object Recognition

Kazunori Higuchi, Martial Hebert, Katsushi Ikeuchi

December 1993
CMU-CS-93-215

Computer Science Depertment
Carnegie Mellon University
Pittsburgh, PA 15213

This work was sponsored in part by the National Science Foundation under IRI-9224521 and in part by the Avionics Laboratory, Wright Research and Development Center, Aeronautical Systems Division (AFSC), U.S. Air Force, Wright-Patterson AFB, Ohio 45433-6543 under Contract F33615-90-C-1465, ARPA Order No. 7597 and Contract F33615-93-1-1282.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NSF, ARPA, or the U.S. government.

Keywords: object recognition, spherical representation, range data, color, pose estimation

Abstract

Both photometric and geometric information are important for 3D object recognition. Traditionally, however, few systems utilized both types of information. This is because no single representation is suitable for both types of information. This paper proposes a method for representing both color and geometric information using a common framework, the Spherical Attribute Image (SAI). The SAI maps the values of curvature and color computed at every node of a mesh approximating the object surface onto a spherical image. A model object and an observed surface are computed by finding the rotation that brings their spherical images into correspondence. We show how this matching algorithm can be used for object recognition using both geometric and photometric information. In addition, we describe how the two types of information can be combined in a way that takes into account their actual distribution on the surface.

Contents

1 Introduction 6
2 Overview of SAI Coordinate System 7
2.1 Definition of SAI Coordinate System 7
2.2 Cartesian Coordinates and SAI Coordinates 8
2.3 SAI Matching 9
3 Matching Using Color Data 10
4 Combining Photometric and Geometric Information 13
4.1 Combining Similarity Measures 13
4.2 Experimental Evaluation 14
5 Conclusion 18
6 References 18

List of Figures

Figure 1 Indexing nodes on a 2-D curve 10
Figure 2 SAI matching based on color information 12
Figure 3 Combining color and curvature information-1 17
Figure 4 Combining color and curvature information-2 17
Figure 5 Combining color and curvature information-3 17
Figure 6 Matching between a complete model and a partial view 17

1 Introduction

This paper addresses the problem of recognizing three-dimensional curved objects in range and color images. More specifically, given a three-dimensional model of an object and a similar description of a surface observed in a range and color image, the problem is to compute the pose which best matches the model and the observed surface.

Traditionally, this problem has been addressed in two different ways. Some approaches use only *geometric* information such as curvature, local shape indicators, surface patches, or surface features such as edges. Other approaches use *photometric* information in the form of regions of uniform hue, color histograms, or edges extracted from reflectance discontinuities in images. There is little in common between these two classes of approaches. They use different types of object models, different matching strategies, and operate under different assumptions. These make variances building an object recognition system using both geometric and photometric information difficult. This paper addresses this problem.

We represent both photometric and geometric information using a common framework, the *Spherical Attribute Image (SAI)*. The SAI coordinate system is based on constructing a quasi uniform regular mesh on the object surface by deforming a spherical mesh. The deformed mesh has a one-to-one mapping to the original reference spherical mesh. We obtain a spherical representation of the object by mapping properties computed at every node of the mesh, such as curvature to the corresponding nodes on the sphere. By comparing the spherical image of a complete model with the partial spherical image extracted from an image, we can obtain the pose of the object in the scene.

In our previous work [5], we used only shape information because we stored only curvature in the spherical image. We also showed how to obtain the globally optimum pose without combinatorial search. These results were encouraging steps toward recognizing 3-D curved objects. However, shape information alone is not sufficient for matching objects in general. Consider the simple case of a painted sphere, for example. It is obvious that it is impossible to compute the transformation between a model and an observation of the object, whereas it is equally obvious that we ought to be able to compute the transformation from color information. Although this is a rather artificial example, it illustrates the need for including photometric information.

In itself, using photometric information would not be an improvement if it required precise color segmentation or the extraction of color features, both of which are difficult problems in themselves. The main contribution of our algorithm is precisely that it does not require any color segmentation or color features because it uses the color distribution on the surface directly. Since we use essentially the same approach with color as with shape information, the matching algorithm also has the same two important properties emphasized in [5]. First, it is not based on a restrictive mathematical surface model and it can be applied to a large class of 3-D curved objects. Second, it can be applied to matching partial views of objects as is normally the case in object recognition problems instead of only full models even though we are using a spherical representation.

The paper is divided into three parts. In the first part, we review our approach to object matching using the SAI. This approach was initially developed for recognizing objects using shape information such as curvature. In this first part, we indicate how non-geomet-

ric information, such as color, can be included in the same framework in a natural way. We do this by drawing a parallel between matching spherical images of 3-D surfaces and correlating 2-D images. In the second part, we discuss the matching of 3-D objects using color information only. We analyze the performance of the algorithm on a cylindrical object whose pose cannot be calculated from shape information only. At the end of this section, we compare the use of color and shape information on an object for which either could be used for recovering object pose. This example is used as an introduction to the third section in which we present an algorithm for combining shape and color information in a way that takes advantage of both. We show results that indicate that the matching algorithm performs better by combining the two types of information than by using either one individually.

2 Overview of SAI Coordinate System

This section reviews the basic ideas behind our representation of surfaces and introduces the extension of the matching algorithm from geometric to photometric information by drawing an analogy between correlating sets of points in the plane and matching surfaces using our representation.

2.1 Definition of SAI Coordinate System

Our basic representation is a mesh of nodes approximating the surface with a fixed topology and certain local regularity properties. Fitting a discrete mesh to range data is done using techniques based on the concept of deformable surfaces [4].

The deformation process involves a local regularity constraint. The local regularity constraint is introduced so that two instances of the same surface in different poses have similar meshes. Specifically, while building the mesh, we enforce a local constraint so that the distribution of mesh nodes on the surface is approximately uniform. This constraint is included in the deformable surface formalism. See [4] for more details.

Figure 1 illustrates an indexing of nodes between the object mesh and its spherical representation in 2D case. Since the mesh has a fixed topology, it can be viewed as a deformation of a canonical reference mesh on the unit sphere. Specifically, for a given size of the mesh, a standard numbering of the nodes can be defined on the mesh. We call the number associated with every node its index. This indexing depends only on the size of the mesh and its topology. Therefore, there is a similar indexing scheme for the reference spherical mesh and for the surface mesh; we can associate with each node $\boldsymbol{N}$ of the surface mesh the node $\boldsymbol{N}_{sph}$ of the spherical mesh with the same index. This defines a standard mapping between a mesh and the unit sphere.

Finally, we can store at a node $\boldsymbol{N}_{sph}$ of the spherical mesh any quantity computed at the corresponding node $\boldsymbol{N}$ of the surface mesh, thus creating a spherical image of the surface. Once the spherical images are built, matching objects involves finding the best rotation that brings the two spherical images in correspondence.

2.2 Cartesian Coordinates and SAI Coordinates

In our previous work [5], we viewed the spherical image as a way to store shape information. Specifically, we introduced the notion of a simplex angle which is a measure of the local curvature of the surface[1], and we called the corresponding spherical image the simplex angle image. In general, we can store in the spherical image any value computed from the object, not just shape information. The rest of the paper focuses on the use of photometric information instead of, or in combination with, shape information.

In order to understand the properties of the SAI representation, it is convenient to view it as a way to define intrinsic coordinates on the surface. Specifically, we can draw an analogy between Cartesian coordinate systems in the plane and SAI coordinate systems. Under this analogy, coordinates of points corresponds to indexing of nodes on a mesh, while correlation of planar sets of points corresponds to correlation of SAI matching.

In conventional Cartesian geometry, three points in a plane and the corresponding three points in a transformed version of the plane define a transformation between the two planes. Namely, the location of a point in the first plane is entirely determined by the location of the corresponding point in the transformed plane with respect to the three basis points.

In the case of SAI representation, the correspondences between three arbitrary nodes on two meshes $\mathcal{M}_1$ and $\mathcal{M}_2$ define a unique mapping between the nodes of $\mathcal{M}_1$ and $\mathcal{M}_2$.

In Cartesian geometry, knowing the correspondences between points defines a change of coordinates. In a similar way, in the SAI coordinate system, knowing the correspondences between nodes defines a transformation of the indices of the nodes. Moreover, just as the only coordinate transforms that preserve distances and angles are the rigid transformations, the only indexing of nodes that preserve the local regularity constraint and the topology of the mesh are the ones that can be generated by a rotation of a spherical mesh. This last property can be better understood in the context of plane curves, in which two different indexings of the vertices on the curve correspond to a rotation of the node numbers on the unit circle (Figure 1).

To summarize, the SAI mesh has the same properties as the 3-D Cartesian space by replacing the words “plane” by “spherical surface”, “coordinates” by “indices”, “points” by “mesh nodes”, “change of coordinates” by “index transformation”, and “rigid transformation” by “rotation of the SAI mesh”. The analogy can be pushed further by saying that two sets of points in the plane are matched if there exists a rigid transformation that brings the two sets into registration and such that the values stored at corresponding points are identical. The analogous property of SAI meshes is that two meshes represent the same object if there exists a valid index transformation of the nodes between the two meshes such that the values stored at corresponding nodes are identical.

1. Although the simplex angle is not exactly the Gaussian nor the mean curvature of the surface, we will refer to it as “curvature” in the rest of paper for simplicity.

2.3 SAI Matching

As a result of the last property, we can see that in order to match two meshes $\mathcal{M}$ and $\mathcal{M}'$ constructed from two views of a surface, we need to find the rotation $\boldsymbol{R}$ such that the value stored at any node $\boldsymbol{N}$ of $\mathcal{M}$ is the same as the value stored at the corresponding node $\boldsymbol{N'} = \boldsymbol{R}(\boldsymbol{N})$. This rotation establishes an index transformation between $\mathcal{S}$ and $\mathcal{S}'$, which in turns establishes a correspondence between the nodes of $\mathcal{M}$ and $\mathcal{M}'$. Finally, the rigid transformation between the underlying surfaces of $\mathcal{M}$ and $\mathcal{M}'$ is computed by minimizing the distance between the corresponding nodes of $\mathcal{M}$ and $\mathcal{M}'$.

Formally, the best rotation is found by finding the maximum of a similarity measure $D(\boldsymbol{R}, \mathcal{S}, \mathcal{S}')$ related to the sum of squared differences between the value of surface curvature c stored at every node $\boldsymbol{N}$ of $\mathcal{S}$ and the curvature c' stored at the corresponding node $\boldsymbol{R}(\boldsymbol{N})$ of $\mathcal{S}'$:

$$D(R,S,S') = \sum_{N \in S} exp\left(-\frac{1}{2}\left(\frac{c(N) - c'(R(N))}{\alpha_c}\right)^2\right) \quad (1)$$

where α_c is a constant value that defines the sensitivity of the similarity to the differences between the corresponding nodes of $\mathcal{M}$ and $\mathcal{M}'$. If α_c is too small, only small differences contribute to $D(\boldsymbol{R}, \mathcal{S}, \mathcal{S}')$. In particular, there is no contribution at the node that has differences beyond $3\alpha_c$. On the other hand, the maximum of $D(\boldsymbol{R}, \mathcal{S}, \mathcal{S}')$ is shallow if α_c is large. We set α_c experimentally as follows: we computed $D(\boldsymbol{R}, \mathcal{S}, \mathcal{S}')$ for a few test cases by changing α_c from a large value until the significant peak in the distribution of $D(\boldsymbol{R}, \mathcal{S}, \mathcal{S}')$ appears. We selected the appropriate value of α_c as the one yielding the sharpest peak of $D(\boldsymbol{R}, \mathcal{S}, \mathcal{S}')$ for the test cases. In practice, α_c is directly related to the average variation of curvature on typical objects.

In our previous work [5], we used the sum of squared differences as a distance measure that is minimum for the best rotation. We use here a similarity measure rather than a distance measure for technical reasons but the results are the same.

One technical difficulty is that the sum in Equation (1) has taken over the entire spherical image even though only a subset of the spherical image should be used, that is, the part that corresponds to the part of the object visible in the range image. In practice, the algorithm used for fitting the mesh to the surface of the object can identify which points of the mesh are computed from actual range data and which points are interpolated. This information is used to determine which subset of the spherical image corresponds to actual data and should be included in the summation. In the rest of the paper, we will write the similarity as a sum over the entire sphere as in Equation (1) but it should be clear that, in the actual implementation, the sum is taken only over the visible subset of the sphere.

Our construction of the discrete mesh and of its associated spherical image allows us to generalize the concept of correlation of two 2-D sets of points to the concept of correlation of two 3-D surfaces. In particular, we have reduced the difficult problem of finding a six-degree of freedom transformation to the easier problem of finding a rotation between spherical images.

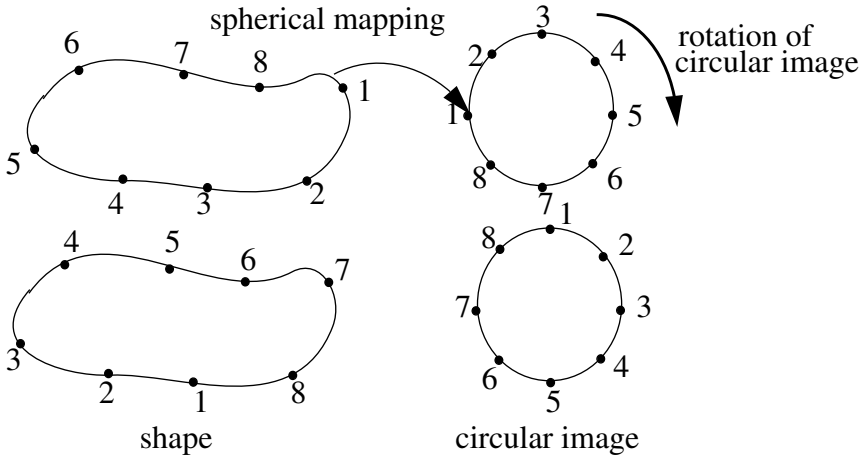

Figure 1 : Indexing nodes on a 2-D curve: Two different regular discrete approximations of a 2-D curve with the same number of points are related by a rotation of the indices on the unit circle.

3 Matching Using Color Data

Since the SAI representation defines a coordinate system over an 3D object surface, we can store not only geometric information but also photometric information such as color. In the work discussed here, we compute the hue at every point on the surface and store it at the corresponding node of the spherical mesh. The hue is the simplest characterization of color. In reality, a more sophisticated model of color formation should be used to account for color constancy, specularity, shadows, etc. However, our emphasis here is on the demonstration of the spherical mapping as a tool for registering object models using photometric information, a task for which the hue is sufficient.

Matching objects using color proceeds in the same manner as in the case of geometric data discussed in Section 2.3. The similarity between the spherical representations of the model and of the observed data is computed for the possible rotations of the unit sphere. The similarity is defined as a function of the sum of the squared differences of the hue value h at every point of the spherical representation instead of the sum of the differences of the curvature values:

$$D(\boldsymbol{R},S,S') = \sum_{N \in S} \left(w \times exp\left(-\frac{1}{2} \left(\frac{h(N) - h'(R(N))}{\alpha_h} \right)^2 \right) \right) \tag{2}$$

where α_h is a constant value that defines a sensitivity of the similarity to the differences. As in the case of geometric data, α_h is chosen experimentally by considering an average variation of hue values of several objects. w is a weight value that takes into account intensity and saturation at each node. Specifically, it is necessary to avoid influences of the nodes at which there is little photometric information. In fact, w is computed as a function

of both intensity and saturation at the corresponding nodes between the model data and the observed data. That is, w becomes large at the node that has a large saturation value and a large intensity value.

As in Section 2.3, the best rotation between $\mathcal{S}$ and $\mathcal{S}'$ is the one that realizes the maximum of D. Once the rotation between $\mathcal{S}$ and $\mathcal{S}'$ is computed, the full transformation between the original surfaces $\mathcal{M}$ and $\mathcal{M}'$ is computed as before by establishing the correspondence between mesh nodes and by minimizing the distance between corresponding nodes.

Let us consider a cylindrical object to verify that this algorithm enables us to correctly register surfaces. Because of the axis of symmetry of the cylinder, it is impossible to recover the pose of this object in an image using shape information alone. However, the pose is correctly recovered using the mapping of hue onto the intrinsic spherical representation of the surface as shown in Figure 2. Figure 2(a) shows the image of a cylinder that has significant photometric features on the surface. Figure 2(b) shows the model of the object as a mesh of points, which is the dual of the 7th subdivided icosahedron containing 980 faces. Figure 2(d) shows the result of mapping hue onto the SAI coordinate system. In this display, the hue is indicated by gray level intensity. The hue is encoded using the scale of Figure 2(c). We will use this scale in all the figures in the remainder of the paper. Also, in the rest of the paper, the SAI will always be displayed as in Figure 2(d), i.e., by placing each node at a distance from the center propotional to the curvature. Here, the curvature distribution over the SAI coordinate system is uniform around the axis direction and no meaningful result can be obtained by correlating the curvature spheres. However, the hue distribution exhibits clear features due to the uneven distribution of hue on the original object.

Figure 2(e) shows the graph of the similarity between spheres of different poses as a function of rotation angles for hue mapping. Since the space of rotations is three-dimensional, we display a two-dimensional view of $D(\boldsymbol{R}, \mathcal{S}, \mathcal{S}')$ by plotting for each of the values of two of the rotation angles, say φ and θ, the maximum value of D over all possible values of the third angle ψ. In the rest of the paper, the graph of similarity measure will be displayed as in Figure 2(e). The units in the graph are 10° in both φ and θ. The range of angles is 0 to 360° in φ and 0 to 180° in θ. Here, we include the graphs of D over the entire space of rotations in order to illustrate its behavior over the entire space. In practice, it is not necessary to explore the entire space at full resolution.

Figure 2(f) shows the result of the matching as a superimposition of the model data and the observed data using the transformation computed from the optimal rotation between spheres. One can see that the correct rotation about the axis of symmetry has been computed. In fact, we took two range images by rotating the object on the rotational table. The rotation angle computed by this matching result is 30.5° where the actual rotation angle is 30.0°. There are two reasons for the residual 0.5° error. First, because of the discretization of a mesh, there is a minimum distance between the node on the model surface and its corresponding node on the other. Second, we assume that the hue value is a photometric invariance; however, the variance is slightly dependent on the direction of the illumination despite the color compensation using a reference object under the illumination.

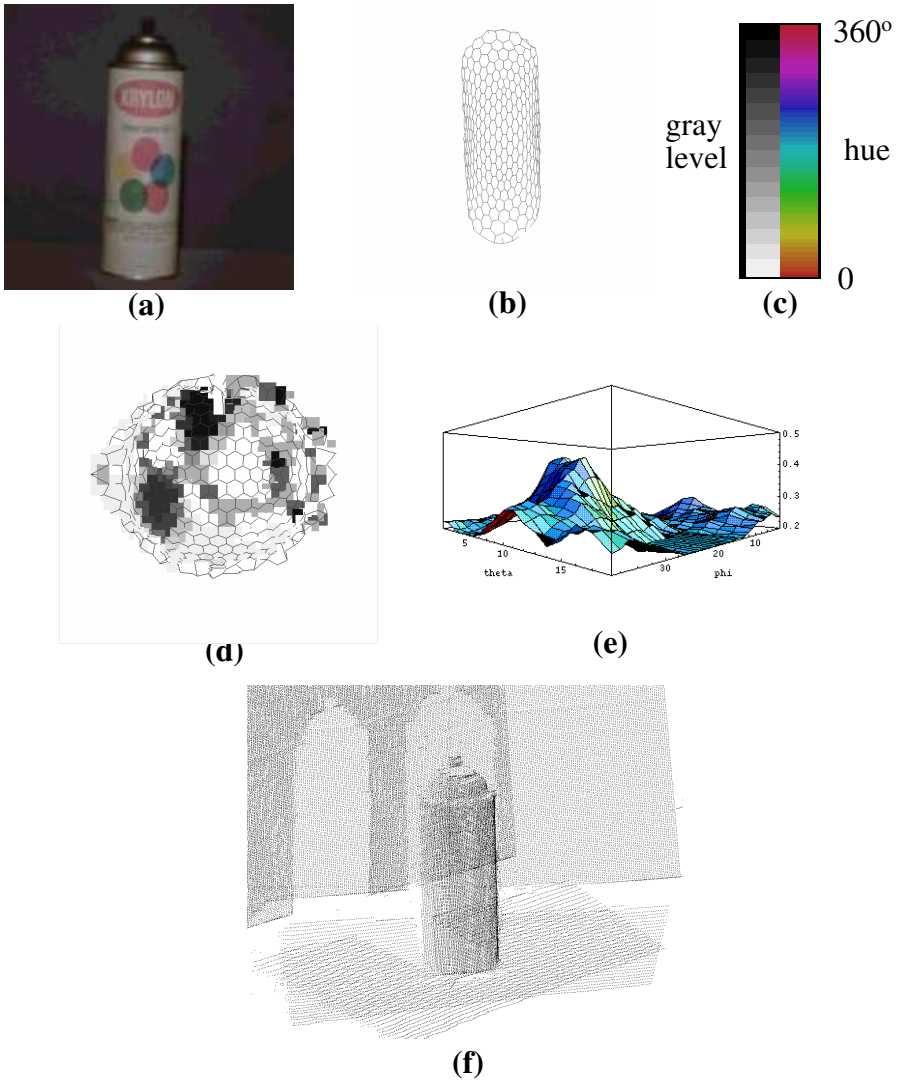

Figure 2 : SAI matching based on color information: (a) Original cylinder; (b) The object is approximated by a regular mesh with N = 980 nodes; (c) Scale of hue; (d) Mapping of hue onto the spherical representation; (e) Graph of similarity function using color; (f) Superimposition of the two registered models using color; The correct pose of the model in the scene is found even though the object has an axis of symmetry.

This example highlights the key features of using the SAI coordinate system for matching surfaces which we stated earlier in the Introduction. First of all, the matching algorithm does not require segmenting the color image or extracting color features, a difficult problem in general. Instead, the algorithm relies solely on the hue distribution on the surface. One consequence is that the algorithm does require combinatorial search for the best combination of feature matches as is the case in most conventional approaches. Secondly, the algorithm can recover full three-dimensional object pose without requiring a mathematical model of the object surface because the underlying model, the discrete mesh, is independent of any mathematical characterization of the surface. Finally, the algorithm can handle partial views and occlusions because of the properties of the spherical mapping.

4 Combining Photometric and Geometric Information

The purpose of combining photometric and geometric data is to take advantage of the type of data that best yields discrimination between object poses. Therefore, we need to evaluate at every node two quantities that measure the amount of information from both types of data and we need to combine them into a single number which is included in the similarity between spherical images, D. We have to solve two problems, quantifying the amount of information at a node, and scaling values from different sources of data in order to compare them.

4.1 Combining Similarity Measures

In order to address these problems, we need to define some notations. We denote the number of points on the mesh by N. We denote by c_i and h_i the values of curvature and hue at a node i of a spherical model $\mathcal{S}$, respectively. In order to simplify notations, we omit the rotation $\boldsymbol{R}$ from the notations when there is no ambiguity. Specifically, we denote by h_i the hue of a node $\boldsymbol{N}$ of $\mathcal{S}$ and by h'_i, using the same index i, the hue value of the corresponding node $\boldsymbol{R}(\boldsymbol{N})$ of $\mathcal{S}'$. Similarly, we will write the similarity $D(\boldsymbol{R},\mathcal{S},\mathcal{S}')$ between two spheres $\mathcal{S}$ and $\mathcal{S}'$ given a rotation $\boldsymbol{R}$ as $D(\mathcal{S}, \mathcal{S}')$, omitting the letter $\boldsymbol{R}$. It should always be clear in both cases that the indices and the similarities are defined with respect to a rotation. We denote $D_c(\mathcal{S},\mathcal{S}')$ (resp. $D_h(\mathcal{S},\mathcal{S}')$) the similarity between $\mathcal{S}$ and $\mathcal{S}'$ computed using curvature (resp. hue). With these notations, the similarity measures are defined as:

$$D_c(S, S') = \sum_i exp\left(-\frac{1}{2}\left(\frac{c_i - c'_i}{\alpha_c}\right)^2\right) \tag{3}$$

$$D_h(S, S') = \sum_i \left(w_i \times exp\left(-\frac{1}{2}\left(\frac{h_i - h'_i}{\alpha_h}\right)^2\right)\right) \tag{4}$$

With these definitions, we combine D_c and D_h into a composite similarity D by taking a weighted sum:

$$D(S, S') = \lambda D_c(S, S') + (1-\lambda) D_h(S, S') \tag{5}$$

In this definition of D, λ is a scalar between 0 and 1 which characterizes the respective contributions of the curvature and hue distributions to the overall distance. Specifically, $\lambda = 1$ means that only the curvature is taken into account in the matching, while $\lambda = 0$ means that only the hue distribution is used.

The problem now is to select an appropriate value of λ. Clearly this number must depend on the data. We choose λ to be a function of the variance σ_c (resp. σ_h) with respect to $\tau\eta\varepsilon$ $\sigma\varepsilon\nu\sigma\iota\tau\iota\varpi\iota\tau\psi$ α_c (resp. α_h). In addition, we add the coefficient w which is computed by averaging weight values taking into account intensity and saturation. The variance σ_c (resp. σ_h) are computed by averaging the local variance α_i^c (resp. α_i^h) around each node over the entire sphere individually. The local variance α_i^c (resp. α_i^h) at node i is computed

by using a neighborhood K_i which consists of the node and its three neighbors. In practice, a larger neighborhood K_i could be used. We denote by N_i the size of the neighborhood. With these definitions, we have the following equations:

$$\lambda = \frac{\left(\frac{\sigma_c}{\alpha_c}\right)^2}{\left(\frac{\sigma_c}{\alpha_c}\right)^2 + \left(w \times \left(\frac{\sigma_h}{\alpha_h}\right)\right)^2} \tag{6}$$

$$\sigma_c = \frac{\sum_i \alpha_i^c}{N} \qquad \sigma_h = \frac{\sum_i \alpha_i^h}{N} \qquad w = \frac{\sum_i w_i}{N} \tag{7}$$

$$\alpha_i^c = \sum_{j \in K_i} \frac{c_j^2}{N_i} - \left(\left(\sum_{j \in K_i} c_j\right)/N_i\right)^2 \qquad \alpha_i^h = \sum_{j \in Ki} \frac{h_j^2}{N_i} - \left(\left(\sum_{j \in K_i} h_j\right)/N_i\right)^2 \tag{8}$$

The coefficients σ_c and α_c can be interpreted as follows: aic quantifies the variation of curvature around node *i,* which is a measure of the quantity of information carried by the curvature at that node *j* is an estimate of the information carried by the curvature over the entire surface. A small value of σ_c means that there is, on average, little variation of curvature over the surface. The interpretation is the same for σ_h and α_h.

The intuitive interpretation of this choice of λ is straightforward: If the values of ${\alpha_i}^c$ are small compared to the values of ${\alpha_i}^h$, then the distance D_c is scaled up by choosing λ close to 1. Conversely, if the values of ${\alpha_i}^c$ are large compared to the values of ${\alpha_i}^h$, then D_c is scaled down by setting λ close to 0.

4.2 Experimental Evaluation

Two criteria should be used in evaluating the combined similarity D. First, we need to make sure that, if the best pose is found correctly from either curvature or color alone, then it will be found as accurately by using a combination of the two. In other words, combining curvature and color should not be worse than using each individually. The second criterion is that the determination of the rotation between the spherical representations is, in general, better with the combined similarity than with each similarity individually. This criterion ensures that D_s and D_h are correctly scaled by λ.

In order to evaluate the performance of the matching algorithm based on these two criteria, we ran the matching algorithm on several test objects. Most of the test objects are deliberately chosen to be extremely simple: painted cylinders, ellipsoids, and spheres. This is because we want to evaluate the algorithm in extreme cases in which there is little or no geometric information. Although these objects are unexciting from an object recog-

nition standpoint, they are necessary for our experiments. We also include in this section an example of matching using a real object with a complex shape.

To evaluate the first criterion, let us consider the cylindrical object shown in Figure 2(a). As we denoted in Section 3, the hue distribution mapped on the SAI representation has more significant information than the curvature distribution (Figure 2(c)). Figure 3(a), (b) and (c) show the graphs of D_c, D_h and D for this object, respectively. In this example, the same optimal rotation is found in both (b) and (c). The computed rotation angle is 30.5^o where the actual rotation angle is 30.0^o. Since the hue variation is larger with respect to the curvature variation, λ is computed at 0.36 using Equation (6).

To evaluate the second criterion, let us consider our second object (Figure 4). Figure 4(a) shows the image of a plastic egg, in which the surface is split in two halves, one painted red, the other one painted green. Figure 4(b) shows the mapping of hue onto the SAI representation. Since the object is symmetrical in both shape and appearance, it is theoretically impossible to register the pose by using either curvature or color information individually. Figure 4(c), (d) and (e) show the graphs of D_c, D_h and D for this object, respectively. The best rotations which are computed by the maximum in the graphs of D_c and D_h are both wrong because of the symmetries. However, we can get the right solution by using the combined similarity. The rotation angle computed by using the combined similarity is 18.8^o where the actual rotation angle is 20.0^o. Since both curvature and hue variations are small, λ is computed at 0.56 by using Equation (6).

To ensure that this algorithm is applicable to an object that has a more complicated shape, let us consider our third object (Figure 5). Figure 5(a) and (b) show the images of a toy duck in two different poses. Figure 5(c) shows the model of the object as a mesh. Figure 5(d) shows the mapping of hue onto the SAI representation. Figure 5(e), (f) and (g) show the graphs of D_c, D_h and D for this object, respectively. Figure 5(h) shows the superimposition of both model data and observed data which is transformed by using the matching result of the combined similarity. Figure 5(i) shows the errors of the estimated pose between the original range data and the transformed model. The length of the line at each node indicates the distance between the node on the mesh and the data point that is closest to the node. The average error over the entire mesh is 0.13mm where the accuracy of this range finder system [11] is 0.1mm. Since we captured two different poses arbitrarily, we cannot estimate the actual rotation angle. However, according to this result, this algorithm enables us to estimate the pose precisely. Table 1 summarizes the error statistics for this object. In this example, λ is computed at 0.89 using Equation (6). The value is higher than in the previous two cases because the object has a more complicated shape.

In order to check the matching between a complete model and a partial model using both color and curvature, let us consider our fourth object (Figure 6). Figure 6(a) shows the images of a painted ball which are obtained by using three calibrated range finders. Figure 6(b) shows another pose using one of the range finders. Figure 6(c), (d) show the mapping of hue onto the object mesh for each pose. Figure 6(e) shows the graph of D for this object. Figure 6(f) shows the transformed object mesh computed from Figure 6(d) by using the matching result. To evaluate this result, we compute the rigid transformation by calculating the gravity point of each marked region and specifying the correspondence between marks manually. The rotation angle by using both color and curvature informa-

tions is 94.5^{o} where the rotation angle from the manual estimation is 92.7^{o}. Since there is no geometric information, there are three rotational degrees of freedom which is the reason why there is such a large error comparing previous cases. In this example, λ is computed at 0.03 using Equation (6), since there is no curvature information for this object.

These examples show that our determination of λ does take explicitly into account the relative distribution of information of the surface. Table 2 summarizes the coefficients computed in these examples.

Table 1: Error statistics

	color	curvature	both
Min. error	0.010	0.014	0.006
Max. error	0.676	0.651	0.665
Std. σ error	0.097	0.082	0.089
Ave. error	0.150	0.132	0.135

Table 2: Summary of coefficients

	cylinder	egg	duck	ball
σ_c	0.0305	0.0228	0.0931	0.0207
σ_h	8.47	4.54	9.35	23.3
w	0.959	0.889	0.668	0.888
λ	0.360	0.562	0.898	0.038

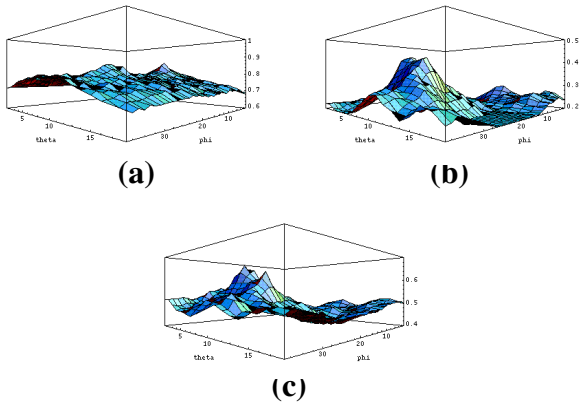

Figure 3 : Combining color and curvature information-1: (a) Similarity using curvature only; (b) Similarity using color only; (c) Combined similarity

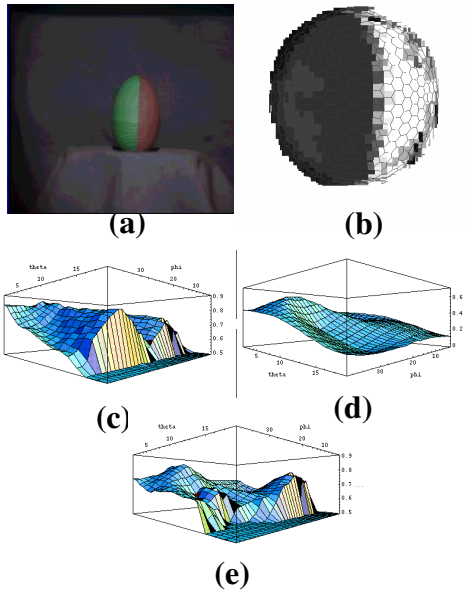

Figure 4 : Combining color and curvature information-2: (a) Image of a painted egg; (b) Mapping of hue onto spherical representation; (c) Similarity using curvature only; (d) Similarity using color only; (e) Combined similarity

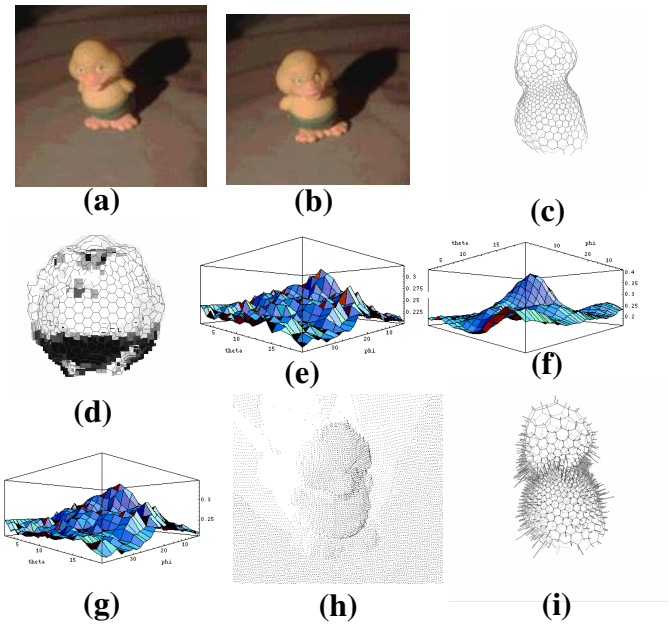

Figure 5 : Combining color and curvature information-3: (a), (b) Image of a plastic toy (two poses); (c) Deformed mesh; (d) Mapping of hue onto spherical representation; (e) Similarity using curvature only; (f) Similarity using color only; (g) Combined similarity; (h) Superimposition of original range data and estimated pose; (i) Matching errors of the estimated pose

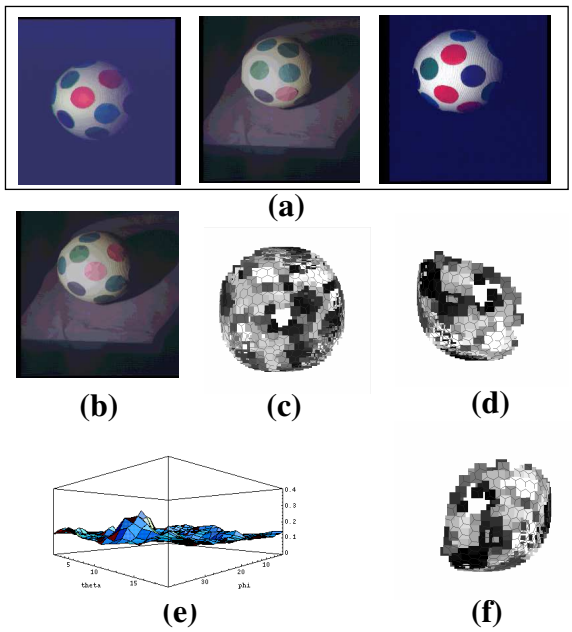

Figure 6 : Matching between a complete model and a partial view: (a) Original images measured from three different views; (b) Image of partial view; (c) Mesh of complete model mapped hue values; (d) Mesh of partial model mapped hue values; (e) Matching using color and curvature; (f) Transformed mesh using the matching result;

5 Conclusion

In this paper, we have shown how photometric information can be incorporated in the SAI representation in addition to purely geometric information. We have shown that the very same algorithm that was used for surface matching based on geometry can be used for surface matching based on color distribution. The only difference is that the SAI stores curvature in one case, and hue in the other. In addition, we have shown that it is possible to combine geometric and photometric information in a way that takes into account the distribution of both on the object. This is done by defining the similarity between the SAIs of two objects as the weighted sum of the similarity measures computed from the shape information and from the photometric information. We have developed a weighting scheme that gives the highest weight in the evaluation of the similarity between SAIs to the type of information that is most significant for a given object. For example, the photometric term of the similarity measure between SAIs is weighted more heavily if the surface of an object includes a large nearly spherical region in which there is no shape information. Conversely, the geometric term is weighted more heavily if the color of the object is nearly uniform across the surface.

The combined shape/appearance matching algorithm has several important properties. First, it merges the two types of information in a natural way whereas traditional techniques use shape and photometric information separately. Second, the matching algorithm is applicable to a wide class of curved objects since the underlying mesh fitting algorithm used for building the SAI representation makes few assumptions on the shape of the object; and, the representation can handle partial views and occlusions owing to its property of connectivity conservation. Finally, the algorithm does not rely on arbitrary thresholds or scaling coefficients because all the parameters, for example, the weights between the two similarity measures, are computed dynamically from the actual data.

Color is only one type of appearance information that can be used. Other types of information, such as various measures of texture, could also be used in matching surfaces. An natural extension of this work is its generalization to other types of information. Another extension is the use of multiple appearance attributes instead of a single one, in our case, the hue. This involves refining the weighting scheme to take into account the relative contributions of multiple attributes.

6 References

[1] Besl, P., Kay, N.D., “A Method for Registration of 3-D Shapes”, PAMI-14(2), 1992.

[2]Champleboux, G., Lavallee, S., Szeliski, R., Brunie, L., “From Accurate Range Imaging Sensor Calibration to Accurate Model-Based 3-D Object Localization”, Proc. CVPR-92, pp. 83-89, 1992.

[3]Chen, Y., and Medioni, G., “Object Modelling by Registration of Multiple Range Images”, Image and Vision Computing, 10(3), April 1992.

[4]Delingette, H., Hebert, M., Ikeuchi, K., “Shape Representation and Image Segmentation Using Deformable Surfaces”, Image and Vision Computing, 10(3), April 1992.

[5]Delingette, H., Hebert, M., Ikeuchi, K., "A Spherical Representation for the Recognition of Curved Objects", ICCV'93, Berlin, 1993.

[6]Ikeuchi, K., "Recognition of 3-D objects using the extended Gaussian image", IJCAI-81, pp. 595-600, 1981.

[7]Kamgar-Parsi, B., Jones, L.J., Rosenfeld, A., "Registration of Muliple Overlapping Range Images: Scenes Without Distinctive Features", PAMI-13(9), 1991.

[8]Kang, S.B., Ikeuchi, K., "Determining 3-D Object Pose Using the Complex Extended Gaussian Image", Proc.CVPR-91, 1991.

[9]Martin, W.N., Aggarwal, J.K., "Volumetric Descriptions of Objects from Multiple Views", PAMI-5(2), 1983.

[10]Parvin, B., Medioni, G., "B-rep from Unregistered Multiple Range Images", Proc. IEEE Robotics and Automation, 1992.

[11]Sato, K., Yamamoto, H., Inokuchi, S., "Range Imaging System Utilizing Nematic Liquid Crystal Mask", Proc. ICCV-87, 1987.

[12]Stein, F., Medioni, G., "TOSS-A System for Efficient Three Dimensional Object Recognition", Proc. DARPA Image Understanding Workshop, 1990.

[13]Zhang, Z., "Iterative Point Matching for Registration of Free-Form Curves", Research Report 1658, INRIA, March 1992.