Measures for Silhouettes Resemblance and Representative Silhouettes of Curved Objects

Yoram Gdalyahu and Daphna Weinshall

Institute of Computer Science, The Hebrew University, 91904 Jerusalem, Israel

Abstract. We claim that the task of object recognition necessitates a measure for image likelihood, that is: a measure for the probability that a given image is obtained from a familiar (pre-investigated) object. Moreover, in a system where objects are represented by 2D images, the best performance is achieved if those images are selected according to a maximum likelihood principle. This is equivalent to maximum stability of the image, or minimal change under a viewpoint perturbation. All of those qualities involve a quantitative comparison of similarity between images. We propose different metric functions which can be imposed on the image space of curved three dimensional objects. We use these metrics to detect the representative views (most stable and most likely views) of three test models. We find the same representative views under all the investigated metrics, suggesting that local maxima of stability and likelihood are metric independent. Our method of image comparison is based solely on the appearance of the occluding contour, hence our method is suitable for object recognition from silhouettes.

1 Introduction

Three dimensional objects can be projected to an infinite number of images due to the infinite number of possible viewing directions. In the absence of any particular symmetry, all those images may be different from each other, but some of them can be thought of as better representing the three dimensional object. This leads to the intuitive notions of "canonical", "characteristic" or "generic" views, which frequently appear in the literature. Recently, an objective definition has been given for those notions [14, 13]. According to it, a representative view has the properties of maximum likelihood (it is similar, as an image, to a large number of other images), and maximum stability (it doesn't vary a lot when the viewing direction is perturbed). We have already shown that a local maximum of the likelihood function is obtained at the same points (views) for which a local maximum of the stability function is obtained [13]. Therefore, the same single view is (locally) both the most likely and the most stable one, and it is naturally proposed as a representative view.

Explicit identification of such views requires an explicit quantitative measure of similarity between images. Such a measure converts the image space into a metric space, where the distance between every two images is proportional to their dissimilarity. Finding an appropriate metric is a research subject of its own [6], and for complicated images it is still an open question how to measure their resemblance. It becomes easier to measure image similarity if the family of the objects is constrained. For example, a discrete object which is made of a cloud of points. This case and other simple cases are reviewed in section 2. Our present work is devoted to the study of the case of curved objects (not necessarily convex) which lack intrinsic feature points. We discuss and compare some possible metric functions to judge the similarity between images of such objects. We apply our proposed metrics to three different test models, perform a numerical study of their views stability, and compute their representative views.

Representative views play an important role in viewer-centered approaches to 3D shape representation, where three dimensional information is not represented explicitly. In these approaches a novel view of the object is recognized by comparing it to stored views, and storing "good" views is essential for successful performance. If an angle is to be recognized, Ben-Arie [2] and Burns et al. [5] already identified the more common views of it (see section 2). A related measure of view genericity, or view likelihood, also underlies the Bayesian decision strategy suggested by Freeman [7] to interpret correctly ambiguous scenes. Freeman showed how to *use* measures of view "genericity" in image understanding, but did not propose a general way to *compute* such measures directly from images and models.

The rest of the paper is organized as follow: In section 2 we summarize our definitions of stability and likelihood, for a given image metric. We also present a few early results obtained with these definitions. In section 3 we propose metric functions which are suitable for the case of curved objects. The application of these metrics to identify representative views of some 3D models is presented in section 4. In section 5 we draw our conclusions.

2 Stability and likelihood of views

2.1 Definitions

Our framework consists of a weak perspective projection scheme, hence the image space of a given object is two dimensional, spanned by the two angles ϑ and φ, respectively specifying the elevation and azimuth of the camera axis with respect to the center of the object. The image space can be therefore mapped to a unit sphere, where each point on that sphere is called a view, and each such point represents a set of images related to each other by a 2D similarity transformation (2D rotation, translation and scale). We call this sphere "the viewing sphere".

The parameterization of the viewing sphere demands some preselection of a pole, relative to which the angles of elevation and azimuth are measured. Since the definitions of stability and likelihood that we are to present involve the comparison between neighboring views, it is useful to measure the elevation and azimuth angles of one view with respect to the other. Let $V_{\vartheta,\varphi}$ be the view whose elevation ϑ and azimuth φ are measured with respect to the z axis pole, and let $V_{\delta,\psi}^{\vartheta,\varphi}$ be the view whose elevation δ and azimuth ψ are measured with respect to the previous view.

By definition, the geodesic distance on the viewing sphere between $V_{\vartheta,\varphi}$ and $V_{\delta,\psi}^{\vartheta,\varphi}$ is δ. However, the similarity between these two views, as 2D images, is not necessarily proportional to the geodesic distance, and we denote their image distance by $d(\vartheta,\varphi,\delta,\psi)$. This image distance can change abruptly with δ, usually if features appear or disappear. In regions of continuous behavior the image distance can be expressed as $d(\vartheta,\varphi,\delta,\psi) = D(\vartheta,\varphi,\psi)\delta + O(\delta^2)$, where $D(\vartheta,\varphi,\psi)$ is the derivative of the distance function with respect to δ in the direction ψ, and $D(\vartheta,\varphi,\psi) \geq 0$

The variability of a view reflects the amount of image change caused by a perturbation of the viewing direction. More precisely, we define the ε-variability of a view $V_{\vartheta,\varphi}$ as the average change in the images which belong to an ε-neighborhood of $V_{\vartheta,\varphi}$:

$$R_\varepsilon(\vartheta,\varphi) = < d(\vartheta,\varphi,\delta,\psi) >_{0\leq\delta\leq\varepsilon,\ 0\leq\psi\leq 2\pi} = \frac{\int_0^{2\pi}\int_0^{\varepsilon} d(\vartheta,\varphi,\delta,\psi) sin(\delta) d\delta d\psi}{\int_0^{2\pi}\int_0^{\varepsilon} sin(\delta) d\delta d\psi}$$

The smaller the view variability is - the grater is its stability. In [14] the ε-stability $s_\varepsilon(\vartheta,\varphi)$ of a view was defined simply as $1/R_\varepsilon$. Here we are primarily interested in the comparison between different metric functions, so we prefer another definition which is scale invariant:

$$s_\varepsilon(\vartheta,\varphi) = \frac{R_\varepsilon^{max} - R_\varepsilon(\vartheta,\varphi)}{R_\varepsilon^{max} - R_\varepsilon^{min}}$$

The notion of stability captures a local property of a view, namely its level of invariance under a small change of the viewing position. On the other hand, the notion of likelihood is defined with respect to the whole image space, and it measures how common a view is. We start by simply "counting" the number of views which are at least ε-similar to a given view, according to some measure d. This yields the definition of an ε-area:

$$A_\varepsilon(\vartheta,\varphi) = \int_0^{2\pi}\int_0^{\frac{\pi}{2}} \chi_{\{d(\vartheta,\varphi,\delta,\psi)\leq\varepsilon\}} sin(\delta) d\delta d\psi$$

where χ_b is 1 when b is true, and 0 otherwise.

The total surface area of the viewing sphere is 4π, hence $A_\varepsilon/4\pi$ can be considered as the conditional probability $f(\varepsilon|\vartheta,\varphi)$ of an image error ε, given a reference view $V_{\vartheta,\varphi}$. Using Bayes rule we get the inverse relation:

$$f(\vartheta,\varphi|\varepsilon) = \frac{f(\vartheta,\varphi) f(\varepsilon|\vartheta,\varphi)}{f(\varepsilon)}$$

where $f(\vartheta,\varphi)$ is the prior distribution of camera positions. If the viewer position is not constrained, then $f(\vartheta,\varphi)$ is of course uniform. The part of $f(\vartheta,\varphi|\varepsilon)$ which depends on the viewing direction is defined as the view ε-likelihood $l_\varepsilon(\vartheta, \varphi)$:

$$l_\varepsilon(\vartheta, \varphi) = f(\vartheta, \varphi) f(\varepsilon|\vartheta, \varphi) \simeq f(\vartheta, \varphi) \int_0^{2\pi} \int_0^{\frac{\pi}{2}} \chi_{\{d(\vartheta,\varphi,\delta,\psi) \leq \varepsilon\}} sin(\delta) d\delta d\psi$$

We note that it is also possible to define both the stability and the likelihood in the limit of $\varepsilon \to 0$, as we do in [13]. This means that stability is measured in the limit of a zero perturbation of camera position, and likelihood is measured in the limit of zero image error.

2.2 General results and some special cases

It can be shown [13] that if the prior distribution of viewpoints of a given object is uniform, namely all camera positions are equally likely, then the most stable and the most likely views are the same view. In other words, the local maxima of the stability function and of the likelihood function are obtained for the same views.

In previous works we applied our definitions to simple objects, for which there is a natural metric function [13, 14]. The first case is that of objects represented by feature points in 3D space. A natural measure for the distance between two images is obtained by measuring residual distances between corresponding points, after the two images have been normalized and aligned. We used two different alignment schemes, with either an affine or a similarity image transformation, and obtained two metric functions (which are positive, symmetric and obey the triangular rule). We showed that under both of them, the most stable and likely view depends only on the three principal moments [1], and not on the specific location of all the features. Moreover, this view is unique and it is the "flattest view", obtained when the three dimensional object has its minimal spread along the viewing direction [2].

A second case which was studied is that of a planar angle that can be viewed from an arbitrary direction in space. This case was studied empirically by Ben-Arie [2] and Burns et. al [5], who found the "picking effect": the most probable value of an angle to be picked from a random viewing direction is the value of the angle itself. We imposed a natural metric on the image space of such an angle, namely the difference between the two projected angles, and we were able to find the same result analytically.

A third metric function is based on extraction of extreme curvature points along the occluding contour of an object. We discuss this metric in section 3.2. Here we note, that for the special case of an ellipsoid, we found analytically that under this metric the views which are local maxima of stability and likelihood are located along the three main axes of the ellipsoid.

[1] We defined the three principal moments of a discrete cloud of feature points as the three eigenvalues of their autocorrelation matrix.

[2] More precisely, if we assume that the z axis is the viewing direction, then the flattest view is the view for which the autocorrelation matrix is diagonal, with the three principle moments ordered in decreasing order.

3 Metrics to compare silhouettes

In this work the 3D objects are assumed to be topologically simple (without holes), which means that any image can be divided into two connected and non overlapping areas: the interior and the exterior regions. The border between those two areas is called the occluding contour, and the 3D curve which lies on the surface and is projected to the occluding contour is called the rim.

The shape of the occluding contour depends only on the object's shape and the viewing direction, in contrast to internal contours that may be sensitive to illumination. Therefore, the analysis of image geometrical variability here is solely based on the shape variations of the occluding contour. It goes without saying that all our objects are recognized according to their silhouettes.

Usually one wouldn't like to distinguish between images which differ by only an irrelevant image transformation. Here we assume a weak perspective projection, hence it is natural to choose the group of similarity transformations in the image plane (translation, rotation and scaling) as the group which divides the images into equivalence classes. Hence, All the images which belong to the same class (equivalent images) are related to each other by a 2D similarity transformation. Those images are represented by the same point on the viewing sphere. Moreover, when two images are not equivalent, the distance between them is determined by alignment, which means that the two closest images in the two respective equivalent classes determine the distance for all the members in those classes. Evidently, the distance between equivalent images is zero.

In the following sections we introduce the metric functions which we use in order to measure distances between views (closed curves). Two different strategies are taken: either a whole curve comparison (section 3.1), or comparison via interest points (section 3.2). The second mechanism permits a huge reduction of information due to the usage of only a zero measure set of points instead of a continuous line, but an undesirable result is the unification of equivalence classes into larger ones: two curves whose sets of interest points are related by a similarity transformation have a zero distance, and it is not necessary that this transformation maps the whole one curve onto the other. On the other hand, methods based on feature extraction more easily deal with occlusions and can be integrated with database indexing methods.

3.1 Whole contour comparison

Our first proposed method to compare two images (two planar closed curves), is simply a comparison of their tangent angle transforms. This transform is a representation of the curve by its tangent angle versus its normalized arc length [12, 1]. To construct such a representation, we start from an arbitrary point on the curve and follow it in an arbitrary direction, measuring in each point the angle of the curve tangent with respect to some arbitrary reference direction. While traveling along the curve we record this angle as a function of the accumulated arc length. We obtain a one dimensional graph which can always be transformed into a "canonical" form, which starts at the point [0,0] and ends at [1,1]. To achieve this form we may need to: (1) normalize the x-axis to one (equivalent to an image scaling operation), (2) reflect the function around its middle (equivalent to selection of travel direction along the curve), (3) shift the y-axis (equivalent to a rotation operation in the image plane), and (4) rescale the y-axis by $1/2\pi$ (which is no more than a cosmetic manipulation). We note that for the first operation to be valid we assume that the object is not occluded (see section 5). Figure 1 shows an example of an occluding contour extracted from an image, and its canonical form.

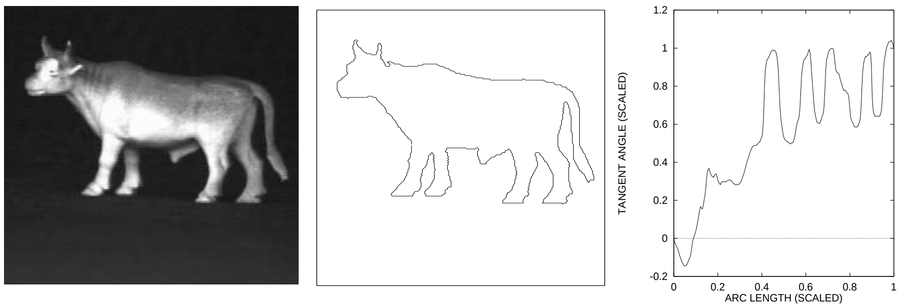

Fig. 1. Example of an occluding contour and a canonical form

A canonical form of a curve is not unique, since it depends on the arbitrary chosen starting point. However, The family of all such forms, obtained for all possible selections of the starting point, is invariant under a similarity transformation, and by sliding the origin along one such form, the family of all canonical forms is obtained. Hence for given two curves, a proper alignment algorithm performs a one dimensional search along one of them, to find the optimal starting point. The power of this alignment algorithm is that it combines the search for the proper rotation and translation into a one dimensional search of the starting point. It requires, however, having an algorithm to travel along curves.

The distance between canonical forms, which is to be minimized by the alignment procedure, is defined as follow: For a given desired resolution, the two functions are represented by two ordered sets of equally spaced points P_1 and P_2. Their distance is $min(d_{12}, d_{21})$, where d_{12} is the sum of distances of P_1 points from the spline formed by P_2, and similarly for d_{21}. The two sets are assumed to be of equal size.

3.2 Contour comparison using features

In the following we describe some methods of feature extraction that we use to locate interest points on a given smooth curve. The evaluation of distance between two curves is then composed of the following stages: feature extraction, alignment of feature points by a similarity transformation, and measuring the average residual distance between corresponding points. Note, however, the difference from the case of an object equipped with intrinsic features: here, two corresponding interest points in two different images *do not* correspond to the same single 3D point on the surface of the object.

Extreme curvature points Points of maximum (and perhaps also minimum) curvature along the curve are probably the most intuitive features that one can think of. However, extreme curvature points are known to be unstable [8] and hard to extract, due to their sensitivity to resolution and to the presence of noise. In general, a reliable extraction requires some smoothing mechanism, which enhances the extraction robustness and reduces its locality. Nevertheless extreme curvature points are the most local feature that we use (i.e, their position is determined by a small fraction of the curve around the point, which is advantageous in a case of occlusion).

By definition, curvature computation involves the second derivatives of the curve with respect to arc length. For digital curves, we find it more robust to use the following method of maximum curvature extraction: let p denotes one of the curve points, and let $\mathbf{u}_n$ and $\mathbf{v}_n$ denote the vector sums of the n vectors pointing from p to its n successive neighbors to the left and to the right. Then the angle between $\mathbf{u}_n$ and $\mathbf{v}_n$ is minimal when the curvature is maximal. This method avoids any computations of arc length or derivatives with respect to it. The vector summation gives more weight to distant points, for which errors due to the finite resolution are less important.

The polygonal approximation which is obtained by the extracted points may be refined by adding more interest points. At each refinement step, the curve point which is most distant from the polygonal approximation is added, until the desired accuracy is obtained. Figure 2 shows the interest points extracted by this method, together with the results of other extraction methods that we describe below.

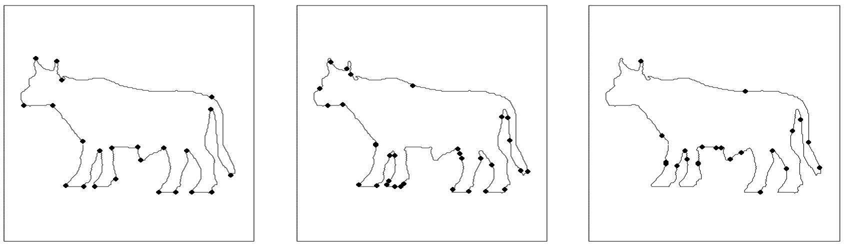

Fig. 2. Feature points which are extracted by maximum curvature plus refinement (left), by transformed self intersections (middle), and by unit signature (right).

Transformed self intersections The occluding contour, which in our view here is the only image content, is a simple closed curve which does not intersect itself. But it was recently shown [9] how such a curve can be transformed into another one, which does intersect itself. In this method the tangent angle transform (see above) is multiplied by a constant factor and a new curve, which in general does intersect itself, is reconstructed (figure 3). The exceptional conditions for which self intersections can't be made, or their number becomes infinite, are discussed in [9] and are ignored here.

The self intersection points can be mapped back to the original image. Note that each intersection is both a starting point and an ending point of a loop on the transformed curve, hence it is mapped back to *two* points on the original curve. Clearly, only the part of the curve which is enclosed between these two points has any influence on the intersection existence. Hence locality is controlled by the multiplication factor. More local features are obtained for larger multiplication factors, at the cost of losing robustness against noise.

Note that for the extraction of self intersections one only needs the first derivative of the curve (the tangent angle), while for the extraction of extreme curvature points the third derivative is needed.

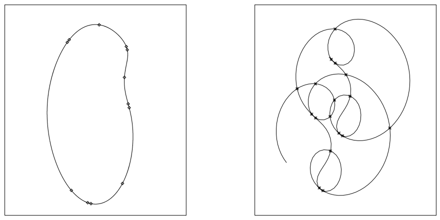

Fig. 3. An original curve (left) and the transformed curve (right) obtained with a multiplication factor of 3. To avoid losing intersections due to end effects, the transformed curve is obtained by tracing the original curve twice.

Points of unit signature The term "invariant signature" was assigned in [4] to functions that can be computed from a given curve, and have an invariance property under a certain transformation group. For example, had we been interested merely in rigid transformation, we could adopt the curvature function as a signature of the curve, and extract features by a certain curvature value. However, since the curvature function is not invariant with respect to scale, we could rely above only on its extreme points, but not on the function values themselves.

More sophisticated signature functions, that remain invariant under similarity, affine and projective transformations, can be found in the literature (see [3] and references therein). Some of them suffer from the involvement of high order derivatives, which is clearly an undesired property when using digitized curves. Our similarity invariant signature function is computed as follow: we place a right angle on a curve point s, such that the angle is symmetric with respect to the normal at s, and it is directed toward the interior region (figure 4). The right angle intersects with the curve at points $s1$ and $s2$, and the signature at s is defined as the ratio of lengths $||s1 - s||/||s2 - s||$. We extract points for which this ratio has the value of one.

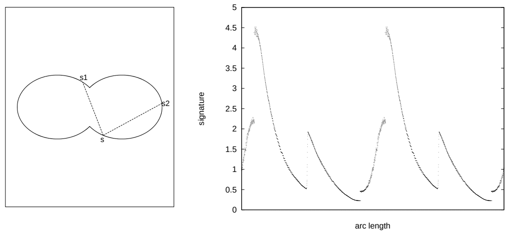

Fig. 4. Left: a right angle positioned on a given curve to measure the signature at that point. Right: the signature function obtained for the given curve.

The features locality can be controlled by the angle sharpness. Robustness against noise is achieved due to the fact that only the first derivative of the curve is needed, and there is no measurement of arc length.

4 Representative views of smooth objects

Imposing a metric on the image space of a three dimensional object is a crucial preliminary stage in the process of identifying representative views. In this work we examine the relation between representative views and the metric being used. We calculate numerically the views stability function for three different test objects, applying for each object the different metric functions which are proposed above. The investigated objects are: an ellipsoid (which obeys reflection symmetries), a sandglass model (which is axial symmetric), and a model of an eggplant (which lacks any kind of symmetry). All the models are represented by implicit functions and they are shown is figure 5.

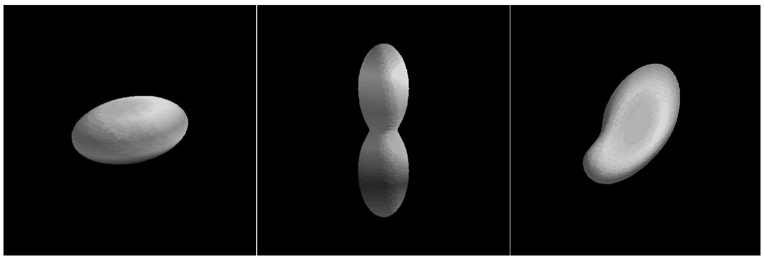

Fig. 5. The three test models: an ellipsoid, a sandglass and an eggplant.

According to our numerical simulations with different test objects and different metric functions, we conjecture that *the representative views do not depend* *on the metric being used* [3]. We find that the positions of the local maxima of the stability function are metric independent, hence the "best" viewing directions are an intrinsic property of the object.

We also find, as should be expected, an inverse relation between robustness and locality: using extreme curvature points is most sensitive to noise, while the other proposed methods have greater noise resistance.

4.1 Ellipsoid model

The ellipsoid is represented by the implicit function $x^2+2y^2+5z^2 = 1$, hence the three axes x, y and z are axes of reflection symmetry. Figure 6 shows the views stability function as calculated numerically using extreme curvature points. It is evident that the function has six local maxima in the entire image space (the two maxima at the poles appear in the graph as two lines). The six local maxima are in fact three pairs of points, since the stability function must obey the same reflection symmetries as the ellipsoid itself. Furthermore, each one of those pairs of viewing points is located on one of the ellipsoid axes. It turns out that the most stable view, which corresponds to the highest local maximum, is obtained when the ellipsoid is viewed along its minor axis (z). The most stable view has minimal depth and maximal image area. These results are in agreement with the analytical study reviewed in section 2.2.

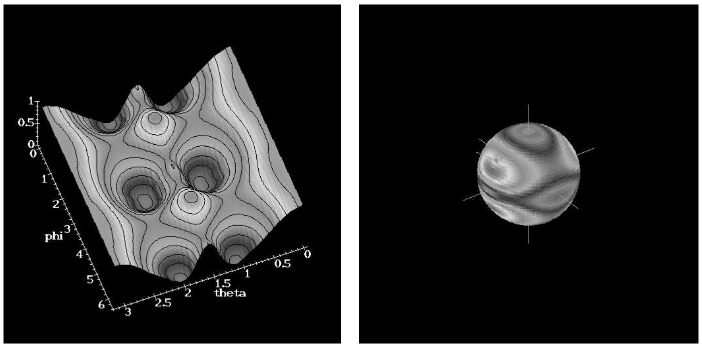

Fig. 6. Stability of ellipsoid views (also plotted on the viewing sphere). Note that the gray levels are not ordered from bright to dark, to enhance the contrast in the figure.

The relation between representative views and maximal image area was pointed out in the past on psychophysical ground. However, assigning to views a "genericity level" which is proportional to their apparent area, necessarily leads to the identification of the least representative view as the one with the minimal area. In the ellipsoid case, the minimal area is obtained when the viewing point is located on the main axis line (x), in sharp contrast to its being locally stable and likely.

[3] We exclude here metrics which are not related to the image content, like the discrete metric that assign a zero distance between an image to itself and a unit distance between any non identical images. For such metrics our conjecture certainly cannot hold.

Using the other metrics to estimate views stability, we get the same results regarding the position of local maximum points. The representative views are therefore invariant under the four examined metrics. The question of which metric to use should be addressed according to the application (section 5).

4.2 Sandglass model

The second object that we consider is represented by the implicit function $(x^2 + y^2 + 0.3z^2 - .09) * (x^2 + y^2 + 0.3z^2) - 0.3z^2 = 0$, and we refer to it as a model of a sandglass (see figure 5). In this representation, the z axis is an axis of rotation symmetry, so views which have the same altitude angle ϑ are all equivalent and are equally stable. It is therefore sufficient to investigate the views stability with respect to this angle alone, as is shown in figure 7 (transformed self intersection points are used). Again, using any one of our proposed metrics

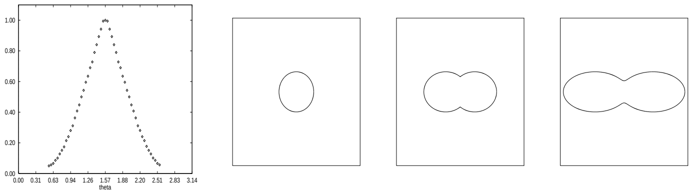

Fig. 7. Sandglass: stability of views and some images obtained for $\vartheta = 0, \pi/6$ and $\pi/2$ (most stable)

produced the same result, that the most representative view of the sandglass model is obtained (like before) when the principle axis of the object lies in the image plane (figure 7).

We note that with perfectly circular curves, all the proposed methods for feature extraction must fail, since the curvature and signature functions become uniform, and the number of transformed self intersections approaches infinity. The stability value for views which are close to the z axis could therefore be computed only with the method of a whole curve comparison, since for those viewing directions the object appears as a circle.

4.3 Eggplant model

The last test object is a model of an eggplant, which is represented by a fourth degree polynomial whose 35 coefficients were fitted experimentally to a real eggplant [10]. The eggplant model lacks any kind of symmetry, except the trivial reflection symmetry of silhouettes which is obtained for any two antipode viewing directions.

The stability map and the corresponding representative and non representative views are shown in figure 8. As can be seen, our results seem to agree with human intuition. Similar results are obtained with the different metric functions, but we note that with the metric based on extreme curvature extraction it was almost impossible to get reliable results with the given amount of noise in the

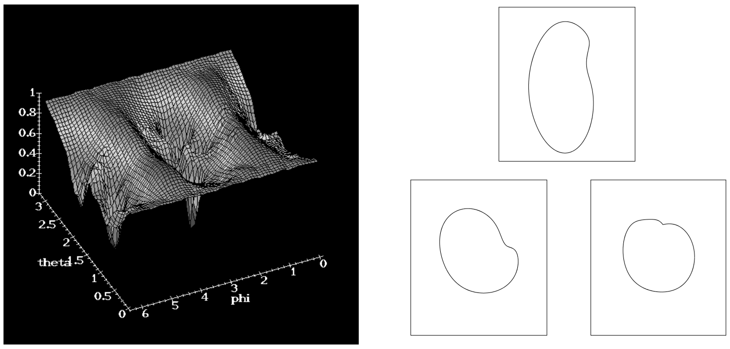

Fig. 8. Stability of the eggplant views. The upper view on the right is found as most stable (hence representative) while the other two views are in local minima of stability.

images. (The origin of noise in our case is the limit put on the accuracy of image generation from the implicit function model. Due to the complexity of the model and the large number of images needed to be generated during the simulation, we chose a rather coarse image resolution, which had introduced noise).

5 Summary and conclusions

Computing the likelihood of views of three dimensional objects can significantly enhance existing object recognition techniques. The question that should be answered by an object recognition algorithm is the identity of an object which appears in an image, but this question is frequently ambiguous, and should be answered statistically. The framework of maximal likelihood decision provides the appropriate statistical tool, and leads to the following obvious rule: among all the possible interpretations of a (noisy) image, select the one which is the most likely.

In this work we demonstrated how to use the quality of view stability to identify characteristic views of general curved objects, that can be recognized according to their silhouettes. Very simple objects could be dealt with analytically, but for more general cases we used numerical methods. We estimated the views stability function for three qualitatively different test models, and obtained their representative views according to the points of local maximum of stability.

Likelihood and stability cannot be evaluated without a quantitative measure of similarity between views. We based our similarity measure on the shape variation of the occluding contour, which depends only on the object's shape and the viewing direction, and does not depend on illumination. Four different metric functions were proposed, either using a whole curve comparison or using feature points. We found the same representative views under all metrics, hence we conjectured that *the representative views do not depend on the metric being used* (as long as the metric is "reasonable"), hence representative views are an intrinsic property of the object.

Since the selection of a metric does not seem to alter the most likely views, the important criteria for a metric selection become its locality and noise resistance. These two properties were shown to be inversely related, as one would expect. Locality is important as it best resists occlusions, where the given contour is not complete. In this work we didn't consider cases of occlusion, hence we could normalize the perimeter of a given curve to a unit length (section 3.1). This operation is crucial only for the metric 3.1 (a whole contour comparison). The other metric functions can be applied also to non normalized curves, and therefore can deal with occlusions.

Another benefit of comparison methods based on feature extraction is the possibility to combine them with database indexing methods [11]. The extracted features can serve not only for distance evaluation against stored arrays of feature points in the memory, but their invariant coordinates can also be used to generate pointers into the database, to enhance the search.

References

1. E. M. Arkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, and J. S. B. Mitchell. An efficiently computable metric for comparing polygonal shapes. *T-PAMI*, pages 209–216, 1991.
2. J. Ben-Arie. The probabilistic peaking effect of viewed angles and distances with application to 3-d object recognition. *T-PAMI*, pages 760–774, 1990.
3. A. M. Bruckstein, R. J. Holt, A. N. Netravali, and T. J. Richardson. Invariant signature for planar shape recognition under partial occlusion. *CVGIP: Image Understanding*, pages 49–65, 1993.
4. A. M. Bruckstein, N. Katzir, M. Lindenbaum, and M. Porat. Similarity invariant recognition of partially occluded planar curves and shapes. *IJCV*, pages 271–285, 1992.
5. J. B. Burns, R. S. Weiss, and E. M. Riseman. View variation of point set and line segment features. *T-PAMI*, pages 51–68, 1993.
6. S. Duvdevani-Bar and S. Edelman. On similarity to prototypes in 3d object representation. Technical Report CS-TR 95-11, Weizmann Institute, July 1995.
7. W. T. Freeman. Exploiting the generic view assumption to estimate scene parameters. In *proc. of ICCV*, pages 347–356, 1993.
8. D. P. Huttenlocher and S. Ullman. Object recognition using alignment. In *proc. of ICCV*, pages 102–111, 1987.
9. N. Katzir, M. Lindenbaum, and M. Porat. Planar curve segmentation for recognition of partially occluded shapes. In *proc. of ICPR*, 1990.
10. D. Keren, D. Cooper, and J. Subrahmonia. Describing complicated objects by implicit polynomials. *T-PAMI*, pages 38–53, 1994.
11. Y. Lamdan, J. T. Schwartz, and H. J. Wolfson. Affine invariant model based object recognition. *T-RA*, pages 578–589, 1990.
12. P. J. van Otterloo. *A Contour-Oriented Approach to Shape Analysis*. Prentice Hall Intrernational, 1991.
13. D. Weinshall and M. Werman. Disambiguation techniques for recognition in large databases and for under-constrained reconstruction. In *proc. of Intr. Symp. on Comp. Vision*, pages 425–430, Coral Gables, USA, 1995.
14. D. Weinshall, M. Werman, and N. Tishby. Stability and likelihood of views of three dimensional objects. *Computing Suppl.*, 11:237–256, 1996.

This article was processed using the LaTeX macro package with ECCV'96 style