Biol. Cybern. 62, 225–236 (1990)

Biological Cybernetics

The Largest Convex Patches: A Boundary-Based Method for Obtaining Object Parts

L. M. Vaina[1,2] and S. D. Zlateva[1]

[1] Intelligent Systems Laboratory, College of Engineering, Boston University, Boston, MA 02115, USA
[2] Harvard-M.I.T., Division of Health Sciences & Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract. The importance of boundaries for shape decomposition into component parts has been discussed from different points of view by Koenderink and van Doorn (1982), and by Hoffman and Richards (1984). The former define part boundaries as parabolic contours, whereas the latter propose that part boundaries should be defined by contours of negative minima (or maxima) of principal curvature. In this article, building on aspects of both approaches, we develop a new method for shape decomposition. This method relies exclusively on global properties of the surface which are fully characterized by local surface properties. We propose that a useful parcellation of shapes into parts can be obtained by decomposing the shape boundary into the *largest convex surface patches* and the *smallest nonconvex surface patches*. The essential computational steps of this method are the following: *(i)* build initial parts from the largest locally convex patches, *(ii)* consider an initial part as a constituent part if it is essentially convex, and *(iii)* obtain the remaining constituent parts by merging adjacent initial parts generated by the largest locally convex and the smallest nonconvex patches of nearly the same sizes. The method is illustrated on both smooth and continuous shapes. We show that the decomposition of shapes into the *largest convex patches* aims to maximize the "thingness" in an object, and to minimize its "non-thingness". The method is conducive to a natural parcellation of shapes into constituent parts useful for recognition and for inferring function.

1 Introduction

For the visual system, Bartlett's phrase "effort after meaning" (Bartlett 1932) implies that it must recognize and identify objects in the visual field, for it is objects that have meaning for us. The world around us is filled with an overwhelming variety of objects. One of the most fundamental human abilities is the perception of similarities and differences of objects and to thereby categorize them. Psychologists from the time of the Gestalt school onwards (Gibson 1979) have argued convincingly that the perception of objects is intimately related to the exploratory, manipulative and orientational motor behavior of the perceiver. That people recognize what they perceive as known objects, and that they can easily figure out possible functions of objects that they have never seen before and, moreover, that these perceptions tend to be consistent across perceivers (Pentland 1987), indicate that there is some natural, stable method for structuring object description. Furthermore, people tend to recover this natural structure from images and use it for recognition. We suggest along with others (Marr and Nishihara 1978; Vaina 1983; Pentland 1986; Biederman 1985) that at the heart of this commonly used method for structuring objects is the notion of "parts". These parts have both a perceptual and a functional prevalence for the perceiver, for they underlie both the distinctiveness of objects and their similarity. Our view is that parts are important because they offer a canonical and stable description of objects which is at once perceptual and functional.

Parts embody the visual meaning of objects and constitute the dominant component of the basic level of reference in the object category (Rosch 1973). Thus, at the basic level of a category the attributes of a chair include seat, legs and back. The basic level categories are the most general categories having members with similar shapes. We submit that object parts constitute the (global) reference space for features useful for recognition, or for determining the compatibility with the requirements imposed by actions (Vaina and Jaulent 1988). For example, a handle, which is the graspable part of a wide class of hand manipulable objects such as tools or kitchen utensils is usually long and thin and its size is compatible with the aperture of the human hand.

The structure of objects into parts and associated perceptual attributes serves as an economical reasoning about the world in terms of simplified category descriptions (Rosch 1973; Vaina and Jaulent 1988).

If we consider that parts are the building blocks of visual knowledge systems, then reasons of efficiency and usefulness impose that relatively simple combinations of parts must form "rough-and-ready models of the objects in our world and of how they behave" (Pentland 1986). We consider that the *practical* goal of perception, which refers to reasoning and to various types of recognition, is to find parts of objects as soon and as reliably as possible in the visual processing of images. The basic premise of this study is that organization by parts is both convenient and sufficient for the type of mental manipulations of objects which are involved in knowledge-based vision.

The idea of breaking objects into their components has been implicit in many systems, but has only recently been addressed explicitly. In recent years there has been considerable effort to formalize the decomposition of objects into parts in terms of differential geometry (Koenderink and van Doorn 1982; Hoffman and Richards 1984). Along this line, we propose a new method for obtaining parts. It differs from the previous methods in that it relies exclusively on global properties of the surface, which are fully characterized by local surface properties.

The problem of characterizing objects by parts, or in short the Parts Problem, consists of several subproblems which must be understood independently. It has been argued that the description of parts should be separated from the problem of finding parts (Hoffman and Richards 1984). In other words, the problem of assigning parts to categories is different from the problem of carving an object into parts, or from deriving parts from the image, or of relating the parts for describing whole objects for recognition, inspection, or manipulation.

We propose a new method of carving objects into parts and hence we shall limit the discussion to this aspect of the "parts problem". The paper is organized as follows: Section 2.1 briefly reviews two fundamental paradigms for decomposition of objects into parts. Section 2.2 presents in detail one of these paradigms, the boundary-based method, which constitutes the point of departure for the decomposition method proposed in this study and discussed in Sect. 3. Section 3.1 illustrates the method with an example. Section 3.2 gives the necessary formal background for developing this method and Sect. 3.3 gives examples of partitions according to the method of the largest convex patches we propose. Section 4 concludes by illustrating the psychological validity of the decomposition method proposed. We shall call our method the *decomposition in the largest convex patches* to distinguish it from Hoffman and Richards' *minima rule* and from Koenderinck and van Doorn's *parabolic lines*.

2 Methods for Obtaining Parts

2.1 Paradigms for Part Decomposition

Three main methods have been proposed for obtaining parts. The first method is usually referred to as *axis-based* (Blum 1968; Binford 1971) and relies on the decomposition of objects into parts based on the axes of symmetry. The second method is called *primitive-based* and relies on defining the possible shapes of the parts. The basic shapes of parts most commonly used are cylinders, spheres, cones or polyhedras. The third method is *boundary-based*, and here parts are defined through their boundary with adjacent parts (Koenderink and van Doorn 1982; Hoffman and Richards 1984; Bennett and Hoffman 1987). Both, the primitive-based and the boundary-based methods define parts on a representation of the object which is independent of the vantage point. Furthermore, they fully specify the 3-D shape, that is the volume occupied by the object or, equivalently, the whole 3-D bounding surface.

Both these methods have advantages and disadvantages. The axis-based method is good for objects which have axes of symmetry which are easily accessible from the earlier processes.

The primitive-based method is very useful for a priori specified classes of objects. Once one decides on the appropriate primitive parts, the task is to find them in the object and to associate them with characteristic metrical properties (e.g. length) and relate them by predicates which express spatial relations (e.g. to the right of). Many of our recognition tasks in the daily life are facilitated by context or by our expectations of what kinds of objects we may be dealing with. Specifically, in the domain of computer vision, for which most of the part-based methods have been developed, this holds quite well.

Naturally, one would prefer a general purpose shape recognizer which does not impose strong demands on the earlier processes. In essence it would be desirable to have a system which could recognize pineapples, tomatoes, cats and chairs equally well, even though in some circumstances some of these objects might be better described by some "special purpose" representations.

Bennett and Hoffman (1987) have demonstrated that the boundary-based methods do give a part definition which is completely general. In addition, we argue that the boundary based methods are better suited for discovering how to use an object in other than its primary function, that is the function that the object is designed for (Vaina and Jaulent 1988; Zlateva and Vaina, in preparation). This is because we visually sample only the parcellation of an object (Koenderink 1987), rather than the whole object. For the perceiver an object is first defined through exploratory movements as the visual response to its parcellation. Hence we argue that one first obtains a *functional* definition of the object, rather than a *description* of it.

Imagine, for example, holding an object in your hand and examining it to see what it is good for. You will first observe its humps, dimples, and furrows and then you will apply your reasoning to them. In this paper we propose a new boundary based method of parcellation. It is useful for relating the perception of an object to potential actions and this is the goal of the *functional recognition.*

2.2 *Boundary Based Methods*

Most representative of the boundary based method are the approaches of Koenderink and van Doorn (1982), and Hoffman and Richards (1984). The basis for obtaining parts in both methods is not dependent on the vantage point.

Both these methods, as well as the decomposition theory proposed in this article, employ techniques from differential geometry. Therefore, we shall briefly introduce the necessary terminology. The intersection of a surface S with a plane containing the surface normal at some point $p \in S$ is called *normal section.* This is a plane curve and its curvature is called the *normal curvature* of the surface in the direction given by the tangent of the normal section at p. The maximum normal curvature k_1 and the minimum normal curvature k_2 are called *principal curvatures.* The corresponding tangents give the *principal directions* of the surface which can be shown to be orthogonal. The Gaussian curvature at p is the product of the principal curvatures, k_1 and k_2, and it constitutes the basis for a classification of the points on the surface independent of its orientation. A point p is planar if $k_1 = k_2 = 0$, parabolic if $k_1 \cdot k_2 = 0$ and either k_1 or k_2 is $\neq 0$, elliptic, if $k_1 \cdot k_2$ is positive and hyperbolic, if $k_1 \cdot k_2$ is negative.

Koenderink and van Doorn (1982) were the first to develop a boundary based method for the decomposition of smooth solid shapes which are connected, bounded volumes with at most a finite number of sharp edges or points. In this method the parts are outlined by the parabolic lines on the surface of objects. These lines divide the surfaces of smooth shapes into elliptic (synclastic) and hyperbolic (anticlastic) patches. In the neighborhood of any point of the elliptic patches the object boundary lies completely on one side of the tangent plane. However, in the case of a hyperbolic, or saddle like patches, the surface cuts the tangent plane and cuts an hourglass region shape of the object. For elliptic patches, if the plane lies on the outside the patch is *convex* and if it lies inside, it is *concave.*

We have seen above that, in geometric terms, the parabolic lines connect points on the surface where the Gaussian curvature is zero. This method gives a part definition which is general and provides a partition of any smooth surface. In this sense we could informally consider that this schema is inherently global. However, Hoffman and Richards (1984) pointed out two limitations of Koenderink and van Doorn's method mainly regarding the unnaturalness of the decomposition it obtains. First, the condition of zero of the Gaussian curvature is not sufficient for all developable surfaces have zero Gaussian curvature but the method does not obtain parts in such cases. Second, this approach does not work in the case of figure/ground reversal since the parabolic lines are independent from the surface orientation.

The second boundary-based decomposition method we shall discuss here is based on the observation that maxima, minima and inflections of principal curvature along lines of curvature form good candidate points for partitioning a surface into units in a viewer independent manner. Hoffman and Richards (1984) observed that in the natural or constituent parts of shapes when 3-D parts are joined to create complex objects, by and large the contour of the joint is concave. This suggested to them that only minima of the normal curvature, not the maxima or inflections, be used to partition a surface into parts. Thus they proposed that, as a general rule, part boundaries are found in places of concave discontinuities of the tangent plane or negative minima of curvature. Hoffman and Richards' method relies in two rules for dividing a surface into parts:

(a) Divide a surface into parts at concave discontinuity of the tangent plane;

(b) Divide a surface into parts at loci of negative minima of each principal curvature along its associated family of lines of curvature *(the minima rule).*

A consequence of the above formulation is that the part boundaries which are not invariant under an orientation reversal for contours of the minima become contours of the maxima and vice versa. This is a psychologically valid point in that there is vast evidence that the human visual system is very sensitive to figure-ground reversal and so, for example, it describes the same curve differently depending on the direction of the traversal of the curve.

However, there are examples for which this method will not give good parts (Fig. 2). One problem is that although Hoffman and Richards' schema may obtain the required information for recognition, it is often not explicit and relations between parts are hard to obtain. Also, the application of the minima rule is not always straightforward while minima of the principal curvature may not necessarily exist on the surface or along the associated lines of curvature it is still possible to perceive distinct structural parts. Consider the shape bounded by the surface of revolution shown in Fig. 4a whose generating curve is composed of three smoothly joining circular arcs. The bounding surface is given by

$$r=(u\cos\varphi, u\sin\varphi, f(u)), \tag{1}$$

where u is the radius of the parallels, φ the angle between the (x, z) plane and the rotation plane and $f(u)$ the parameterization of the generating curve. The principal curvatures are correspondingly:

$$k_1=\frac{\dfrac{d^2f}{du^2}}{\left(\left(1+\left(\dfrac{df}{du}\right)^2\right)^{3/2}\right)}, \tag{2a}$$

$$k_2=\frac{\dfrac{df}{du}}{u\sqrt{1+\left(\dfrac{df}{du}\right)^2}}=\frac{\cos\Theta}{u}, \tag{2b}$$

where, θ is the angle between the axis of rotation and the tangent to the generating curve (Fig. 1a). The surface is smooth with two contours of discontinuity of k_1.

Taking into consideration that k_1 is just the curvature of the generating curve and k_2, the reciprocal of the length of the normal intercepted between $f(u)$ and the z-axis. It is easy to see that there are no negative minima of k_1 and k_2 is positive everywhere. Even if one smooths the surface the minima would be rather weak and difficult to localize.

However, segmentation contours like those shown in Hoffman and Richards (Fig. 1b) could be obtained when the rule is extended as to define part boundaries at loci of positive maxima if the two principal curvatures take opposite signs. For example in Fig. 4a, this would yield the segmentation contour occurring at $\theta=0$ corresponding to places informally characterized as being if minimal width.

The shape in Fig. 4a is more likely to be perceived as composed of two spheres connected by a smooth neck. The decomposition according to the extended minima rule, results in two spheres and half of the joining neck which is rather unnatural.

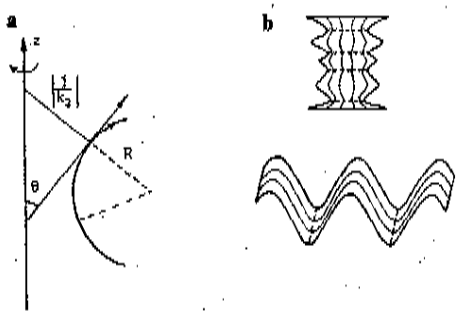


Fig. 1. a Geometrical characterization of a surface of revolution whose generating curve $f(u)$ is a circular arc of radius R. The principal curvatures k_1, k_2 of any surface of revolution are correspondingly the curvature of the generating curve and the reciprocal of the length of the normal intercepted between $f(u)$ and the axis of rotation z. In the above example k_1 is constant everywhere ($k_1=R$). At places characterized as being of minimal width k_2 has a positive maximum [$=0$, with Θ the angle between z and the tangent of $f(u)$]. **b** Segmentation contours according to the minima rule (from Hoffman and Richards 1984, Fig. 8)

2.3 *Criteria for the Evaluation of the Decomposition Methods*

The previous discussion and the general goal of this study to obtain parts which are functional, suggest general criteria which must be satisfied by an adequate method of part definition:

The first criterion is *robustness*. This criterion says that the method must be insensitive to small changes of shape due to differences among the members of a class or to small differences in the viewing position. Robustness is achieved by requiring that the part definition should rely on global features of the spatial structure.

The second criterion is *computability* and requires that ultimately parts must be extracted from the image. To satisfy computability, it is assumed that selected global features build on local properties extracted from representations obtained earlier in the image processing (Marr and Nishihara 1978; Brady 1983; Brady and Asada 1984; Biederman 1985).

The third criterion refers to the *scale*, or the resolution of the description. Parts can be defined at several scales and the parts defined at coarser scales tend to play a more dominant role in the description of the object, comparable to object sizes. However, for determining the implementation of an action with the object, parts expressed at finer scales are often more relevant. The multiple scale structure involves a description which is based on the use of parts, and the smaller parts are expressed in terms of coarser ones. This induces a hierarchical indexing of the object that makes efficient both the search and the representation of useful attributes (Vaina and Jaulent 1989). For sawing, for example, the wiggly edge of the blade implements the physical constraints required by the action-consequence. The blade, however, tells us that it is a "cutting" object.

The fourth criterion requests that the representation have *rich local support* (Brady 1983). By rich, it is meant that the representation should be information preserving so that a close approximation to the original shape is recoverable from the representation. *Local support* means that the information should be derivable from local parts of the shape.

The fifth criterion refers to *smooth extension and subsumption* (Brady 1983). Smooth extension requires that local information should give rise to global representation through processes of grouping relating local representations into a single global representation. The subsumption criterion suggests that one local representation is better suited than others in order to yield a more concise and natural representation of the shape.

In this paper we shall propose a method for part identification from three dimensional surfaces which satisfies these criteria. We shall prove formally and illustrate with several examples that our approach circumvents the difficulties met by the decomposition methods discussed above.

3 A Convexity Based Method of Surface Decomposition into the Largest Convex Patches

Both the mathematical concept of smooth shapes and plastic art emphasize the importance of the elliptic patches which have a *thing-like* characteristic. By contrast, the hyperbolic patches are only surfaces of transition have a *nonthing-like* characteristic (Koenderink and van Doorn 1982), they are the "glue" that holds the parts together. Viewing plastic art as the expression of knowledge which, by and large, is gained through visual perception, Koenderink and van Doorn suggested that "natural" models for shape description will have to take into account these observations, and use as descriptors mathematical formulation of elliptic patches. The decomposition into elliptic and hyperbolic patches works well for smooth shapes. Many of the objects surrounding us have continuous shapes in the sense that derivatives of higher order do not exist (unless one does smoothing of the shape). A consistent extension of the local behaviour of the elliptic patches is, however, captured by the property of local convexity of the surfaces.

Taking into consideration the tendency of the visual system to close curves and surfaces in the most conservative way [1], the "inward" surface of a part may be defined as the surface of minimum area *(minimal surface)* that closes the locally convex patch. Following these observations, and considering the importance of the space enclosing properties of surface patches for part perception (Gibson 1979), we propose the hypothesis that a natural decomposition of a shape must *maximize* its thing-like properties and *minimize* its non-thing like characteristics.

We suggest a decomposition of the shape boundary into the *largest convex surface patches* and the *nonconvex surface patches* which, as we shall prove, with additional assumptions gives a unique surface partition. More specifically, we first decompose an object into sets of points "enclosed" [2] by the *largest* convex and *smallest* nonconvex surface patches of comparable sizes. We argue that such a decomposition maximizes the "thingness" in an object, and minimizes its "nonthingness".

The surface decomposition into the largest convex and the smallest nonconvex patches will then generate three types of shape regions, which we will refer to as initial parts: *(i)* convex parts if the set of points enclosed by the convex surface patch builds a convex body; *(ii)* nonconvex parts generated by the largest convex patches; *(iii)* nonconvex parts generated by the smallest nonconvex patches.

3.1 An Example

We will discuss a simple example which illustrates the different initial parts and how they may relate to the constituent parts of the object. A *suitcase*, such as that depicted in Fig. 2, has two component parts: the container and the handle. The container surface has two largest locally convex parts (P^1, P^2) and three smallest nonconvex surface patches (P^3, S^1, S^2). As discussed above the significance of the two nonconvex patches S^1 and S^2 is limited to "glue" or to relate the constituent parts.

We see that by closing P^1 through the minimal surface, we obtain exactly a constituent part. The same procedure applied on the P^2 results in an awkward segmentation of the handle. In order to derive the handle shape the two parts obtained from P^2 and P^3, by closing them through the minimal surface, must be merged.

Two observations emerge from this example. First, the largest nonconvex patches are conducive to obtaining parts. Second, two distinct strategies can be employed for obtaining parts, and which one should be used is determined by the specific structural characteristics of the surface. Thus, as we have seen in the

[1] This property has been used in the work on surface representation (Horn 1986; Brady and Asada 1984; Brady et al. 1985). However we use it here not to obtain surface representations consistent with the available sparse data, but to trace part boundaries

[2] By "enclosed" it is meant that the shape's points belong to the subset of R^3 bound by the corresponding surface patch and the minimal surface which closes it

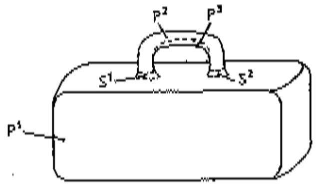


Fig. 2. The surface of the suitcase shape displays two connected subsets P^1 and P^2 which are locally convex everywhere. Closing P^1 in a "conservative" way, e.g. through a surface of minimal area, one obtains a convex set of 3D points which defines a constituent part the container. The other constituent part – the handle – builds a nonconvex set and is defined by the locally convex P^2 as well as the nonconvex P^3 of nearly the same size as P^2. The remaining two nonconvex sets S^1 and S^2 are of small size and are overridden by the constituent parts

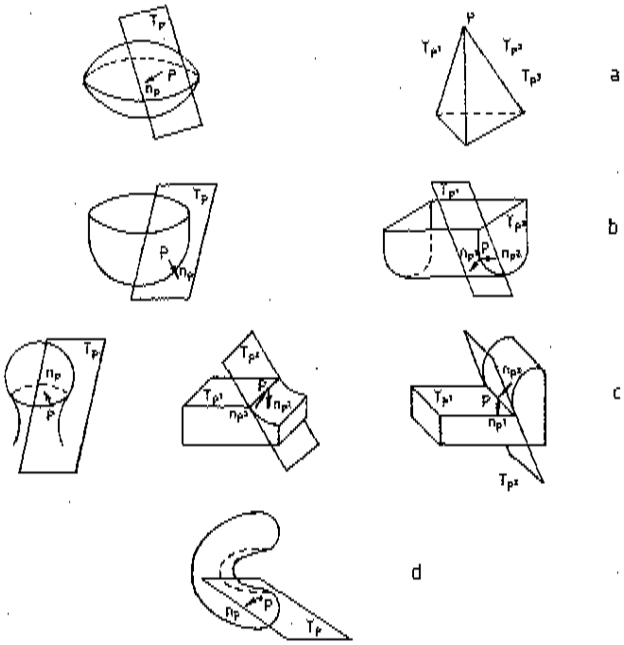


Fig. 3a–d. Local and global convexity of surface patches. Under the assumption that the surface normal points towards the shape the boundary S at p is: **a** locally positive convex; **b** locally negative convex or concave; **c** locally nonconvex. The left column illustrates the above notions for smooth and the right column for continuous shapes. The surface in **d** shows that local convexity does not entail global convexity: the surface patch P^1 is locally convex in all its points and nevertheless there is a tangent plane which cuts P^1 (P^2 is locally nonconvex and thus globally nonconvex)

suitcase example, the part generated by closing P^1 is locally and globally convex. That is, the whole part always lies on one side of any of its tangent planes, whereas the part generated by P^2 is (globally) nonconvex. The locally nonconvex patch P^3 has size parameters of the same order as the locally convex surface patches, whereas the nonconvex patches S^1 and S^2 are significantly smaller in size. Methods based on the *parabolic lines* would divide the handle into several regions, and likewise the *minima rule* would decompose the handle into several parts.

In the following section we shall give the formal characterization of the class of shapes under discussion and we will present a new decomposition method for smooth shapes.

3.2 The Largest Convex Patch Method of Shape Decomposition

The class of shapes discussed in this paper is limited to smooth objects with well defined physical boundaries. Therefore it includes man-made objects and animals, but not shapes of rivers, clouds or mountains.

Such objects are bounded volumes formally defined as connected compact (closed and bounded) subsets of the three dimensional Euclidian space, R^3. For the present investigation we assume that the boundary is a single connected surface. For smooth objects (for which the first derivative exists everywhere), the boundary has a tangent plane defined everywhere. This differs from the case of continuous shape for which first order derivatives do not always exist everywhere and hence the tangent plane is not always defined. We note that methods from differential geometry cannot always be directly applied for the shapes defined above as they are developed for surfaces with partial derivatives of all orders.

The local convexity property that our decomposition method builds upon can be defined uniformly for smooth and continuous shapes as follows: a surface S is *locally convex* at a point p, $p \in S$ if there exists a neighborhood $V \subset S$ of p, such that V is entirely contained in one of the closed half spaces determined by the tangent plane(s) T_{pi} at p $(i=1, \ldots, i_p)$, where i_p is the number of sided tangent planes at p (usually p is equal to 2). Correspondingly, if there is a tangent plane T_{pk} with $k \in \{1, \ldots, i_p\}$ such that V_p lies in both half-spaces determined by T_{pk} the surface is nonconvex at p.

For a fixed direction of the surface normal n_p at p, (e.g. pointing towards the shape) we can differentiate two cases of local convexity. A convexity is locally *positive* if and only if for all tangent plane(s) T_{pi} both the normal n_{pi} to T_{pi} and some neighborhood V_p are contained in one of the half-spaces determined by T_{pi}. Otherwise the convexity is locally *negative* convex or concave (Fig. 3a–c).

A connected subset of the surface is referred to as a surface patch and is said to be locally convex (positive or negative) if the surface is locally convex in all points of the patch. It is clear that global convexity implies local convexity but the converse is not necessarily correct, as illustrated in the example of Fig. 3d where the locally convex patch P^1 is locally but not globally convex.

Global convexity is a strong requirement which is computationally sensitive to very small perturbations of the object surface. In order to make the method more robust we propose to compare the convexity of a set against a model (e.g. the convex hull of the set [3]). One possibility is to threshold the difference between the volume of the considered subset of R^3 and the volume of its convex hull. This results in an essentially convex subset (Philipps et al. 1986). A subset K of R^3 is said to be essentially convex if

$$Volume_k - Volume_{convex\,hull(K)} < Threshold.$$

The input to the decomposition method is a canonical representation of the three-dimensional surface of the whole object. Also, as with the other boundary-based decomposition methods, the surface representation is viewer independent and hence the parts obtained will be defined independent of the viewpoint.

We are ready now to present our method for surface decomposition according to the largest convex patches and the smallest nonconvex patches.

The decomposition is performed in four stages:

3.2.1 Initial Surface Decomposition $[S]^0$ *into Smooth Surface Patches* S^i. Build the set $[S]^0$ of the connected subsets S^i with area $(S^i) \neq 0$ and with the first and second derivatives defined everywhere. $[S]^0$ is referred to as the *initial surface decomposition*. A label $l(S^i)$ of the ordered pair of signs of principal curvatures is uniquely assigned to each S^i, with

$$l(S^i) = (sign(k_1)\, sign(k_2))^i \quad \text{and} \quad l(S^i) \in \{-1, 0, 1\}^2$$

with $(t_1, t_2, n)_p$ building a positively oriented local coordinate frame at $p \in S^i$. The t_1, t_2 are the tangents in the principal directions and n is the surface normal pointing towards the shape.

The initial decomposition $[S]^0$ contains the following classes of surface patches:

- positive (or negative) elliptic patches for which $l(S^i) = (11)$ [respectively $l(S^i) = (-1-1)$]
- positive (or negative) parabolic patches for which $l(S^i) \in \{(01), (10)\}$ [respectively $l(S^i) \in \{(0-1), (-10)\}$]
- planar patches for which $l(S^i) = (00)$
- hyperbolic patches for which $l(S^i) \in \{(-11), (1-1)\}$.

The positive (negative) elliptic and positive (negative) parabolic patches are locally positive convex (negative locally convex or concave) everywhere and the hyperbolic patches are locally nonconvex everywhere.

[3] A set K ($K \in R^3$) is convex if for every pair of points $p_1 p_2$ belonging to K, the whole segment $p1p2$ belongs to K or equivalently, if K is the intersection of all closed half-spaces that contain it (e.g. Serrat 1982). The convex hull of a set is defined as the smallest convex set containing the set

The boundary e^{ij} between two patches $S^i, S^j \in [S]^0$ is a curve with the tangent uniquely defined everywhere (S was assumed to have a tangent plane uniquely defined everywhere). The surface S in e^{ij} has discontinuity in the curvature (the Gaussian curvature is not defined) or at least one principal curvature is equal to zero. In the following, we shall refer to e^{ij} as the *edges* between two patches. Locally an *edge* is uniquely assigned a label $l(e^{ij}) = (+, -, n)$ depending on whether the surface is locally positive convex, concave or nonconvex in e^{ij} (Faux and Pratt 1978).

For smooth shapes the signs of the principal curvatures of S^i and S^j constrain the possible labels of the edge e^{ij} between them. If there is a neighborhood V_p for $p \in S$ in which the principal curvatures do not take opposite signs, then the surface is convex in p and we have the following possible situations:

(a) $\forall l(S^i), l(S^j)\,\{(11), (10), (01), (00)\} \quad l(e^{ij}) = +$

(b) $\forall l(S^i), l(S^j)\,\{(-1-1), (-10), (0-1), (00)\}$
$l(e^{ij}) = -$.

The patches S^i of the initial surface decomposition $[S]^0$, have the following property:

$$\sum_{S^i \in [S]^0} \text{area}(S^i) = \text{area}(S) \qquad S^i \cap S^j \neq 0 \quad \text{for} \quad i \neq j.$$

In general $S^i = S$, and thus the decomposition $[S]^0$, is not a partition.

3.2.2 Surface Decomposition $[S]$ *into Largest Locally Convex and Smallest Nonconvex Patches.* Build a surface decomposition $[S]$ of the largest convex (positive and negative) and smallest nonconvex surface patches P^k and label them with

$$l(P^k) = (+, -, n)$$

where

$-l(P^k) = + \Leftrightarrow P^k = \cup(S^i \cup S^j)$

for $l(S^i), l(S^j) \in \{(11), (10), (01), (00)\}$ and $l(e^{ij}) = +$

$-l(P^k) = - \Leftrightarrow P^k = \cup(S^i \cup S^j)$

for $l(S^i), l(S^j) \in \{(-1-1), (-10), (0-1), (00)\}$ and $l(e^{ij}) = -$

$-l(P^k) = n \Leftrightarrow P^k = S^i \quad \text{for} \quad l(S^i) \in \{(1-1), (-11)\}$.

The surface decomposition $[S]$ is uniquely defined if there is no planar patch S^i which joins a positive convex patch S^m in a positive convex edge and, concomitantly, it joints a concave patch S^n in a concave edge, i.e.

$$\forall S^i, S^m, S^n, l(S^i) = (00)$$

$$\neg\,(l(e^{mi}) = + \quad \text{and} \quad l(S^m)\,\{(11), (01), (10)\} \text{ and}$$
$$l(e^{ni}) = - \quad \text{and} \quad l(S^n)\,\{(-1-1), (0-1), (-10)\}).$$

The surface decomposition $[S]$ is trivially uniquely defined for surfaces in which there is no occurrence of patches S^i with $l(S^i)=(00)$ and area $(S^i)\neq 0$.

Note that, even in the last case, the shapes under consideration build a larger class than the classical surfaces with partial derivatives of all orders. Similarly to $[S]^0$ the surface decomposition $[S]$ has the property:

$$\sum_{P^k\in[S]} \text{area}(P^k)=\text{area}(S) \quad \text{and} \quad P^k\cap P^l=\emptyset$$

for $k\neq 1$.

3.2.3 Shape Decomposition into Initial Parts (IP^m). Build the closed subsets IP^m of R^3 with boundary, P^k, size parameters of the same order as the largest locally convex patches and a minimal surface that closes them. The IP^m are called initial parts and are assigned the same labels as the generating P^k. The set of all initial parts is denoted by $[IP]$.

3.2.4 Shape Decomposition into Constituent Parts (CP^n). (1) Consider all essentially convex initial parts IP^k as essentially convex constituent parts CP^n. In this case the label of the constituent part is the same as the label of the corresponding initial part. Thus, accordingly, a constituent part labelled with + (respectively with −) is a *positive (negative)* constituent part.

(2) Nonconvex constituent parts CP^n are obtained in two ways. Firstly, they result from the union of all pairs of adjacent nonconvex initial parts in which one initial part is generated by a largest locally convex patch, and the other initial part is generated by a smallest nonconvex initial patch (e.g. the *handle* in Fig. 5). Secondly, they are obtained from the remaining nonconvex initial parts which are of comparable size and are generated by a single smallest nonconvex patch (e.g. the graspable part of the *handle* in Fig. 4a and b). Nonconvex constituent parts are always positive and hence they are always labeled with +.

We wish to emphasize here that the nonconvex surface patches P^k of small sizes, as compared to the convex patches do not generate initial parts. The small negative Gaussian curvatures correspond to joints but large non-negative curvatures correspond to nonconvex constituent parts which should be merged with convex parts. Hence, from this decomposition into constituent parts $[CP^n]$, an approximation of the shape may be obtained by building the difference between the union of the positive parts and the union of the negative parts.

In summary, the basic idea of the decomposition proposed here involves the following steps: *(i)* build initial parts from the largest locally convex patches; *(ii)* consider an initial part as a constituent part if it is essentially convex and *(iii)* obtain the remaining

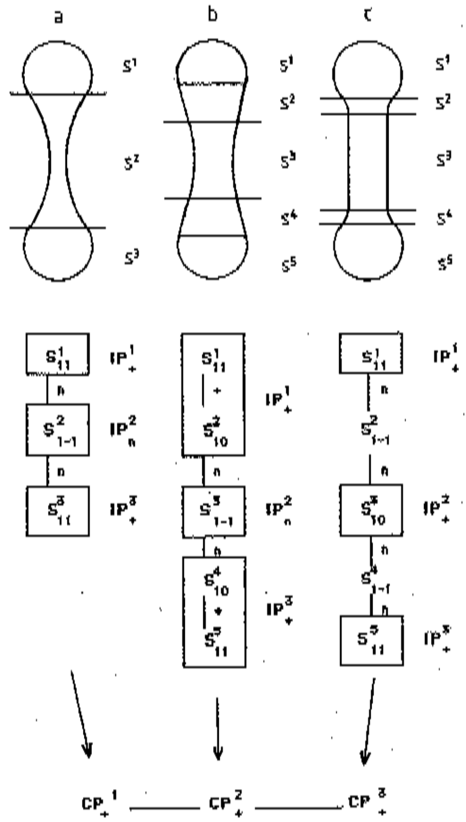


Fig. 4a–c. The dumbbell shapes shown in **a**, **b**, and **c** have the same constituent parts – two weights and a joining neck for grasping – but different bounding surfaces. Thus the initial decompositions are different for each shape. This is depicted as a graph with nodes, the surface patches S^i of nonzero area, and second derivatives defined everywhere. The ordered pair of signs of principal curvatures of S^i is assigned as its lower index. Two patches S^i and S^j are joined by an edge in the graph if there is an edge e^{ij} on the surface. The edges of the graph are labeled with the corresponding $l(e^{ij})$. The boxes indicate only the largest locally convex and smallest locally nonconvex patches P^k which generate initial parts IP. In this example all initial parts are positive and are generated by a single P^k. Again the label of the constituent part is assigned as lower index to the CP

constituent parts by merging adjacent initial parts generated by the largest locally convex and the smallest nonconvex patches of nearly the same sizes.

We shall illustrate this method by several examples of decomposition of smooth objects.

3.3 Examples of Partitions According to the Method of the Largest Convex Patches (LCP)

In Fig. 4 we see three differently shaped dumbbells and a corresponding graph representation of the initial decomposition $[S]$. The elliptic, parabolic, and hyperbolic patches S^i are mapped onto the nodes, and the

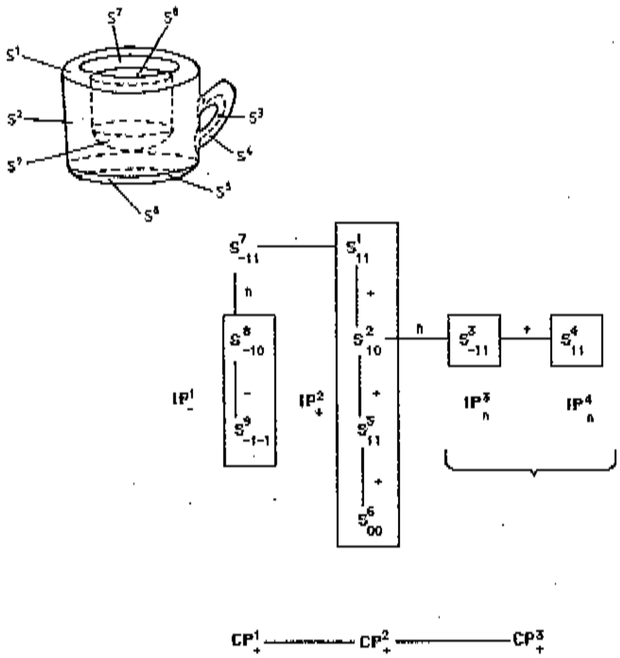


Fig. 5. Surface decomposition of a cup with the corresponding graph representation. The blocks outline the largest convex patches which generate initial parts. The convex initial parts IP^1 and IP^2 define a negative and a positive constituent part correspondingly. The two adjacent nonconvex parts IP^3 and IP^4, the one generated by a largest convex and the second by a smallest nonconvex patch of comparable size, are then merged to yield the handle

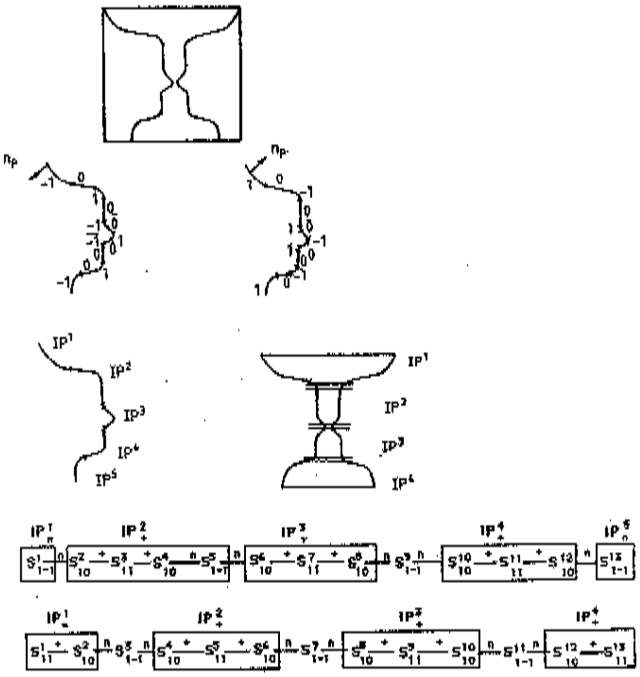


Fig. 6. Figure-ground reversal leads to change in part boundaries: The local properties of the surface along the bounding contour can be inferred from the projection of the bounding contour taking into account that for points along the contour *(i)* the sign of the Gaussian curvature is the same as the sign of the curvature of the projection (Koenderink 1984; Brady et al. 1985) and *(ii)* one of the principal curvatures is always positive. The smooth patches S^i of the initial decomposition $[S]^0$ can then be easily inferred from the corresponding portions of the contour and grouped in largest locally convex and smallest locally nonconvex patches according step 2 of the method and corresponding part boundaries found. With reversal of the direction of the surface normal the number of elements of $[S]^0$ remains the same but their local convexity characteristic changes which entails a change in part boundaries. The graph representations of the initial decomposition with the blocks indicating the P^k leading to initial parts are shown for the face (a) and the goblet (b)

edges e^{ij} onto the graph edges. The nodes are denoted by S^i, with the corresponding label $l(S^i)$ as lower index, and the graph edges are labelled as $l(e^{ij})$. We see that although the initial decomposition $[S]^0$ is different for each shape, the method always yields the correct constituent parts. This satisfies the *robustness* criterion. Figure 5 illustrates the decomposition of a cup which displays a negative part.

Figure 6 demonstrates that, by and large, our method accounts for figure-ground reversal. This figure represents the 2D view of an object that can be perceived either as a face or as a goblet. Since the curvature of the contour has the same sign as the Gaussian curvature of the surface (Koenderink and van Doorn 1982; Brady and Asada 1984) and one of the principal curvatures is always positive for all the points of the bounding contour, it follows that the largest locally convex patches project in a contour of zero or positive curvature. With the reversal of surface orientation, negatively curved sections of the contour become positively curved and associate the adjacent sections of zero curvature. In this way, a change in part boundaries is entailed [this has been extensively discussed Brady et al. (1984) and by Koenderinck (1987)].

Hence the *largest convex patches* (LCP) method of shape decomposition gives essentially the same results as Hoffman and Richard's *minima rule* when the constituent parts are separated by nonconvex parts of small sizes or by a closed contour of concave discontinuity.

4 Psychological Validation

The above method of shape decomposition based on finding the largest convex surface patches and the smallest nonconvex patches yields two types of parts: essentially convex (positive and negative) and essentially nonconvex parts. We have shown (Sect. 3) that in either case the structural feature used for part discrimination is the largest locally convex surface patch

P^k. This generates an essentially convex constituent part if the set of space points enclosed by P^k is essentially convex and a nonconvex constituent part otherwise.

In this section we shall provide empirical evidence that the *largest convex parts* tend to correspond to the constituent parts of objects. Thus, if the hypothesis that visual recognition of objects may be carried out on representations of shapes by *constituent parts* is correct, it is natural to assume that recognition would be best achieved in those views in which most object parts can be identified. According to our method this occurs in views in which the outline of the parts can be obtained, and their convexity or nonconvexity attributes can be determined.

Although in everyday life objects are seen in many orientations, it is clear that some orientations are more familiar than others. Warrington and James (1986) have investigated the ability of normal subjects and patients with right hemisphere lesions to identify common objects viewed in three-dimensions at different angles of rotation. The motivation for their experiment was Warrington's earlier results (Warrington and Taylor 1973) which showed there is a "favoured" view for efficient object recognition, informally called the prototypical view. Furthermore, brain damaged patients with lesions affecting the posterior right hemisphere are less able to tolerate a deviation from the prototypical view. However, such patients had no difficulty in the recognition of object photographs in which the angle of view was considered prototypical. The results were interpreted as a deficit of perceptual categorization, which is a process whereby multiple stimuli (e.g. the same object, but different views) are judged to belong to the same class.

Only two views were considered in these experiments, and thus the question of what object view might constitute the recognition threshold was not addressed. The recognition threshold is measured by the angle of rotation for correct identification. Warrington and James (1986) addressed the threshold for the identification of 3 Dimensional objects by both normal subjects and patients with right hemisphere lesions. Ten common objects, each with a well defined base, were used in this experiment and they were presented in three-dimensions by a shadow image projector (Gregory 1964). The objects were placed, one at the time, on a turn table rotated by a stepping motor in 1 1/2° steps. The turn table was situated between the light sources and the screen and, by using crossed polaroid glasses, the subjects viewed the three-dimensional shadows of objects presented in various degrees of rotation. The thresholds for object recognition were measured under two conditions: *(a)* by rotating the object from its dorsal view through its vertical axis and *(b)* by rotating the object from its basal view through its horizontal axis (Fig. 7).

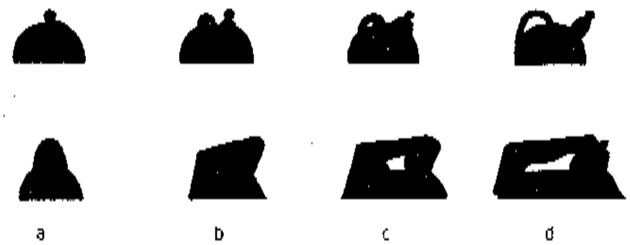
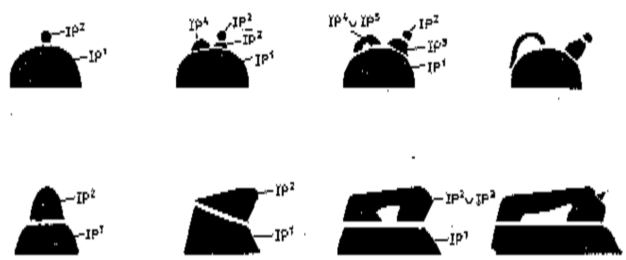


Fig. 7a–d. (From Warrington and James 1986) Projected images of two objects – iron and kettle: **a** initial image of the object in lateral rotation conditions; **b** image of objects correctly recognized by 50% of the control group; **c** image of the objects correctly recognized by 100% of the control group; **d** image of objects fully rotated through 90°

Starting from two initial positions (90° rotation about the vertical axis and 90° rotation about the horizontal axis with respect to the lateral view) the objects were rotated in single steps back to the lateral view (through a lateral and base rotation correspondingly) and the angle at which correct recognition was achieved was recorded as the "object recognition threshold". Independence of the measurements was guaranteed by recording once from each subject for each rotation condition.

The experimental results indicated that very different recognition thresholds measures were obtained with respect to task difficulty (as measured by the angle of rotation) and intersubject variability. The data showed that in both normal control subjects and patients with right hemisphere lesions, the manipulation of the angle of view did not have a systematic effect on the object recognition. The authors interpret these results as a challenge to an aspect of Marr's theory of object recognition, namely that the object representation depends on axes of symmetry or elongation (for objects which do have such axes) which are obtained from a viewer centered surface description (the 2 and 1/2D sketch in Marr's theory). Such axes give rise to a viewer independent coordinate system to describe the geometry of the spape. While agreeing with Marr's view that object recognition must rely on descriptions which are independent of the viewer, Warrington and James argue, however, that the object's principal and component axes may not form the basis for describing the geometry of the shape. They suggest instead that the perceivers may rely on *distinctive features*, and that recognition occurs when sufficient features have been processed to specify the object and to differentiate it from other similar objects. That is, when there is enough information for obtaining perceptual categorization.



a b c d

Fig. 8a–d. Decomposition of views of the shadowed three-dimensional images depicted in Fig. 7. The initial parts IP^k generated by the largest locally convex patches on different views of the two objects are shown. The columns correspond to initial position, views of 50% and 100% recognition and lateral view

A detailed debate on whether axes or features are most useful for recognition is outside the scope of this article. Our focus is to show first that the decomposition into the *largest convex patches* proposed here defines formally the *distinctive features* which usually constitute component parts of the object, and second, to discuss the circumstances in which recognition occurs.

The local surface convexity is viewpoint independent. Thus, applying our decomposition method on a particular view of the object surface will result in outlining the constituent parts in view. As might be expected, if parts are missing or nonconvex parts are perceived as convex the recognition becomes more difficult.

We will use the shapes of the *iron* and the *kettle* (Fig. 8) from Warrington and James' experiment to show that the *recognition threshold function* is strongly related to the identification of the largest locally convex patches and the convexity-nonconvexity attribute of the set of enclosed points which determine the parts derivation process.

In the initial position the surface of the kettle in Fig. 8a displays only two locally convex patches which, closed through the minimal surface, define convex initial parts IP^1 and IP^2. In Warrington and James's experiment, these yield only 10% recognition. In the second view the remaining parts generated from IP^3 and IP^4 are outlined but recognition rises only to 50%. This is because the alternative essentially convex-nonconvex for IP^4 is undecided. This is resolved in the third view which corresponds to 100% recognition.

Similarly, in the initial position the iron (Fig. 8b) shows two locally convex patches generating IP^1 and IP^2, with IP^2 mistaken as defining an essentially convex part. We see that this view achieves 0% recognition. In the following views as the recognition ability rises (50% and 100%), IP^2 is identified as nonconvex and thus it leads to a nonconvex part.

These considerations bring experimental support to the naturalness of the decomposition method proposed in this paper. It must be pointed out, however, that we do not claim that only the largest locally convex patches and the convexity-nonconvexity alternative of the set of enclosed points may be used to characterize the object parts and to achieve recognition. The rise from a 0% to 50% recognition rate in the first two views of the kettle is perhaps related more to the perception of elongation of the two parts obtained in the same view, rather than to the fact that in the second view IP^2 is not unequivocally determined as essentially convex.

This suggests that part discrimination and characterization relies on several global features which are computed in parallel and integrated to yield recognition.

5 Conclusion

A new approach to the decomposition of objects into parts has been proposed. It is suggested that parts are obtained by decomposing the object into sets of shape points enclosed by the largest convex and the smallest nonconvex patches. The method, called *the decomposition into the largest convex patches* (LCP) is based on the global features of the surface which, in turn, are built upon surface features extracted by local operators. The basic idea of this decomposition involves the following steps: *(1)* obtain the initial parts from the largest locally convex patches, *(2)* consider essentially convex initial parts as constituent parts, *(3)* obtain the remaining constituent parts by merging adjacent initial parts. These initial parts are obtained from the largest convex and the smallest nonconvex patches of nearly the same sizes.

We show in a series of examples that the LCP decomposition method maximizes the "thingness" characteristics of the object and minimizes its "non-thing-like" characteristics. Thus, it gives structural parts which also are *constituent parts*. We are planning further experiments along the line of those of Warrington and James to evaluate the psychological validity of our theory and to contrast it with other existing theories for extracting object parts. Our view is that several representations based on different salient characteristics of the object (e.g. axes, surface patches) are maintained in parallel in the human visual system. The one which is actually used depends on the specific object one looks at. A comparison of these various representations and their effectiveness on the recognition of common objects will be presented in a forthcoming paper (Zlateva and Vaina, in preparation).

References

Bajcy R, Solina F (1987) Three dimensional object representation revisted. Proc First Int Conf Comput Vision: 231–241

Bartlett FC (1932) Remembering: a study in experimental and social psychology. Cambridge University Press, London New York

Bennett B, Hoffman D (1987) Shape decompositions for visual recognition: the role of transversality. In: Richards W, Ullman S (eds) Image understanding. MIT Press, Cambridge, pp 215–256

Biederman I (1985) Human image understanding recent research and a theory. Comp Vis Graph Imag Process 32:29–73

Binford TO (1971) Visual perception by computer. IEEE Conference on Systems and Control, Miami

Blum H (1973) Biological shape and visual science. J Theor Biol 38:205–287

Brady MJ (1983) Criteria for representations of shapes. In: Rosenfeld A, Hope B, Beck J (eds.) Human and machine vision. Academic Press, New York

Brady JM, Asada H (1984) Smoothed local symmetries and their implications. Int Robot Res 3

Brady JM, Ponce J, Yuille AL, Asada H (1985) Describing surfaces. Comp Vis Graph Imag Process

Faux ID, Pratt MJ (1978) Computational geometry for design and manufacture. Ellis Horwood Press

Gibson J (1979) The ecological approach to visual perception. Houghton Miflin, Boston

Gregory RL (1964) Stereoscopic shadow images. Nature (London) 193:1407–1408

Hilbert D, Cohn-Vossen S (1932) Anschauliche Geometrie. Springer, Berlin Heidelberg New York

Hoffman D, Richards W (1984) Parts of recognition. Cognition, pp 1865–1896

Horn BKPH (1986) Robot vision. McGraw-Hill, New York

Koenderink J (1987) An internal representation for solid shape based on the topological properties of the apparent contour. In: Richards W, Ullman S (eds) Image understanding, pp 257–285

Koenderink J, van Doorn A (1982) The shape of smooth objects and the way contours end. Perception 11:129–137

Marr D, Nishihara, HK (1978) Representation and recognition of the spatial organisation of three dimensional shapes. Proc R Soc London B 200:269–294

Pentland A (1986) Perceptual organization and the representation of natural form artificial intelligence 283:1–16

Pentland A (1987) Recognition by parts. First International Conference on Computer Vision, pp 612–621

Phillips T, Cannon R, Rosenfeld A (1986) Decomposition and approximation of three-dimesnional solids. Comput Vision, Graph, Image Process 33:307–317

Rosch E (1973) On the internal structure of perceptual and semantic categories. Cognitive development and the acquisition of language. Academic Press, New York

Rosenfeld A (1988) An overview of computer vision (keynote address). SPIE Apl Artif Intell 937:382–384

Serrat J (1982) Image analysis and mathematical morphology. Academic Press, New York

Vaina LM (1983) From shapes and movements to objects and actions. Synthese 54:3–37

Vaina LM, Jaulent MCh (1989) Object structure and action requirements: a compatibility model for functional recognition. Int J Intell Syst 7:73–92

Warrington EK, James M (1986) Visual object recognition in patients with right hemisphere lesions: axes or features? Perception 15:355–366

Warrington EK, Taylor AM (1973) The contribution of the right parietal lobe to object recognition. Cortex 9:152–164

Weatherburn CE (1927) Differential geometry of three dimensions. Cambridge University Press, Cambridge

Received: October 20, 1988
Accepted in revised form: August 15, 1989

Dr. Lucia M. Vaina
Harvard-MIT Division of Health Sciences and Technology
MIT 20A-121
Cambridge, MA 02139
USA