

Recovering Shape and Determining Attitude
From Extended Gaussian Images
by
James Joseph Little

Technical Report TN 85-2
April 1985

Department of Computer Science
The University of British Columbia
Vancouver, B.C. V6T 1W5

This report was submitted as a thesis in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

Abstract

This dissertation is concerned with surface representations which record surface properties as a function of surface orientation. The Extended Gaussian Image (EGI) of an object records the variation of surface area with surface orientation. When the object is polyhedral, the EGI takes the form of a set of vectors, one for each face, parallel to the outer surface normal of the face. The length of a vector is the area of the corresponding face.

The EGI uniquely represents convex objects and is easily derived from conventional models of an object. An iterative algorithm is described which converts an EGI into an object model in terms of coordinates of vertices, edges, and faces. The algorithm converges to a solution by constrained optimization. There are two aspects to describing shape for polyhedral objects: first, the way in which faces intersect each other, termed the adjacency structure, and, second, the location of the faces in space. The latter may change without altering the former, but not vice versa. The algorithm for shape recovery determines both elements of shape. The continuous support function is described in terms of the area function for curves, permitting a qualitative comparison of the smoothness of the two functions. The next Section describes a method of curve segmentation based on extrema of the support function. Because the support function varies with translation, its behaviour under translation is studied, leading to proposals for several candidate centres of support. The study of these ideas suggests some interesting problems in computational geometry.

The EGI has been applied to determine object attitude, the rotation in 3-space bringing a sample object into correspondence with a prototype. The methods developed for the inversion problem can be applied to attitude determination. Experiments show attitude determination using the new method to be more robust than area matching methods. The method given here can be applied at lower resolution of orientation, so that it is possible to sample the space of attitudes more densely, leading to increased accuracy in attitude determination.

The discussion finally is broadened to include non-convex objects, where surface orientation is not unique. The generalizations of the EGI do not support shape reconstruction for arbitrary non-convex objects. However, surfaces of revolution do allow a natural generalization of the EGI. The topological structure of regions of constant sign of curvature is invariant under Euclidean motion, and may be useful for recognition tasks.

Contents

Abstract	ii
List of Tables	vi
List of Figures	vii
Acknowledgement	ix
1 Introduction	1
2 Object Models for Vision	6
2.1 Introduction	6
2.2 Models for General Vision Tasks	6
2.2.1 Models and Representations	7
2.3 Orientation	8
2.3.1 Curvature	9
2.4 Area as a Function of Orientation	9
2.5 Geometric Models of Objects	10
2.5.1 Object Representation in Vision	10
2.5.2 Shapes in CAD/CAM	11
2.6 Polyhedral Models	12
2.6.1 Homotheticity	13
2.6.2 Support Functions	13
2.6.3 Orientation and Polytopes	14
2.6.4 Combinatorial Types	15
2.6.5 Metric Properties	16
2.6.6 Extended Gaussian Images of Polytopes	18

3	Recovering Shape from an Extended Gaussian Image	18
3.1	Introduction	18
3.2	Previous Work	18
3.3	Direct Methods	19
3.3.1	Area Variation	21
3.4	Minkowski's Fundamental Theorem	23
3.4.1	Linear Operations on Polyhedra	23
3.4.2	Volume of Mixtures	25
3.4.3	Brunn-Minkowski Theorem	26
3.4.4	Minkowski's Theorem on Polytopes with Given Area Functions	28
3.4.5	Example of Minimization	28
3.5	The Iterative Method	29
3.5.1	Constructing $P(H)$	29
3.5.2	Restoring Feasibility	29
3.5.3	Determining a Minimizing Step	29
3.5.4	The Method	31
3.5.5	Deficient Input	31
3.6	Complexity	32
3.7	Performance	32
3.7.1	Errors	34
3.8	Reconstruction from Partial Information	38
4	Support and Area Functions	41
4.1	Introduction	41
4.2	The Relation Between Support and Area Functions	41
4.2.1	Continuous Curves and Surfaces	41
4.3	Curve Segmentation	43
4.3.1	Centre of Support	43
4.3.2	Segmenting Curves at Extrema of \mathcal{H}	46
5	Determining Object Attitude from the EGI	48
5.1	Introduction	48
5.2	Determining Attitude	48
5.2.1	Quantizing Orientation	49

5.2.2	Quantizing Attitude	52
5.2.3	Attitude Determination	52
5.2.4	Attitude from Graph Matching	53
5.2.5	Attitude by Comparing Area Functions	53
5.2.6	Determining Attitude with Mixed Volumes	54
5.3	Experiments in Determining Attitude for Complete EGIs	56
5.3.1	Using Visible Hemispheres	60
5.4	Non-Convexity and Attitude	64
6	Non-convex Objects	66
6.1	Introduction	66
6.2	EGIs for Non-Convex Bodies	67
6.2.1	Curvature and the Gauss Map	67
6.2.2	Segmenting Surfaces	68
6.3	Reconstruction	70
6.3.1	Surfaces of revolution	72
6.4	Attitude Determination	73
6.5	Curvature Graphs and Recognition	73
6.6	Estimating Curvature	74
7	Conclusions and Open Questions	75
7.1	Open Questions	75
7.2	Future Work	77
	Bibliography	79
	A Volumes of Mixtures and Mixed Volumes	86
	B Matching EGIs Discretely	89
	C Example Polytopes	90

List of Tables

3.1	Error statistics for reconstructions	36
3.2	Combinatorial Structures	37
5.1	Errors with varying resolution of orientation	55
5.2	Errors at frequency 2 with varying attitude, axis (1,0,0)	55
5.3	Errors at frequency 2 with varying attitude, axis (0,1,0)	55
5.4	Errors at frequency 3 with varying attitude	56
5.5	Errors at frequency 5 with varying attitude	57
5.6	Errors with subsets at frequency 3 with varying attitude (8 faces)	59
5.7	Errors with subsets at frequency 3 with varying attitude (40 faces)	59
5.8	Errors with subsets at frequency 3 with varying attitude (21 faces)	60
5.9	Errors with subsets at frequency 3 with varying attitude (80 faces)	60

List of Figures

1.1	Stereo View of the EGI of a Distorted Octahedron	3
1.2	View of the Original Polyhedron	3
1.3	Initial Polyhedron	4
1.4	Reconstructed Polyhedron	4
2.1	The Gauss Map	8
2.2	Construction of the Support Function	14
2.3	The dual sketch	14
2.4	Constructing graph of support function	15
2.5	A polytope and its dual	17
3.1	The EGI of a Convex Polygon and Its Reconstruction, rotated by $-\pi/2$	19
3.2	Adjacency changes with position of faces	21
3.3	The sum of a square and a triangle	24
3.4	The mixture of a triangle and a square ; $\lambda = 0.66666$	24
3.5	Volume Contours, by The Brunn-Minkowski Inequality	27
3.6	Rectangle used in minimization example	27
3.7	Minimizing step remains close to constraint surface	30
3.8	Stereo View of the EGI of a Distorted Octahedron	33
3.9	View of the Original Polytope	33
3.10	Initial Polytope	33
3.11	Reconstructed Polytope	34
3.12	Stereo View of the EGI of a Polytope with 21 Faces	34
3.13	View of the Original Polytope	35
3.14	View of the First Estimate Polytope	35
3.15	View of the Reconstructed Polytope	35

3.16	Log of Error vs. Iteration for Example 1	38
3.17	Log of Error vs. Iteration for Example 2	39
3.18	Visible Hemisphere of a Polytope, and the Contour Generator	39
4.1	Lobes of a curve, bounded at minima of the support function	46
4.2	Slabs on a polygonal curve; region of maximum overlap is shaded	47
5.1	Icosahedron and Dodecahedron	49
5.2	Icosahedra Subdivided at Frequency 2 and Frequency 3	50
5.3	Form of entries in the comparison tables	55
5.4	Mixed volume (solid) vs. area (dotted), 21 faces, $f=3$, 20 deg.(1,0,0)	57
5.5	Polytope (40 faces) visible edges solid, invisible dotted	58
5.6	Polygon Completion: shortdashed=direct, dashed=spread,dotted=reflection	61
5.7	Non-convex object	63
6.1	Plane Curve, Hoffman decomposition, not 1-1 Gauss map	67
6.2	Plane curve and Curvature Regions Segmented at Vertical Normals	68
6.3	An object for which multiple-map EGI is not uniquely invertible	69
6.4	Elliptic parabolic and hyperbolic regions	69
6.5	Curvature Regions on a bat	72
A.1	Construction showing mixed volumes are equal	84
C.1	Polytope with 21 faces	88
C.2	Polytope with 40 faces	88
C.3	Polytope with 40 faces on ellipsoid	88
C.4	Polytope with 80 faces on ellipsoid	89

Acknowledgement

Thanks to Bob Woodham, my supervisor, who has been much more than a supervisor over the years we have worked together, mentor, supporter and friend. I have been helped by his inspiration and his advice on the enterprise of science. Thanks to Alan Mackworth, without whom the whole enterprise would not have even begun. Thanks to Dave Kirkpatrick, an exemplary researcher, who always urged me to sharpen the question. I am grateful to the other members of my committee, Jim Varah, and Andrew Adler, for their helpful comments.

Many graduate students have helped me to understand my work, Artificial Intelligence and Computer Science. Thanks especially to Jay Glicksman, Raimund Seidel, Bob Mercer, Jan Mulder, Marc Majka, and Randy Goebel.

The Laboratory for Computational Vision has been an exceptional resource, whose fine facilities allowed me to explore many uncharted regions.

Thanks to all the staff of the department, who provided relief, and the many members of the Department who have taught me so much.

Finally, thanks to Debra, my wife, who had to wait.

This work was supported by a National Science and Engineering Research Council of Canada Postgraduate Scholarship and a University of British Columbia Graduate Fellowship.

Chapter 1

Introduction

Problem solving by computer requires, first of all, a model of the problem domain. The relevant properties of the elements in the domain must be characterized and their relations must be analyzed. Then the elements of the model must be represented symbolically. Effective and efficient computation depends on the right choice of representation. For robotics and for computer vision, the question of representation arises at all stages in the process of converting sensed quantities into assertions about the world, and then into actions. The usefulness of a representation can be judged by the extent to which it facilitates the solution of the problem.

Computational vision seeks to model, first, the way images are formed by the interaction of light with the objects in a scene. For this purpose, the properties of objects and their surfaces which affect their appearance in an image are important. Albedo, surface roughness, and shape, among other properties, enter into the description of image formation. The work of Horn [1975], in particular, demonstrated how the appearance of an object depends significantly on its shape. The portion of computer vision known as image analysis or early vision seeks to devise methods for recovering the shape and location of visible surfaces from an image or images. This dissertation is concerned with surface representations which record surface properties as a function of surface orientation. Transformations among representations of objects, specifically polyhedral objects, are examined. A particular representation, the Extended Gaussian Image (EGI), is studied, and an algorithm for converting the EGI into a more conventional representation is given. The use of the EGI in determining the attitude of an object in space is explored, using the concepts developed in converting the EGI to conventional polyhedral representations.

Because of the recent successes of computer vision in computing surfaces from images, computer vision systems can now provide maps of surface orientation in a scene. Specifically, needle maps[Horn,1982], the "2½D sketch" [Marr,1976], and intrinsic images[Barrow and Tenen-

baum,1978] represent the orientation of a surface at the points of the image. Orientation maps can be generated by the output of stereo processing from several images [Grimson,1981, Baker and Binford, 1981], photometric stereo [Woodham,1980], or any of the so-called "shape from" methods, such as shape from shading [Horn,1975, Ikeuchi and Horn,1981], shape from contour [Marr,1977], shape from texture [Kender,1979, Witkin,1981], shape from edge interpretation [Mackworth,1973, Kanade,1981, Sugihara,1982], and by differentiation of laser range images [Brou, 1983]. By translating the surface normals of an object to a common point of application, a representation of the distribution of surface orientation is formed, called the Extended Gaussian Image (EGI) [Smith,1979]. The EGI of an object records the variation of surface area with surface orientation. The discrete version of the EGI is a histogram of area versus orientation.

Ikeuchi [Ikeuchi,1981] discussed recognizing objects in an industrial environment using the EGI. The EGI of the visible portion of an unknown object is formed by a constrained optimization method applied to data from photometric stereo[Ikeuchi and Horn,1981]. The prototype EGI which best matches this EGI identifies the object. The EGI rotates in the same fashion as the object in space. By comparing a sensed EGI with a prototype EGI of the same object, attitude can be determined [Ikeuchi et al.,1983].

Chapter 2 presents an overview of models and representations for objects and surfaces in computer vision. Particular attention is given to models using orientation to reference object properties. The EGI is such a model, describing area as a function of orientation. This study presents concepts based on orientation, in particular, the support function which describes the distance from the origin of a tangent plane. Lastly, this Chapter defines the properties of polyhedral objects and their surfaces used in the exposition.

In *Chapter 3*, the EGI as an object representation is studied. The EGI uniquely represents convex objects [Minkowski, 1897] and is easily derived from conventional models of an object. The inversion problem is to convert an EGI into an object model in terms of coordinates of vertices, edges, and faces. A iterative algorithm for reconstructing surface shape from an EGI is described. The algorithm converges to a solution by constrained optimization[Little,1983]. The algorithm employs a geometric construction, the mixed volume, which was used in Minkowski's proof of the existence of an inverse. The mixed volume is the product of the area function of one object and the support function of another. The objective function in the minimization is the mixed volume of the area function specified by the EGI and the support function of the reconstructed object. There are two aspects to describing shape for polyhedral objects: first, the way in which faces intersect each other, termed the adjacency structure, and, second, the location of the faces

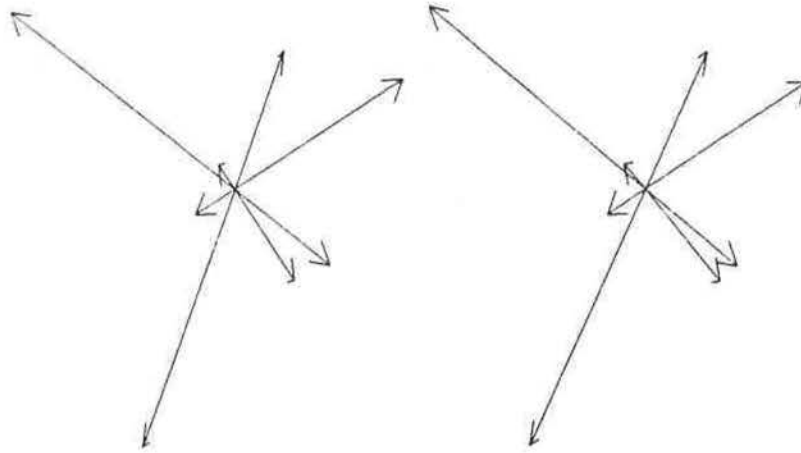


Figure 1.1: Stereo View of the EGI of a Distorted Octahedron

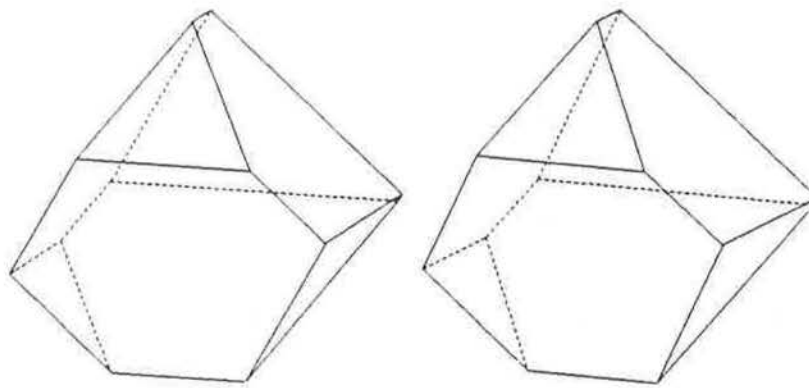


Figure 1.2: View of the Original Polyhedron

in space. The latter may change without altering the former, but not vice versa. The algorithm for shape recovery determines both elements of shape. The mixed volume itself is a measure of similarity of shape.

An example EGI is shown in Figure 1.1. The polyhedron to which the EGI corresponds is shown in Figure 1.2. The faces of the original polyhedron are parallel to those of a regular octahedron, but the distances of the faces from the origin have been altered. The algorithm initially constructs a regular octahedron, shown (in stereo) in Figure 1.3, which places all the faces tangent to the unit sphere. In the regular octahedron, each face is adjacent to three others. During minimization, the shape of the polyhedron constructed changes, and, at an early stage, the adjacency structure becomes identical to that of the original polyhedron. The final reconstructed polyhedron is shown in Figure 1.4. The reconstructed polyhedron has the same adjacency structure as the original polyhedron.

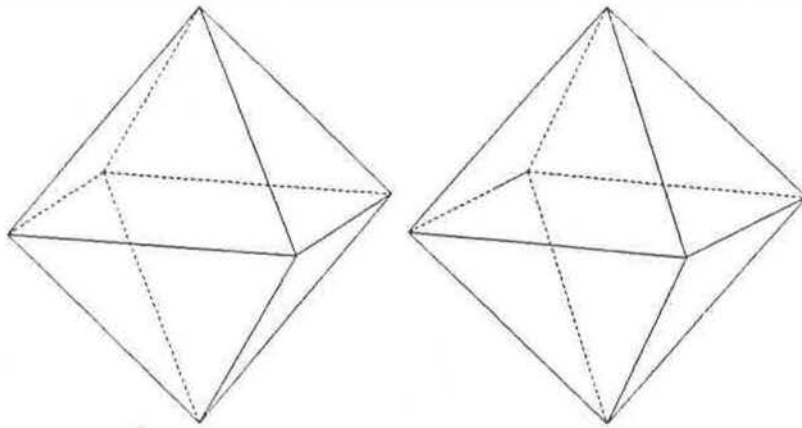


Figure 1.3: Initial Polyhedron

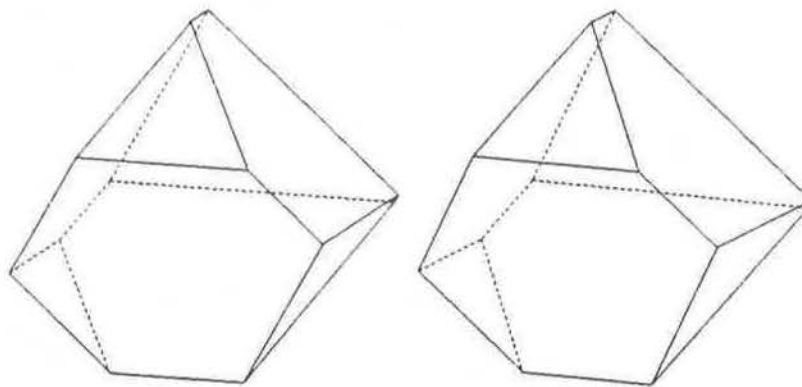


Figure 1.4: Reconstructed Polyhedron

The reconstruction algorithm shows how the support vector for a polytope can be determined from its area vector. In *Chapter 4*, to facilitate arguments about mixed volumes in *Chapter 5*, the continuous support function is described in terms of the area function for curves, permitting a qualitative comparison of the smoothness of the two functions. The next Section describes a method of curve segmentation based on extrema of the support function. Because the support function varies with translation, its behaviour under translation is studied, leading to proposals for several candidate centres of support. The study of these ideas suggests some interesting problems in computational geometry.

The EGI has been applied [Horn and Ikeuchi, 1984] in determination of object attitude, the rotation in 3-space which will bring a sample object into correspondence with a prototype. *Chapter 5* examines previous methods which have relied on direct comparison of the sensed EGI with the prototype EGI. The methods developed for the inversion problem can be applied to attitude determination. The rotation which minimizes the mixed volume of the support function of the prototype EGI and the sensed area function identifies the attitude of the sensed object. Experiments show attitude determination using mixed volume to be more robust than area matching methods. The mixed volume method can be applied at lower resolution of orientation, so that it is possible to sample the space of attitudes more densely, leading to increased accuracy in attitude determination. Experiments with non-convex planar figures suggest that an extension of the proposed method to non-convex objects is possible.

Mathematical results about area functions, support functions and EGIs have mainly concerned convex bodies. In *Chapter 6*, the focus is broadened. On non-convex objects, surface orientation is not unique. The accumulation of area in an EGI must account for this in some fashion. One solution is to segment the object surface by the sign of curvature, keeping a separate EGI for each region. For surfaces of revolution, such segmentation, in principle, allows shape reconstruction. An analysis of the requirements for general solutions is also given. The topological structure of the curvature region graph is invariant under Euclidean motions, and may be useful for recognition tasks.

Chapter 7 concludes with suggestions for future work and open problems. These centre, first, on improvements to the reconstruction algorithm and extensions to smooth surfaces. Secondly, a host of computational geometry problems emerge from the investigations. Lastly, generalizations of the EGI and other uses of the EGI are discussed.

Chapter 2

Object Models for Vision

2.1 Introduction

Models and representations are fundamental for problem solving. Models select characteristic properties of an object. Representations describe the object properties selected by the model to facilitate solution of a class of problems. In early vision, a model must account for several types of sensed quantities, i.e., brightness, hue, saturation, depth, orientation, sparsity or density of measurements. When sensed data and a model are in a similar form or can be easily transformed into similar forms, computation of their relationship simplifies.

New methods of acquiring surface information (photometric stereo, structured light and laser range scanners) allow acquisition of dense measurements of surface orientation. The availability of these data prompts investigation of models using orientation as reference coordinate, to take direct advantage of these data. Data on surface orientation can be readily transformed into representations such as the Extended Gaussian Image, which are well-suited to solution of certain practical problems. To investigate orientation-based models, the common object models of vision and Computer-Aided Design (CAD) are discussed. Planar-faceted models are studied, since they permit simple formulation of problems under investigation. Definitions of polyhedral models are given, stressing properties used in later investigations.

2.2 Models for General Vision Tasks

A model facilitates the representation of aspects of reality useful in a particular problem domain [Bolles et al.,1983]. The utility of a representation is influenced by the combination of the sensor, the task and the object domain. In this dissertation, attention is restricted to a situation in which

dense information about surface orientation is available.

2.2.1 Models and Representations

In a model for a class of individuals, abstraction and simplification occurs. The information necessary to identify an individual is selected. Models divide the world into equivalence classes by the mappings induced by characteristic properties. Modelling is a purposive activity, always with a view toward the application.

A representation is a symbol system in which elements are identified and the basic operations on elements and transformations among elements are defined. The mapping between the representation and the model provides the semantics for the representation. When a representation is chosen, the boundary between the explicit and implicit properties is drawn, often by the computation necessary to make explicit that which is implicit. Throughout this dissertation, the focus will be on transformations on representations on the surfaces of objects. These transformations are computations, the complexity of which will enter the discussion. Algorithms are descriptions of computations, which deliver the result of the computation in a finite number of steps, each representing a primitive operation. The exact model of computation will not be critical here, but it may be useful to compare the amount of time algorithms require. The time complexity of an algorithm is given as a function of the size of the input n , using the following notation for functions of n :

A function $f(n)$ will be termed $O(g(n))$ if there exist c, n_0 such that

$$f(n) \leq cg(n), \text{ for } n > n_0$$

A function $f(n)$ will be termed $\Theta(g(n))$ if there exist c_0, c_1, n_0 such that

$$c_0g(n) \leq f(n) \leq c_1g(n), \text{ for } n > n_0$$

A function $f(n)$ will be termed $\Omega(g(n))$ if there exist c_0, n_0 such that

$$c_0g(n) \leq f(n), \text{ for } n > n_0$$

A function $f(n)$ will be termed *exponential in n* if there exist c_0, c_1, k_0, k_1, n_0 such that

$$c_0k_0^n \leq f(n) \leq c_1k_1^n, \text{ for } n > n_0$$

Returning to modelling, the connection between model, representation and data can be seen by considering the circle:

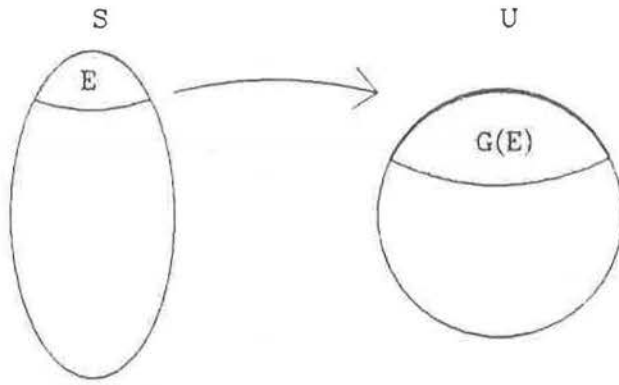


Figure 2.1: The Gauss Map

- a convex figure of constant curvature
- the locus of points at a fixed distance from a point

Both are descriptions or models of circles but their suitability for a particular problem domain depends on the form of the available data. Where curvature is primitive or easy to compute, the first should be used; where distance between elements is simple, the latter should be used.

Vision is concerned with the interaction of light with surfaces to create images. The properties of albedo, surface roughness, shape, and location enter directly in the description of the image formation process. Horn's pioneering work on shape from shading [1975] showed the connection between shape and brightness modulation. Shape is a global property which emerges from the accumulation of the local orientation of the surface. Orientation will serve as a unifying reference property for the models presented herein, so now a precise definition is given.

2.3 Orientation

At each point on a smooth surface S a unit surface normal ω is defined. The direction of the normal vector is the *orientation* of the surface at that point. A unit normal vector can be uniquely associated with a point on the unit sphere U . The map taking surface normals of S onto U is the Gauss map [DoCarmo, 1976]:

$$G(p) = \omega, p \in S, \omega = \text{unit normal at } p, \omega \in U \quad (2.1)$$

Let E be a portion of S bounded by a closed curve. The image under the Gauss map of E , $G(E)$, is the *Gaussian image* of E (Figure 2.1).

2.3.1 Curvature

Curvature describes, informally, the rate of change in the orientation of a curve or surface. Let a plane curve $c(s)$ be defined:

$$c(s) = \{(x(s), y(s)), 0 \leq s \leq s_{max}\}$$

such that $|c'(s)| = 1$. The curve c is then said to be parametrized by arc length. The number $|c''(s)| = k(s)$ is called the *curvature* of c at s . The curvature of a plane curve at a point p is the reciprocal of the radius of curvature at p , which is found by taking the limit of the radii of circles through p and two points approaching p from both sides along the curve. At any point on a surface, a normal section is the intersection of the surface with a plane containing the normal at the point. The curvature of the surface in a given direction is the curvature of the normal section in that direction. At any point on a surface, there are two orthogonal directions at which the curvature is, respectively, maximum and minimum. These directions are the principal directions and the curvatures are the principal curvatures. The Gaussian curvature, K , is the product of the principal curvatures. Gaussian curvature can be described using the Gauss map: let $|R|$ be the area of a region R , then

$$K(p) = \lim_{|E| \rightarrow 0} |G(E)|/|E|, E \text{ a compact portion of } S \text{ enclosing } p \quad (2.2)$$

On an object bounded by planar facets, Gaussian curvature is zero on the facets, and undefined at the intersection of the facets. The Gauss map will take all points on a facet into a single point.

2.4 Area as a Function of Orientation

On a smooth surface of strictly positive curvature, the Gauss map is unique. Any point on the surface can be uniquely identified by its orientation: let $G^{-1}(\omega)$ be the inverse image of ω under G , then $p = G^{-1}(\omega)$. An area function A can be defined for any orientation, in an analogous manner to the definition of the Gaussian curvature from the Gauss map in Equation 2.2.

$$A(\omega) = \lim_{|F| \rightarrow 0} |G^{-1}(F)|/|F|, F \text{ a compact portion of } U \text{ enclosing } \omega \quad (2.3)$$

The area function on U describes the variation of surface area with orientation and so it is analogous to the Extended Gaussian Image (EGI) of an object. Later, the EGI for objects bounded by planar facets will be described.

2.5 Geometric Models of Objects

2.5.1 Object Representation in Vision

Vision systems utilize many different representation schemes, each tailored to a sensor mode and a task. The needs of vision systems arise primarily from the demands of recognition. Recognition is posed as matching descriptions derived from images with models of objects. Object properties which can be reliably determined from images must be described. Polyhedral models have been with us since the early days of blocks-world research [Guzman, 1968] [Huffman, 1971] [Waltz, 1972]. The interpretation of edges and junctions of edges in line-drawings employed the constraints in polyhedral representations. Many schemes rely on simple classes of object models, such as polyhedral models, or schemes with a set of generic objects (polyhedra, second-degree surfaces such as spheres, cylinders or ellipsoids), or objects composed of patches with generic shape, either planar patches or quadric patches. The features for recognition allowed by a scheme of course depend on the vocabulary in which they are expressed. 3DPO [Bolles et al., 1983] hypothesizes the existence of cylindrical portions of objects from the conjunction of circular arc features and edges at the occluding boundary of the cylinder.

The requirement of natural descriptions for image understanding motivates attention to volumetric properties of objects, characterization of the distribution of the mass and general trends of the object shape [Nishihara, 1981]. Surface properties, in contrast, are necessarily local. Many systems use the notion of a central axis of an object. An early scheme focussing on the object axis was the *Symmetric Axis Transform* of Blum [1967], which derives an object axis as the locus of the centres of maximal enclosed circles. It has been extended to R^3 by Blum [1979] and studied by Nackman [1982]. In R^3 it has the drawback that the axes generated are not curves but actually symmetry surfaces, so that the descriptive power is somewhat diluted. Binford [1971] described an alternative method, the *generalized cylinder* which consists of a central axis (a curve in R^3), a cross-section shape, and a sweeping rule specifying the change in scale of the cross-section shape with distance along the axis. The representation was fundamental to the ACRONYM system [Brooks, 1981] which analyzed the projection of generalized cylinders (*ribbons*) to provide cues for recognition [Lowe, 1984]. Brady and Asada [1983] use *smoothed local symmetries* to represent boundaries of regions, capturing significant changes in curvature and deriving a local symmetry axis for parts of the regions.

In addition, the description of geometric features of objects can support attitude determination. Here the requirements of vision systems merge with those of manipulation systems. The

same features can be used for both recognition and manipulation - corners or junctions between planar patches are good indicators for both tasks.

Systems vary in their ability to compute accurately the properties described here, orientation and area. Generalized cylinders are useful in recognition tasks [Lowe, 1984], but computing orientation and area can be problematic; their definition does not prevent self-intersection, which must be determined prior to area computation. Another volumetric system, [Mohr and Bacjcsy, 1983], describes a shape by the centres of spheres packed into the enclosing volume. The graph induced by adjacencies among the spheres characterizes the shape. Clearly such a representation attends to volumetric considerations at the expense of surface properties.

Bajcsy [1980] has criticized the EGI on the grounds that it "does not preserve connectivity and therefore part/whole relationships are hard to identify". It is exactly because the EGI ignores connectivity that it simplifies processing. The shape recovery method to be described shows how connectivity can be recovered. Part/whole relationships in the EGI, as in all representations, rely on accurate segmentation. Feature-based methods for determining attitude can be inaccurate when there is difficulty in precisely localizing the required features. Methods based on the EGI are potentially more robust since attitude determination becomes a global computation. Further, the EGI can deal with surface orientation directly, thus avoiding the integration step necessary to pass from surface orientation to a surface depth map. This is useful with techniques like photometric stereo where surface orientation is determined directly.

2.5.2 Shapes in CAD/CAM

Manipulation poses new problems for a vision system, beyond the construction of an interpretation. The system must not only recognize objects, but also, for example, determine appropriate grasp points, and feasible approaches for a grasping arm. Many tasks envisioned for a robot with a vision apparatus occur in an industrial environment. Computer-based methods have had a strong impact on the manufacturing process. Recently, Computer Aided Design (CAD) and Computer Aided Manufacture (CAM) have expanded from research environments into production situations. Since the objects of robotic manipulation are the products of a computer-aided manufacturing process, it is natural to consider representational methods from these disciplines. Knowledge and understanding of Solid Modelling, the representation and construction of 3-dimensional objects by computer, is currently advancing rapidly in CAD/CAM [Requicha, 1980]. Three-dimensional objects have been described by wireframe models tracing the edges (wires) of the planar faces of the model. These representations were extended to boundary representations which specify the

intersections of the (simple) surfaces bounding an object. With both wire-frame and boundary models, it is necessary to check that the resultant surface descriptions fulfill topological and geometrical consistency constraints, for example, that there are no holes not explicitly described (topology) and that no edges cross (geometry).

Swept-volume methods describe a volume by an axis, a cross-section area, and a sweeping rule relating the scaling of the cross-section to position on the axis. Constructive Solid Geometry (CSG) [Brown and Voelcker, 1979] describes objects by the Boolean combination of primitive 3-dimensional elements such as spheres and rectangular parallelepipeds. Its advantages lie in the assurance that the resultant models fulfill topological and geometrical consistency constraints. The primitive vocabulary of CSG usually consists of elements which can be described by surface equations at most of the second degree, i.e., quadric surfaces.

The form of object representation for computer vision remains an open question. The methods of CAD/CAM facilitate precise measurement and description of the volumetric and surficial properties of an object, but lack the ability to describe simple generalization or hierarchy, properties useful for a system for recognition and description [Mulder, 1985]. Nevertheless, by providing a description of the surfaces of objects these systems can serve as models for calculating surface orientation and area for Extended Gaussian Images of prototypical objects.

2.6 Polyhedral Models

Extended Gaussian Images uniquely represent convex objects. A set C is *convex* if and only if for every pair of distinct points $a, b \in C$, the closed segment with endpoints a and b is contained in C . A set is *strictly convex* if and only if each point on the closed segment ab is contained in the interior of C , i.e., is surrounded by an open ball entirely within C .

The treatment of polyhedra will follow the terminology of Grünbaum [1967]. A minimal number of geometric terms will be defined in order to provide enough terminology for discussion. Additional definitions can be found in Grünbaum [1967]. The term *polyhedron* often loosely refers to any body whose boundary is composed of a finite number of planar facets. In general polyhedra correspond to the class of objects representable by wire-frame models. A plane J can be represented as:

$$J = \{x \mid \langle \omega, x \rangle = c\}, \text{ where } \omega \text{ is a unit vector normal to the plane} \quad (2.4)$$

A plane forms the boundary of a *half-space*, $\{x \mid \langle \omega, x \rangle < c\}$, for a suitable choice of orientation ω and c . The intersection of a finite number of half-spaces forms a convex polyhedron; its boundary

is composed of planar facets. A bounded convex polyhedron will be termed a *polytope*. The term *d-polytope* is an abbreviation for *polytope of dimension d*. Under these definitions, a polytope has a definite location in space. For the purposes of all further discussion, the normal to a facet of a polytope will be outward facing, i.e., it will point away from the half-space bounded by the plane in which the facet lies.

2.6.1 Homotheticity

Two 3-polytopes P and Q are *homothetic* if

$$P = \{x \mid x = \lambda * y + t, y \in Q, \lambda \in R, \lambda > 0, t \in R^3\}$$

This relation is similar to the relation of congruence, which is invariance under translation, scaling, and rotation. Homotheticity is invariance only under translation and scaling. This notion is appropriate for the present setting in which measurements are tied to orientations of faces, and rotation is a parameter which must be determined. In particular, the location in space of a polytope is not relevant to its description for the context of EGLs. It may be taken to be situated so that its centroid coincides with the origin. However, any rotation of a polytope is critical, since it changes the orientation of facets of the polytope.

2.6.2 Support Functions

A *support plane* J for a convex body C is a plane such that all of C lies on one side of the plane and C has at least one point in common with J , that is, J is tangent to the surface of C . Recall the definition of a plane (Equation 2.4). The *support function* of C can be defined:

$$h(\omega) = \max\langle \omega, x \rangle, x \in C.$$

Then the plane $J = \{x \mid \langle \omega, x \rangle = c\}$ *supports* C at the orientation ω . Figure 2.2 shows the construction of the support function for a polygon. For an orientation ω corresponding to the normal of an edge of the polygon, the support function coincides with the value c in the equation of the supporting plane through that edge. For orientations ω not identified with some edge of the polygon, the supporting plane $\langle \omega, x \rangle = c$ is incident upon a vertex v_j of the polygon. Thus

$$h(\omega) = \langle \omega, v_j \rangle$$

Between two adjacent faces, $h(\omega)$ can be rewritten as

$$h(\omega) = |v_j| \langle \omega, \omega_j \rangle \tag{2.5}$$

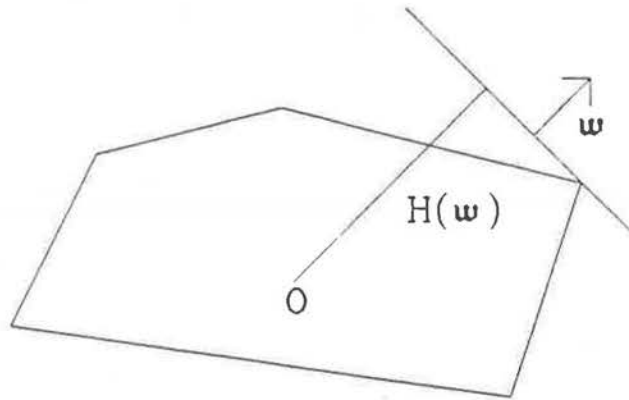


Figure 2.2: Construction of the Support Function

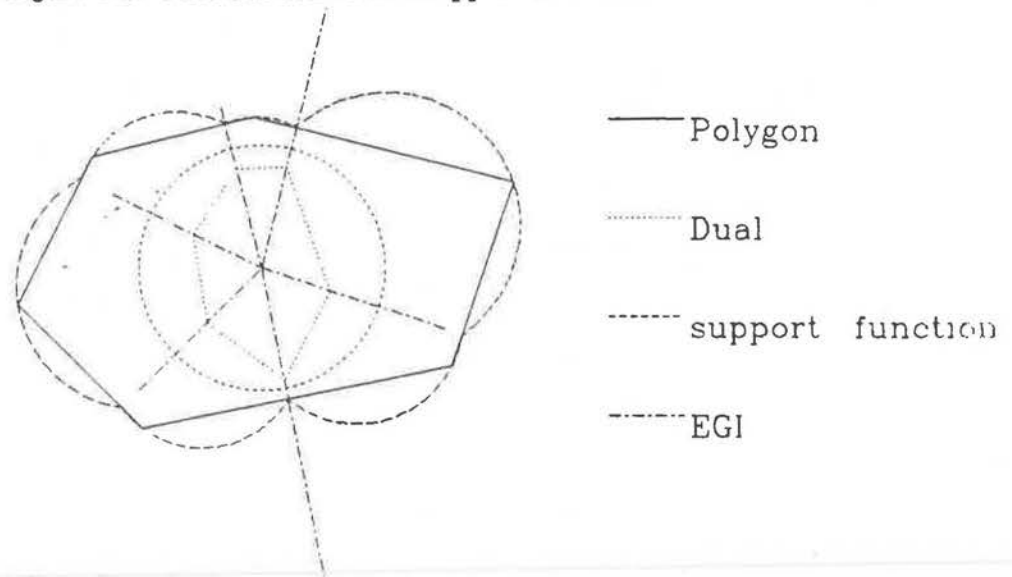


Figure 2.3: The dual sketch

where ω_j is the normalized direction vector of v_j . So $H(\omega)$ varies as the cosine of the angle between ω and ω_j . Drawn as a function of ω , it is composed piecewise of portions of circles with centres at the midpoints of the position vectors of the vertices v_j , with radius $|v_j/2|$ (see Figure 2.4).

Figure 2.3 shows a polygonal curve, its EGI, its dual (to be described in Section 2.6.5) and its support function shown as a function of ω . Figure 2.4 shows the support function of the polygonal curve alone.

2.6.3 Orientation and Polytopes

The set of orientations of the faces of a polytope is termed Ω . The EGI implicitly specifies Ω for the polytope from which it is derived. This set of orientations will be referenced by indices from $1 \dots n$ without implying any particular ordering. When it becomes necessary to compare

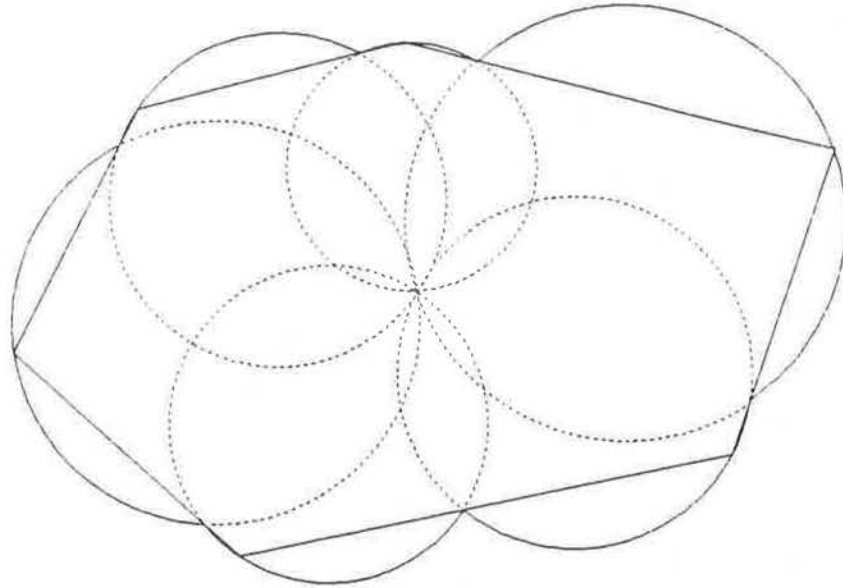


Figure 2.4: Constructing graph of support function

two polytopes, P_1 and P_2 , with face orientations Ω_1 and Ω_2 , the two sets of face orientations can be merged into a common set of orientations Ω . To describe the EGI of P_1 in terms of Ω , simply augment its list of faces with new faces of zero area for ω_i not in Ω_1 . For all polytopes, faces will be referenced by the index of the orientation vector ω_i . The composition of Ω should be clear from the context.

2.6.4 Combinatorial Types

A *face* F of a polytope P is the intersection of a supporting plane of P with P . When the intersection is a point, it is called a *vertex* of P , when it is a line segment, an *edge*, and when it has dimension 2 it is termed a *facet* of P . Two polytopes P and P' are *combinatorially equivalent* or of the same *combinatorial type* if there is a one-to-one function ϕ between the set $\{F\}$ of all faces of P and the set $\{F'\}$ of all faces of P' , such that ϕ is inclusion-preserving, i.e. for $F_1, F_2 \in \{F\}$, $F_1 \subset F_2$ if and only if $\phi(F_1) \subset \phi(F_2)$. Hereafter, the term *face* will be used interchangeably with *facet*. For the purpose of this discussion, the combinatorial type describes the incidence relations in a polytope: which faces share an edge, edges meet in a vertex, vertices lie on a face. The terms *adjacency structure* and *combinatorial structure* will also denote this set of incidence relations. In many cases, the simplest description of the combinatorial structure, a list of the face incidences, will be used. In this list, for each face, the set of faces incident upon the face is specified.

2.6.5 Metric Properties

Many polytopes realize the same combinatorial type, but differ in the placement of the vertices, and the orientation of the faces. There are many alternative formulations of the metric properties of a polytope. The specification can involve the location of the vertices, the location of the planes bounding the polytope, the lengths of the edges in the polytope, etc. Consider the set of polytopes in which the orientations of the faces are restricted to a given finite set of orientations. Each face lies in a bounding plane of the polytope. A plane with fixed orientation has one degree of freedom, (see Equation 2.4). The simplest representation of a polytope when face orientations are fixed is the location of the planes bounding the polytope. The faces of the polytope can be referenced by their indices $1 \dots n$ corresponding to the face orientation; the values of the support function $\mathcal{H}(\omega)$ at the face orientations form an n -tuple $H = (h(\omega_1), h(\omega_2) \dots h(\omega_n))$. A polytope with n faces having orientations in Ω corresponds to a point H in the space described by the generalized coordinates $(h_1, h_2 \dots h_n)$. The space described by the generalized coordinates is termed *support space*. Each point H in support space specifies the location of n planes in R^3 . The point H corresponds to a polytope, $P(H)$, constructed by their intersections. $P(H)$ has a definite location in R^3 , but this location is irrelevant for descriptions based on surface orientation. Consider the coordinates in R^3 of the centroid c of a polytope $P(H)$. Let the coordinates in support space of the polytope when its centroid is translated to the origin be termed H_0 ; there the coordinates are all positive. The difference between H and H_0 is, where c is the centroid:

$$h_i = \mathcal{H}(\omega_i) = \mathcal{H}_0(\omega_i) + \langle \omega_i, c \rangle$$

Let $\omega_{i1}, \omega_{i2}, \omega_{i3}$ be the x, y and z coordinates of direction ω_i , then a translation by 1 in the j^{th} coordinate corresponds to addition to the support vector of:

$$\nu_j = (\omega_{1j}, \omega_{2j}, \dots, \omega_{nj}) \quad (2.6)$$

Points in support space representing translations of $P(H_0)$ by $t = (t_1, t_2, t_3)$ can be described in the form:

$$H = H_0 + \sum_{j=1}^3 t_j \nu_j$$

where the spanning vectors are ν_j . The set of polytopes which are similar except for translation form a flat of dimension 3 in support space. All computations described hereafter will implicitly recognize this fact. In some cases, the exact location of the polytope is irrelevant; otherwise, without loss of generality, the polytope can be considered to have its centroid at the origin.

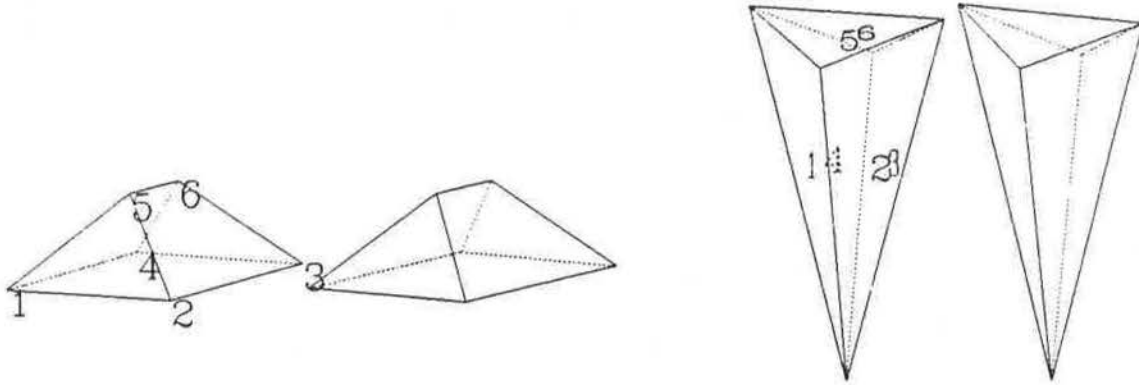


Figure 2.5: A polytope and its dual

The *skeleton* of a polytope lists the combinatorial structure. The full specification then describes the coordinates of the vertices. Each vertex is associated with the list of edges incident upon it and each face with the list of its bounding edges. Each of these lists can be ordered, for example, the bounding edges by counter-clockwise order around the face, in the sense determined by the surface orientation. The skeleton of a 3-polytope can be transformed into a description by bounding planes in $\Theta(V)$ time, where V is the number of vertices, since the number of faces is $\Theta(V)$ and the description of each bounding plane supporting a face can be determined from 3 vertices on the boundary of the face.

The inverse transformation moves from the plane locations H to the vertex description. To construct a polytope $P(H)$, form the intersection of the n half-spaces specified by the vector H . Brown [1978] describes a method for transforming the problem of intersecting n half-spaces into a convex hull problem, using the dual transform, described in the vision literature by [Huffman,1971, Mackworth,1973, Draper,1981]. The dual transform maps a plane with equation

$$\langle \omega_i, x \rangle = h_i \quad (2.7)$$

into the point ω_i/h_i in R^3 (see Figure 2.5). Providing h_i is not 0 for any i , the planes containing faces of $P(H)$ do not pass through the origin, so Equation 2.7 will be defined for all faces. The n planes forming $P(H)$ correspond to n points in R^3 , for which the algorithm of Preparata and Hong [1977] determines the convex hull in $\Theta(n \log n)$ time. Any face of the convex hull of the dual points corresponds to a vertex of P . Any two points incident on an edge in the dual of P correspond to a pair of faces of P which share an edge. The adjacency information in the dual

provides the adjacency information for P . Hence the locations of the vertices and edges of P can be calculated.

2.6.6 Extended Gaussian Images of Polytopes

The areas of the faces of a 3-polytope and their orientations can be computed, from the description of the skeleton of a polytope and the location of its vertices, in time linear in the number of faces. This follows directly from the fact that the area of a face can be computed in time linear in the number of incident vertices and the fact that, for a 3-polytope, the numbers of faces and vertices are linearly related. The definition of area for continuous surface (Equation 2.3) does not cover polytopes, since curvature vanishes on the faces of polytopes. The analogous area function for polytopes is equal in value to the area of the face having orientation ω when ω is the normal to a face, else it has the value zero. For orientations which do not coincide with a face of the polytope, $A(\omega)$ is 0. This function is non-zero only at a finite set of points on U , and can be represented as a vector A of elements corresponding to the values of $A_i = A(\omega_i)$, where the area function is non-zero. Fenchel and Jessen [1938] showed that it is possible to construct a measure on U for a convex body C which is the area function of C , and which reduces to the continuous area function described above as well as to the area of facets in a polytope. So, the two functions can be considered specializations of this generalized measure.

From a image containing surface orientation, area can be calculated as the product of the pixel size and the inverse of the cosine of the angle between the viewing direction and the surface normal. Each surface normal is multiplied by the area, and moved to the origin. Should several vectors coincide, they are added as vectors. This representation of the distribution of surface orientation is termed the Extended Gaussian Image (EGI). The set of orientations present in an EGI derived from an image will lie in a hemisphere of the Gaussian sphere centred on the orientation recording the view direction. This forms a *visible hemisphere* of the EGI. The term EGI will refer to the map of the area function on the entire Gaussian sphere, unless otherwise noted. It is assumed, for the sake of the exposition, that only a single object is imaged.

The Extended Gaussian Image of a polytope P with m faces can be described by a set of vectors $N(P)$, indexed by orientations of the faces of P :

$$N(P) = \{n_i | 1 \leq i \leq m\} \quad (2.8)$$

$$n_i = \omega_i A_i, \quad A_i = \text{the area of face } i$$

Minkowski[1897] showed that if

$$\sum_{i=1}^m n_i = 0$$

then N uniquely represents a polytope, up to translation.

Chapter 3

Recovering Shape from an Extended Gaussian Image

3.1 Introduction

An iterative algorithm for recovering surface coordinates from the EGI of a polytope is given. The algorithm utilizes a geometric construction, the *mized volume*, arising from the theory of mixtures of polytopes. The reconstruction algorithm employs constrained optimization to recover surface shape. Its implementation and performance are discussed and evaluated. This algorithm is generalized for shape recovery from partial information.

3.2 Previous Work

Minkowski proved the existence and uniqueness of the polytope corresponding to an EGI. It is easy to derive the skeleton of the polytope from its support function, which need only be specified at the faces (see Section 2.6.5). The support function describes the locations in R^3 of the half-spaces forming the polytope $P(H)$. What must be determined is the vector $H = (\mathcal{H}(\omega_1), \mathcal{H}(\omega_2), \dots, \mathcal{H}(\omega_n))$. The problem is to compute the support function H from the area vector A and the vector of face orientations Ω .

Ikeuchi[1981] proposed an algorithm for reconstructing a polytope from its EGI. The problem is subdivided into n distinct cases; in the i^{th} case, face i is farthest from the origin. In case i , h_i is set to 1.0; all other h_j vary between 0.0 and 1.0. The $n - 1$ dimensional space of distances is quantized (at spacing $d < 1$). Each of the d^{n-1} locations in this space specifies a possible location of the n faces in R^3 . The polytope can be constructed, and the areas of its faces determined.

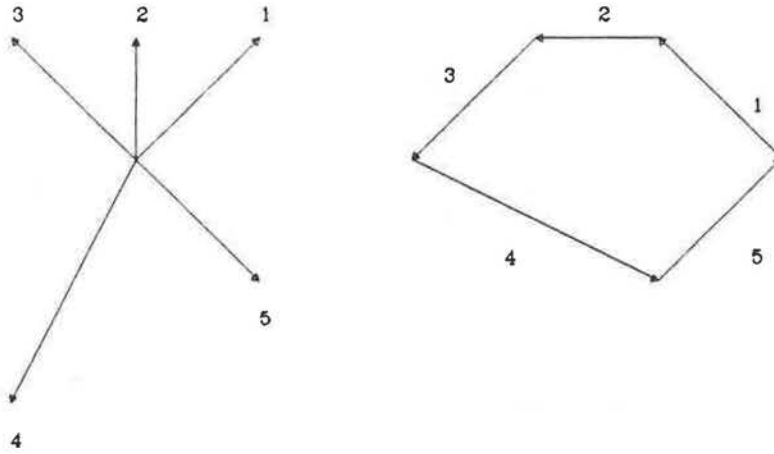


Figure 3.1: The EGI of a Convex Polygon and Its Reconstruction, rotated by $-\pi/2$

These areas are scaled and compared with the objective.

No analysis of the accuracy of the algorithm is supplied. The method minimizes the sum of the squared differences between the calculated areas of the polytope and the given areas in the EGI. It is not clear that the polytope which results from this minimization (after normalizing) will have the same combinatorial structure as the desired polytope. In addition, the method is very expensive. The polytope is constructed at each evaluation point. If the resolution in the location of the bounding planes is doubled ($D = d/2$), the the number of evaluation points increases to $D^{1-n} = d^{1-n}2^{n-1}$, which is an increase by a factor exponential in n , the number of faces.

3.3 Direct Methods

A direct solution by a geometric construction is possible for the two-dimensional case. The EGI of a polygon is a system of vectors emanating from the origin. Any system summing to zero represents a convex polygon. Figure 3.1 shows a two-dimensional EGI and the reconstructed polygon.

Mackworth [private communication, 1982] noted the following simple procedure for constructing the polygon from the system of vectors:

The vectors $\{u_i\}$ are given in anti-clockwise order. Rotate u_1 anti-clockwise by $\pi/2$ and place its tail at some point in the plane. In order, rotate u_i anti-clockwise by $\pi/2$

and place its tail at the head of u_{i-1} . Because the system sums to zero, the head of u_n will close with the tail of u_1 . By definition, the length of each vector is the length of the corresponding edge, and its orientation is normal to that of the edge. Each edge in the reconstructed polygon will be the correct length and at the proper orientation.

The two-dimensional method does not directly extend to higher dimensions. In particular, in two dimensions, the adjacencies among the edges is explicit in the EGI. In three dimensions the adjacency relationships are not explicit in the EGI and must form part of the solution.

The number of different adjacency relations for polytopes with n triangular faces is asymptotically exponential in n [Tutte, 1962]. The number of general polytopes (with faces having any number of sides) is larger. Hence any method which examines all possible adjacency relations will, in the worst case, take exponential time.

A particular combinatorial type may be realized by various assignments of orientations to faces. The assignment of orientations from Ω to the faces in the combinatorial type corresponds to a labeling of the graph induced by the type. There are $n!$ labelings of the incidence graph; many, however, are unrealizable.

A combinatorial type, together with an assignment of face orientations, determines a set of linear equations. These equations indicate the incidence of the vertices of the polytope on the faces of the polytope. The incidence of a vertex v_m on a face i (see Figure 3.2) determines an equation:

$$\langle \omega_i, v_m \rangle = h_i$$

The number of equations involving a vertex is the degree of its incidence specified by the combinatorial type. The skeleton of a 3-polytope is planar and 3-connected [Steinitz, 1922], so there are $O(n)$ equations specifying the incidence of vertices on faces on a 3-polytope with n faces. For a vertex v_m not incident on a face j , there are $O(n^2)$ inequalities

$$\langle \omega_j, v_m \rangle < h_j$$

many of which are redundant. The cardinality of the set of redundant constraints can be determined by the methods of Sugihara [1982]. When these equations do not have any solution, the combinatorial type is not realizable for the particular assignment of orientations to faces. Whether these linear equations have a solution satisfying the linear inequalities can be solved by the methods of linear programming [Papadimitriou and Steiglitz, 1982].

Each combinatorial type forms a convex region in the n -dimensional space referenced by H . Any convex combination of two polytope vectors H_0 and H_1 will also satisfy the systems of

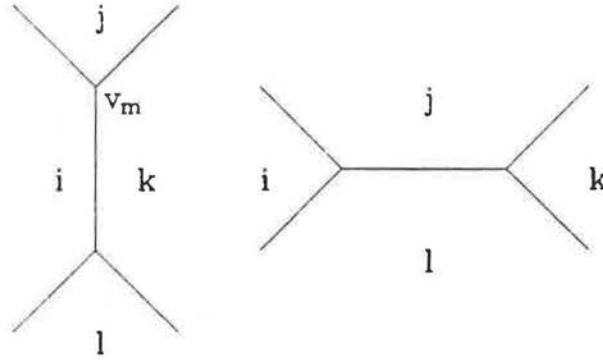


Figure 3.2: Adjacency changes with position of faces

equalities, since they involve only linear functions. Since all regions are convex, they must be bounded by hyperplanes.

3.3.1 Area Variation

The combinatorial type of a polytope P with n faces at given orientations Ω varies with H . For each given combinatorial type, the area of the faces in P will also vary with H . The combinatorial type of P specifies the incidences of faces on vertices. The area of a particular face is a function of the coordinates of the incident vertices, and can be determined by projection onto a plane parallel to the face, where the formula for plane figures can be used. When such an area equation is expanded in terms of the variables h_i , the resulting equations for the area of the i^{th} face is:

$$A_i = \sum_{j \in \{\text{Neighbors}(i) \cup i\}} c_{ij} h_i h_j \quad (3.1)$$

The coefficients c_{ij} are constants which derived from the normal vectors of the faces. The number of terms in this formula is small. The total number of non-zero elements in the matrix (c_{ij}) is $O(n)$. Each coefficient is generated by an edge in the polytope; the number of edges is $O(n)$. Over all A_i , the number of non-zero elements is also $O(n)$, since each edge only occurs in the expression for the area of two faces.

Equating these formulae with the n area values specified by the EGI generates, for each combinatorial type, a system of n second degree equations in H . Closed-form solution of such a system is in general possible for $n \leq 2$. For $n > 2$ these equations may be solvable in closed-form, but a suitable reduction has not been found. Thus approximate numerical solution seems indicated. Morgan [1982] describes a suitable procedure using a continuation method. The form of these equations varies with combinatorial type.

It is interesting to note that all tetrahedra with specific orientations to their faces are homothetic, and thus the only free variable in their reconstruction is scaling. Another way of determining the number of degrees of freedom is to recall that the flat of points representing polytopes similar up to translation is of dimension 3 (see Section 2.6.5). Since the space has dimension 4, that leaves one degree of freedom for scaling.

The volume of a polytope P is a function of H . By triangulating the faces, and connecting these triangles to the centre of gravity of P , the volume can be subdivided into tetrahedra. The volume of each tetrahedron is:

$$V = 1/3A * h$$

where A is the area of the triangular base, and h the height of the tetrahedron. Taking the centre of gravity of P as the point from which the support function $h_i = \mathcal{H}(\omega_i)$ is computed, and aggregating the areas of the triangles into the areas of faces of P , the volume can be expressed:

$$1/3(A_1 h_1 + A_2 h_2 + \dots + A_n h_n) = V(P) = 1/3 \sum_{i=1}^n A(\omega_i) \mathcal{H}(\omega_i)$$

Consider the gradient of V expressed as a function of coordinates in support space:

$$\nabla V = (\partial V / \partial h_1, \partial V / \partial h_2 \dots \partial V / \partial h_n) = 1/3(A_1, A_2 \dots A_n)$$

The gradient of the volume is proportional to the area vector of the polytope. In an EGI, the vector A is specified, and, in reconstruction, the values of H are sought. The volume of a polytope links these two. This fact would support a naive method for recovering shape from an area function: search all polytopes with specified face orientations, choosing that polytope with gradient proportional to the given area vector. The volume of the polytope is irrelevant, since changing volume scales the gradient. The intuition provided by this informal analysis underlies the proof of Minkowski's theorem, which is analyzed in the following Section.

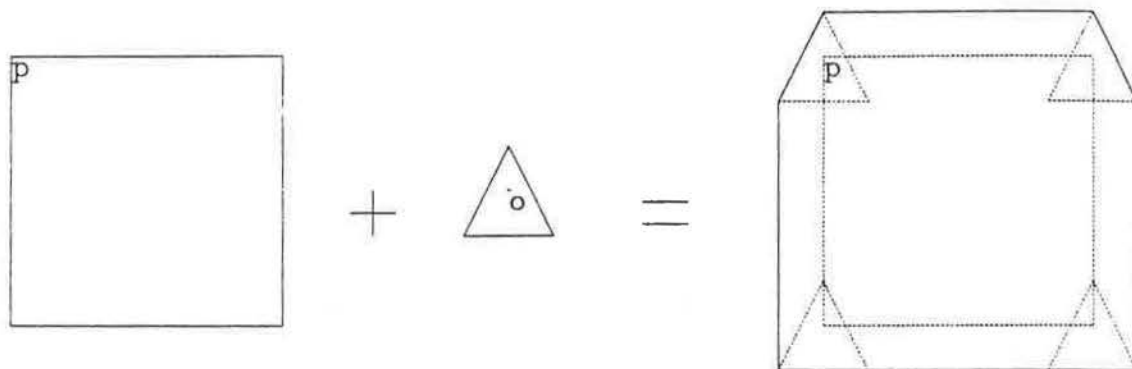


Figure 3.3: The sum of a square and a triangle

3.4 Minkowski's Fundamental Theorem

Minkowski's proof provides clues for finding a reconstruction method. The original proof considers polytopes in any dimension d ; here the proof is described with $d = 3$ for clarity. For a 3-polytope P , the set of vectors is formed, as described in Equation 2.8. A set of vectors N is *equilibrated* if and only if it sums to zero and no two vectors are positively proportional, i.e., no two are positive linear multiples of a common unit vector. An equilibrated set of vectors N is *fully equilibrated* if and only if it spans R^3 . Minkowski's polytope reconstruction theorem shows that

1. if P is a polytope in R^3 not contained in any plane then the $N(P)$ is fully equilibrated and
2. if N is a fully equilibrated system of vectors, then there exists a polytope P unique within a translation such that N is the EGI of P .

This description is taken from [Grünbaum,1967,p.332].

3.4.1 Linear Operations on Polyhedra

The proof of Minkowski's theorem depends on several facts derived from the study of the behavior of polytopes under linear operations. The Minkowski sum of two polytopes P and Q is:

$$P + Q = \{x + y \mid x \in P, y \in Q\}$$

A polytope is considered a set of points, but it will shortly be seen how the sum of two polytopes can be expressed using the notion of support function and therefore support planes (bounding planes).

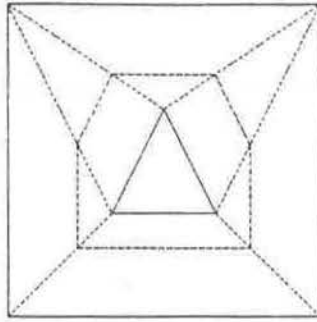


Figure 3.4: The mixture of a triangle and a square ; $\lambda = 0.66666$

Figure 3.3 shows the addition of a square and a triangle. The origin lies at the centre of S . Move point o in the triangle T to coincide with point p on the boundary of the square S , and draw the boundary of T . The boundary of the triangle drawn at that position represents the boundary of the set of points Q of the form:

$$Q = \{x + p \mid x \in T\}$$

By moving the triangle around S , the boundary of their sum is generated. This operation can be described as a convolution [Guibas, Ramshaw and Stolfi,1983]. The sum of T and S would superimpose on the square S , but is drawn separated for clarity.

Likewise, scalar multiplication for polytopes is defined using multiplication in R^3 :

$$\lambda * P = \{\lambda * x \mid x \in P\}$$

The convex sum or *mixture* of two polytopes P and Q is

$$\lambda * P + (1 - \lambda) * Q = \{\lambda * x + (1 - \lambda) * y \mid x \in P, y \in Q, 0 \leq \lambda \leq 1\}$$

Figure 3.4 shows the mixture of the triangle T and the square S , in the form :

$$\lambda * T + (1 - \lambda) * S, \text{ where } \lambda \text{ is } \frac{2}{3}$$

Their mixture is drawn as a figure lying between T and S . The vertices of their mixture are the mixtures of the corresponding vertices of the two polygons. Correspondence is established by orientation. The dotted lines in Figure 3.4 show connections between corresponding vertices.

Support functions of mixed polytopes behave simply under mixing, as described in the following theorem, stated in Lyusternik[1963]:

Let $\mathcal{H}_P(\omega)$ be the value of the support function of P evaluated at ω and let $\mathcal{H}_Q(\omega)$ be the value of the support function of Q , then the support function of their mixture, $R = \lambda * P + (1 - \lambda) * Q$ can be expressed as the mixture of the support functions:

$$\mathcal{H}_R(\omega) = \lambda * \mathcal{H}_P(\omega) + (1 - \lambda) * \mathcal{H}_Q(\omega)$$

The area function of a polygon is simply the length of the side at the given orientation; if there is no such side the value is zero. An argument using similar triangles shows that the area function for the mixture $R = \lambda * P + (1 - \lambda) * Q$ is:

$$A_R(\omega) = \lambda * A_P(\omega) + (1 - \lambda) * A_Q(\omega) \quad (3.2)$$

that is, the lengths of the faces of the mixture are equal to the mixtures of the face lengths of P and Q .

3.4.2 Volume of Mixtures

Consider the expression for the volume of the mixture of two 2-polytopes in terms of λ . It is shown in Appendix A that the volume of the mixture $V(R)$ is:

$$V(R) = \lambda^2 V(P) + 2\lambda * (1 - \lambda) V_2(P, Q) + (1 - \lambda)^2 V(Q) \quad (3.3)$$

where

$$V_2(P, Q) = 1/2 \langle H_P, A_Q \rangle$$

and A_Q is vector of edge lengths in Q , H_P is the vector of support values for P evaluated at the orientations in Ω . The notation is from Appendix A, where the proof is developed for polygonal curves.

The expression $V_2(P, Q)$ is called the *mixed volume* of polygons P and Q , by analogy to the expression of volume (area) of a polygon. It is shown in Appendix A that $V_2(P, Q) = V_2(Q, P)$. Minkowski [1897] showed that a similar expression arises in the formula for volume of 3-polytopes. The generalization of Equation 3.3 to 3 dimensions is:

$$V(R) = \lambda^3 V(P) + 3\lambda^2 * (1 - \lambda) V_3(P, Q) + 3\lambda * (1 - \lambda)^2 V_3(Q, P) + (1 - \lambda)^3 V(Q) \quad (3.4)$$

The coefficients appearing in the formula for $V(R)$, the mixed volumes, are defined as follows:

$$V_3(P, Q) = 1/3 \langle H_P, A_Q \rangle$$

These are expressions got by replacing the area function of one polytope with that of the other. In $V_3(P, Q)$ the area function of Q replaces the area function of P . In 3 dimensions, in contrast to 2, except when P is a translate of Q ,

$$V_3(P, Q) \neq V_3(Q, P)$$

The terminology $V_i(P, Q)$, $i = 2, 3$ departs from the more generalized conventional notation, but is used here since the only results necessary for the exposition occur in 2 and 3 dimensions and involve only two bodies.

The mixed volume of P and Q is independent of the point from which the support function is calculated. Here the support function of Q is denoted by \mathcal{H}_Q . If Q is translated by t , becoming Q' , the support values of its faces are altered:

$$H_{Q'_i} = \mathcal{H}_{Q'}(\omega_i) = \mathcal{H}_Q(\omega_i) + \langle t, \omega_i \rangle$$

$V_2(Q', P)$, after translation, is:

$$\begin{aligned} \sum_{i=1}^n A_{P_i} H_{Q'_i} &= \sum_{i=1}^n A_{P_i} H_{Q_i} + \sum_{i=1}^n A_{P_i} \langle t, \omega_i \rangle \\ &= \sum_{i=1}^n A_{P_i} H_{Q_i} + \sum_{i=1}^n \langle t, A_{P_i} * \omega_i \rangle \\ &= \sum_{i=1}^n A_{P_i} H_{Q_i} + \langle t, \sum_{i=1}^n A_{P_i} * \omega_i \rangle \end{aligned}$$

and, by the closure condition,

$$\sum_{i=1}^n A_{P_i} * \omega_i = 0$$

so

$$V_2(Q', P) = V_2(Q, P) \tag{3.5}$$

and the mixed volume is invariant under translation.

3.4.3 Brunn-Minkowski Theorem

The Brunn-Minkowski theorem describes the relation between the volume of the mixture of P and Q and the volumes of the individual polytopes. The Brunn-Minkowski theorem states, in R^3 , that

$$V(R)^{1/3} \geq \lambda V(P)^{1/3} + (1 - \lambda)V(Q)^{1/3} \tag{3.6}$$

The exponent is $1/d$, where d is the dimension of the minimal imbedding space. Brunn showed that the inequality holds. Minkowski showed that equality holds if and only if P and Q are homothetic.

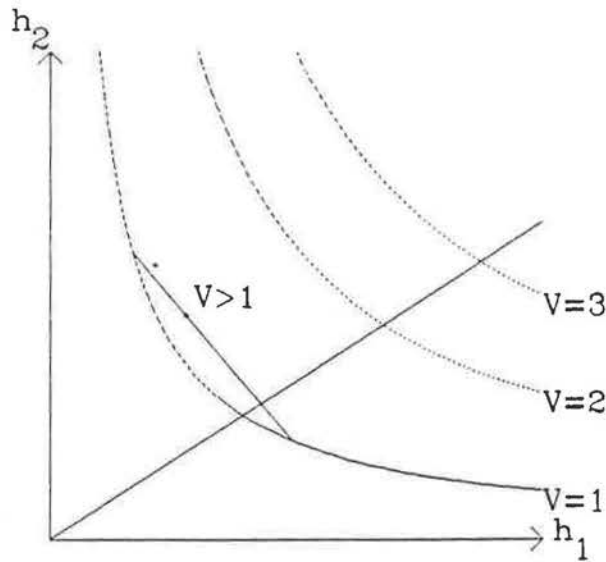


Figure 3.5: Volume Contours, by The Brunn-Minkowski Inequality

Figure 3.5 shows the volume contours in a support space of two dimensions. Assume, for example, that the two support values represent the location of edges of a rectangle with orientations $\omega_1 = (1, 0)$ and $\omega_2 = (0, 1)$. The other two edges, with orientations $(-1, 0)$ and $(0, -1)$, are assumed to pass through the origin in R^2 . Each point having positive coordinates represents a rectangle with edge lengths h_1 and h_2 . The rectangle is depicted in Figure 3.6; it corresponds to the point (b, a) in support space. The contours of constant volume (area) in support space are hyperbolae with equations:

$$h_1 h_2 = c$$

The polygons represented by points along the ray from the origin are homothetic. Points along the chord indicate a mixing between the two polygons represented by the endpoints. Let $V(H)$ be stand for $V(P(H))$, where $P(H)$ is the polytope described by H . Any interior point on the chord H_R has volume $V(H_R) > 1$, by the Brunn-Minkowski inequality.

After substituting the right hand side of Equation 3.4 in Equation 3.6, suitable algebraic manipulation yields:

$$(1/3\langle H_P, A_Q \rangle)^3 = V_3(P, Q)^3 \geq V(Q)^2 V(P) \quad (3.7)$$

Equality holds only when P and Q are homothetic. Homotheticity is very important; if two polytopes are homothetic, their area functions can be made equal by the proper scaling.

An interpretation of these formulas is that the mixed volume captures the relation between the shapes of the two polytopes; when the mixed volume is minimal, the shapes are homothetic, and mixing the two polytopes does not result in shape change, only in scaling. Otherwise, the mixing results in both shape and scale change.

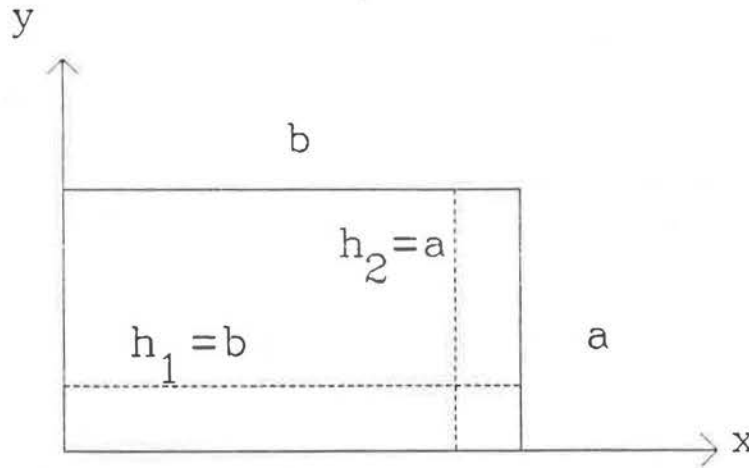


Figure 3.6: Rectangle used in minimization example

3.4.4 Minkowski's Theorem on Polytopes with Given Area Functions

Equation 3.7, which follows from the Brunn-Minkowski theorem, was used by Minkowski to transform a question of uniqueness into a minimum problem. Let Q be the polytope with area vector A_Q . A polytope P will be constructed which is homothetic to Q . Consider only polytopes which have unit volume. Then, by Equation 3.7,

$$(1/3\langle H_P, A_Q \rangle)^3 = V_3(P, Q)^3 \geq V(Q)^2$$

$V(Q)$ is not known, but is fixed. The Brunn-Minkowski theorem says that the polytope P , having unit volume, that minimizes $V_3(P, Q)$ is homothetic to Q . Scaling the minimizing polytope P so that $V(P) = V(Q)$ results in a polytope whose area vector is equal to the desired vector A_Q . The set of polytopes $\{P(H) \mid V(H) \geq 1\}$ is convex, as a consequence of the Brunn-Minkowski inequality. The objective function, $\langle A_Q, H_P \rangle$, since it is linear, is convex, so the minimum will lie on the boundary of the convex set, where $V(H) = 1$. By convexity, a local minimum of $\langle A_Q, H_P \rangle$ is the global minimum. Thus the polytope which minimizes the mixed volume is unique.

The proof relies on the minimization to establish uniqueness. In addition, Minkowski argues that the polytope satisfying the minimization criterion in fact has the desired area vector. In the following discussion, it will be shown that known methods of constructing a polytope from its support vector and known minimization techniques can be combined to construct the support function of P so that $A_P = A_Q$.

3.4.5 Example of Minimization

Consider the set of rectangles whose support space, previously defined, is pictured in Figure 3.5. A particular rectangle, with edge lengths (a, b) at orientations (ω_1, ω_2) corresponds to the point

in support space at (b, a) . Its area is ab , and its area vector $A = (a, b)$ (see Figure 3.6). After substituting the right hand side of Equation 3.4 in Equation 3.6, One can solve explicitly for the point on $h_1 h_2 = 1$ which minimizes $\langle A, H \rangle$; its coordinates are $(\sqrt{b/a}, \sqrt{a/b})$. Scaling by $A^{1/2} = \sqrt{ab}$ recovers the correct coordinates for $(h_1, h_2) = (b, a)$. In general, explicit solution is not possible, so the problem is solved by constrained minimization.

3.5 The Iterative Method

In an iterative solution to a constrained optimization problem, a sequence of feasible points, i.e. points satisfying the constraints, is generated, which converges to the optimum [Gill et al., 1981]. The sequence of polytopes is generated, using a procedure for constructing a polytope $P(H)$ from its support vector H . Each point H is transformed into a point satisfying the volume constraint by computing its volume, $V(H)$, and scaling $P(H)$ so that its volume satisfies the constraint. This permits the generation of a convergent sequence of feasible points by starting from an initial point, taking a step toward the minimum, restoring feasibility, and repeating.

3.5.1 Constructing $P(H)$

A polytope $P(H)$ can be constructed in $\Theta(n \log n)$ time from the intersection of the half-spaces specified by the vector H , as described in Section 2.6.5.

3.5.2 Restoring Feasibility

Once $P(H)$ has been constructed, it is straightforward to determine a corresponding feasible point H' . The volume $V(H)$ of a 3-polytope $P(H)$ is a homogeneous polynomial in H of degree 3. Any $P(H)$ can be scaled by $V(H)^{1/3}$, yielding a corresponding polytope $P(H')$ with unit volume, (see Figure 3.5). The formula for $\nabla V(H)$ can be derived directly from the formula for $V(H)$. The gradient ∇V is used in computing the minimizing step.

3.5.3 Determining a Minimizing Step

Constrained optimization is a well-studied problem, so many methods are available for determining the step direction and magnitude [Gill et al., 1981]. The reduced gradient method is a simple method which was chosen for implementation. A step in support space in the hyperplane perpendicular to $\nabla V(H)$ remains close to the constraint surface $V(H) = 1$. The step is in the direction which minimizes $\langle A, H \rangle$, that is, in the direction of the projection of the vector $-A$, the

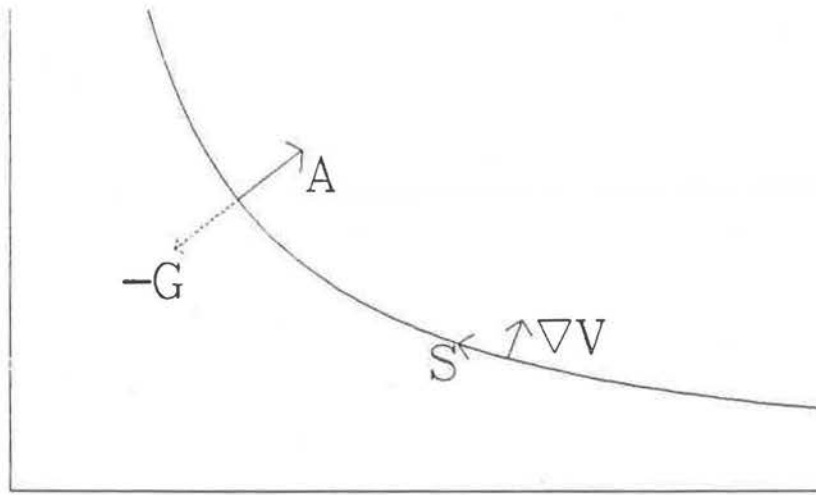


Figure 3.7: Minimizing step remains close to constraint surface

n-vector of areas of the faces given by the EGI, onto the hyperplane perpendicular to $\nabla V(H)$. This step S is a multiple of:

$$\langle A, \nabla V(H) \rangle \nabla V(H) - A \quad (3.8)$$

The minimizing step is depicted in Figure 3.7, in an analogous situation to that shown in Figure 3.5.

Determining the step size is an important factor in the minimization. Initially, step size is set to unity, A is normalized and the first evaluation point is at $(1, 1, \dots, 1)$. The magnitude of the step is adjusted during minimization. Should the step be too large, the value of the objective function for the next iteration may increase when the point is restored to feasibility. This is termed *overshoot*. When the step size is too small, the rate of convergence decreases. There are several ways of handling overshoot – one obvious way is to reduce the step size. In this particular problem, the gradient of the constraint surface is available at each step, as well as the values of the objective function at p_i and p_{i+1} , the positions in support space of the i^{th} and $i + 1^{\text{th}}$ evaluation points. The line $p_{i+1} - p_i$ and the direction A describe a 2-dimensional plane. In this plane, the behavior of the constraint can be approximated by a parabola using the gradient, p_i and p_{i+1} . The minimum of this parabola is used as a new p_{i+2} when overshoot occurs. Essentially, this method computes a local approximation to the Hessian, the matrix of second derivatives of the constraint surface. More sophisticated methods, such as quasi-Newton methods, build up over all iterations an approximation of the Hessian for second-order information. Using the local parabola is effective in improving the convergence when overshoot occurs. After overshoot, the step size is reduced to the length of the vector connecting p_i and p_{i+2} . When several steps have been taken at this new size and no overshoot occurs, the step size is gradually increased.

In this particular problem, the actual Hessian is sparse; each element is:

$$G_{ij} = \partial V / \partial h_i \partial h_j = c_{ij}$$

where the c_{ij} are described in Equation 3.1 above. Solving a sparse system of linear equations may be linear, depending on the structure of the matrix, so moving to a method based on second-order information could be advantageous in this problem.

3.5.4 The Method

The iterative method for reconstructing a convex polytope from its EGI combines the procedures described above. It is a minimization which terminates when the objective function $\langle A, H \rangle$ decreases by less than a prescribed ϵ in some step. Initially, all faces are adjacent to the unit sphere in R^3 , so H is $(1, 1, \dots, 1)$. The process of generating a polytope, scaling, and moving in a minimizing direction is repeated until $\langle A, H \rangle$ decreases by less than ϵ :

1. Construct $P(H)$:
 - (a) Map the n planes given by H into M , a set of n points in R^3 , using the dual transform.
 - (b) Compute the convex hull of M , $CH(M)$.
 - (c) Determine the adjacency relations of $P(H)$ from $CH(M)$. Calculate the locations of the vertices of $P(H)$.
2. Compute $V(H)$ and $\nabla V(H)$. Scale H by $V(H)^{1/3}$ to make its volume unity.
3. Compute a step using Equation 3.8 and update H .

3.5.5 Deficient Input

The necessary condition on A for the existence of the convex reconstruction of shape is:

$$\sum_i A_i \omega_i = 0$$

which in effect is three conditions in R^3 of the form

$$\sum_i A_i \omega_{ij} = 0$$

where ω_{ij} is the j^{th} coordinate of ω_i . The vectors formed by collecting the ω_{ij} are the spanning vectors of the subspace of polytopes related by translation (see Equation 2.6). Should, in practice, the vector A not satisfy the closure condition, the minimization will not succeed. There is no point

H in the space for which $\nabla V = A$. Essentially, A has a component in the translation subspace, and ∇V is always orthogonal to that subspace. The minimization may cycle and terminate with a result where the area vector is far different from A .

The simplest remedy is to solve, instead, some similar vector A^* , such that

$$\sum_i A_i^* \omega_i = 0$$

and A^* is suitably close to A , say that it minimizes

$$\sum_i (A_i^* - A_i)^2$$

3.6 Complexity

The requirements of the reconstruction procedure can be factored into two components: the number of iterations required to find an acceptable solution and the number of operations per iteration. Each iteration requires $O(n \log n)$ operations to compute the convex hull of the n dual points. In addition, $O(n)$ operations are necessary to evaluate the volume. Each iteration thus requires $O(n \log n)$ computations. The number of iterations depends on the constrained minimization method. The convergence rate of an iterative method is linear if the error at step i , ϵ_i , satisfies:

$$|\epsilon_{i+1}| \leq \gamma |\epsilon_i| \text{ for } i \text{ large enough} \quad (3.9)$$

where $\gamma < 1$. A reduced gradient method [Gill et al., 1981] was implemented; its convergence rate is linear. To achieve quadratic convergence, i.e., $|\epsilon_{i+1}| \leq \gamma |\epsilon_i|^2$ for i large, the Hessian matrix of $V(H)$ or an approximation to the Hessian must be used; the Hessian in this problem is sparse (see Section 3.3.1). This was not implemented but could be advantageous. Using the Hessian might impose excessive computational demands, suggesting a quasi-Newton approach.

It is possible to start at each step with the hull of the previous dual polytope. The change in location of the primal planes in most cases is very small, so the change in the adjacency structure will be small. Recall that each point p_i in the dual polytope corresponds to face i in the primal, and will lie along a ray which is the extension of ω_i . In the initial stages of the minimization, adjacencies change rapidly, later little is altered. As implemented, the hull of the dual polytope is reconstructed anew at each step, in $O(n \log n)$ steps. It would be expected that starting from the previous hull, the new hull would be easy to compute, since the change in the position of the dual points is small. However, Seidel [1984] demonstrated that, in the worst case, constructing the 3-d hull of a set of n points, even when starting from the previous hull, requires $\Omega(n \log n)$

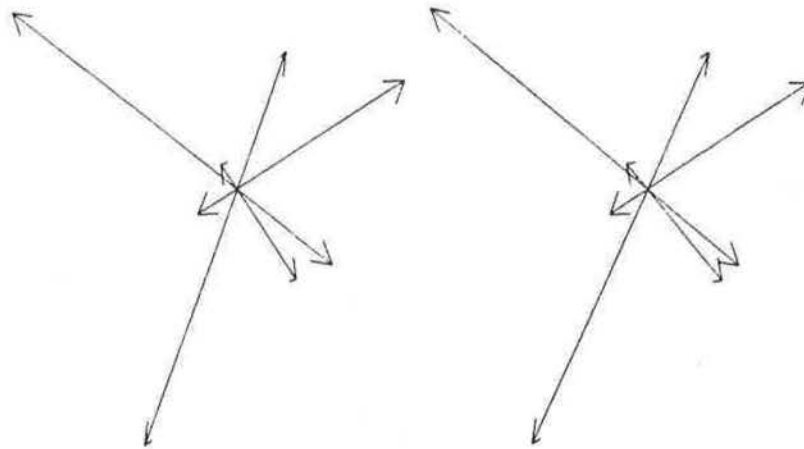


Figure 3.8: Stereo View of the EGI of a Distorted Octahedron

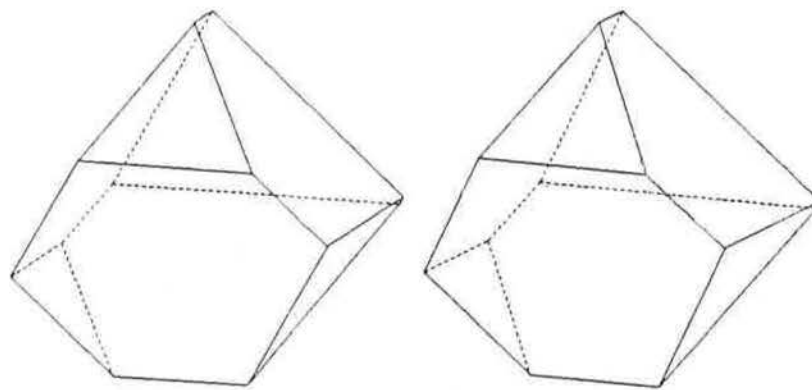


Figure 3.9: View of the Original Polytope

operations. In practice, starting from the previous polytope may speed things up considerably, but this was not implemented.

3.7 Performance

An example polytope P has been reconstructed from its EGI, (Figure 3.8). The polytope P to which the EGI corresponds is shown in Figure 3.9.

The faces of the polytope are parallel to those of a regular octahedron, while the distances of the faces from the origin have been altered. The polytopes generated during reconstruction will be denoted $P_i, i = 0 \dots s$, where the subscript indicates the step number. The polytope constructed initially, P_0 , is shown (in stereo) in Figure 3.10.

P_0 is an octahedron, in which each face is adjacent to three others. During minimization, the polytopes generated change in adjacency structure. The adjacency structure at an early stage

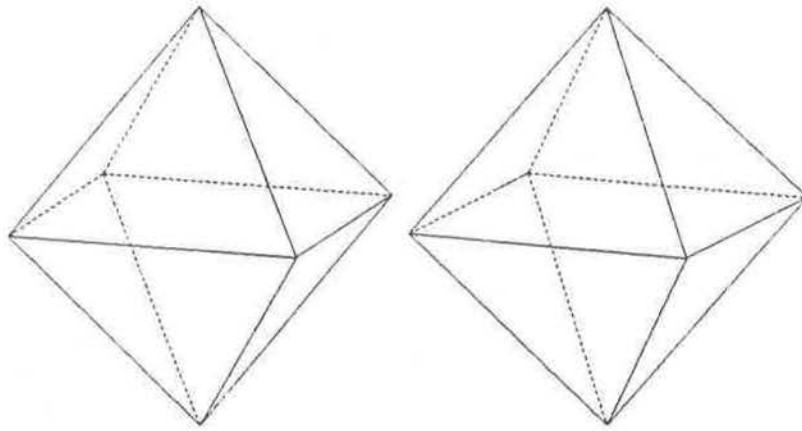


Figure 3.10: Initial Polytope

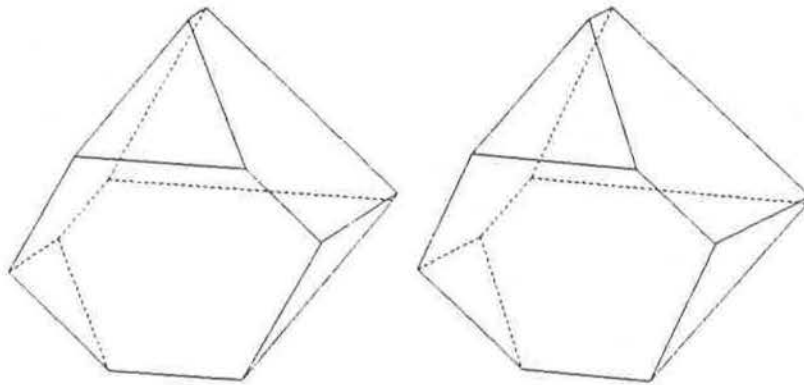


Figure 3.11: Reconstructed Polytope

becomes identical to that of P . The final reconstructed polytope P_8 is shown in Figure 3.11. P_8 has the same adjacency structure as P .

Another polytope, Q , having 21 faces intersecting out of 40 randomly positioned planes, is reconstructed from its EGI (in Figure 3.12). Q is shown in Figure 3.13. Q_0 is shown in Figure 3.14. Q_{18} is shown in Figure 3.15. Reconstruction terminated after 22 steps when $\langle A, H \rangle$ had decreased by less than $1e - 5$ on successive steps.

3.7.1 Errors

Let A and H be the calculated areas and support values, and A and H the actual areas and support values.

The statistics for the two reconstructions are presented in Table 3.1. The procedure for the polytope with 8 faces required 8 steps, and 9 evaluations of intermediate polytopes, terminating when the mixed volume reached 1.240818, and the actual value of the mixed volume for the original polytope is 1.240818. The combinatorial types of the original and reconstruction are

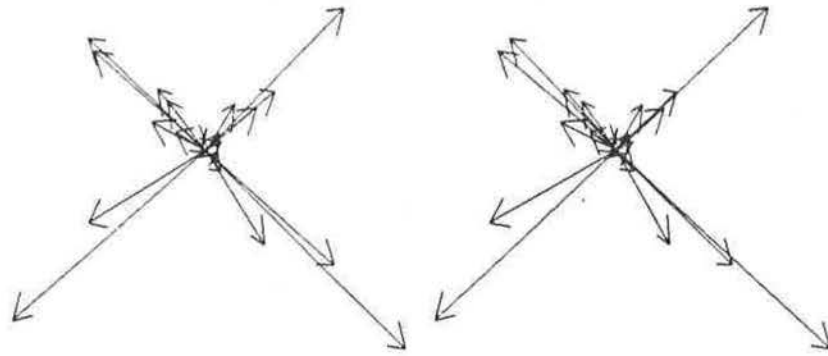


Figure 3.12: Stereo View of the EGI of a Polytope with 21 Faces

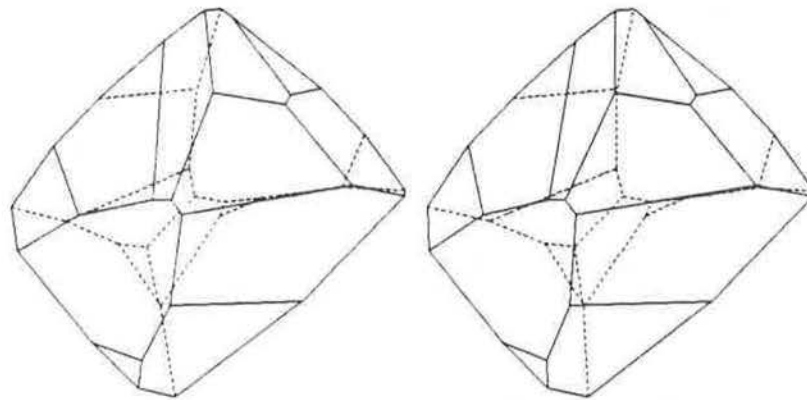


Figure 3.13: View of the Original Polytope

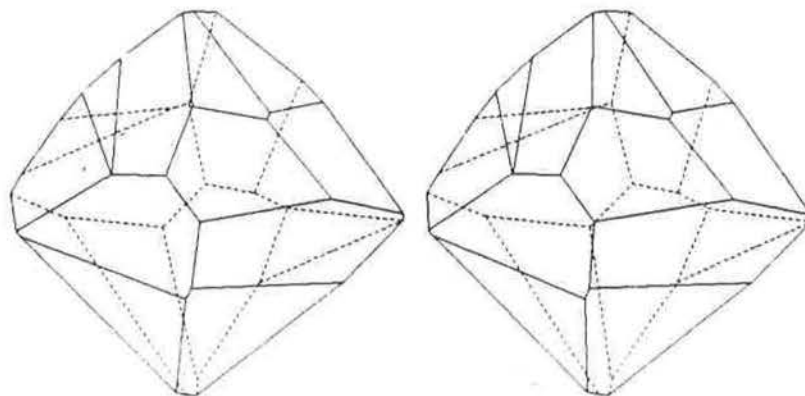


Figure 3.14: View of the First Estimate Polytope

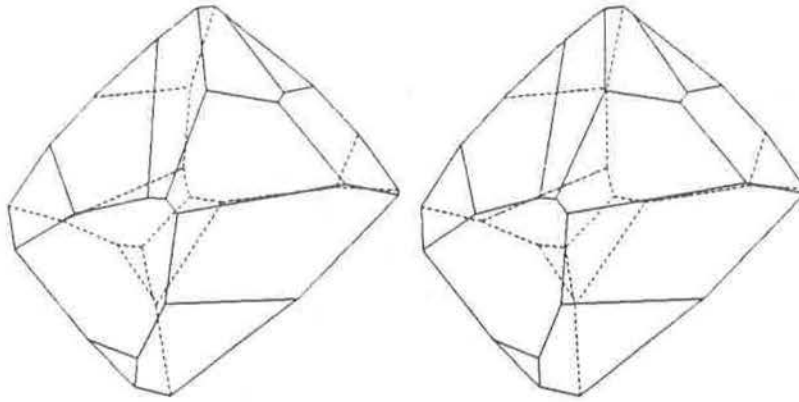


Figure 3.15: View of the Reconstructed Polytope

Object	Area error / Area	Average ΔA	Max ΔA	Average ΔH	Max ΔH
	$\frac{\sum A_i - \hat{A}_i }{\sum A_i}$	$\frac{\sum A_i - \hat{A}_i }{n}$	$\max_i \frac{ A_i - \hat{A}_i }{A_i}$	$\frac{\sum H_i - \hat{H}_i }{n}$	$\max_i \frac{ H_i - \hat{H}_i }{H_i}$
8 faces	0.08 %	0.151 %	0.62 %	0.06 %	0.16 %
21 faces	1.60 %	6.60 %	100 %	1.70 %	30.1 %
20 faces	0.61 %	1.93 %	20.8 %	0.28 %	0.63 %

Table 3.1: Error statistics for reconstructions

identical.

The procedure for the polytope with 21 faces required 22 steps, and 26 evaluations of intermediate polytopes, terminating when the mixed volume reached 1.917301, while the actual value of the mixed volume for the original polytope is 1.91655. There were 4 more evaluations than steps because of overshoots, which result in reduction of the step size. The discrepancy between the achieved minimum and the actual minimum is reflected in the difference between the combinatorial types of the original and the reconstruction. The polytope denoted “20 faces” is an entry showing the effects of considering just those faces which occurred in the reconstructed version, for which the errors are less.

Table 3.2 displays the combinatorial structure of the 21 faced polytope, as lists of face incidences. The absence of one face, with small area, from the reconstruction distorts the combinatorial structure. In fact, the only error in the combinatorial structure is the absence of that face; were it to be introduced into the adjacency lists where it would become tangent, the structures would be identical. Metric comparison of the original and the reconstruction shows that measures involving the support function H are more robust, and are appropriate measures for comparing

Face	Original	Estimated
1	(2 14 21 16 15 9 12 13)	(2 14 16 15 9 12 13)
2	(1 13 14)	(1 13 14)
3	(7 15 8 16 14)	(7 15 8 16 14)
4	(7 11 17 15)	(7 11 17 15)
5	(10 20 14 19 13 18)	(10 20 14 19 13 18)
6	(9 15 17)	(9 15 17)
7	(3 14 20 11 4 15)	(3 14 20 11 4 15)
8	(3 15 16)	(3 15 16)
9	(1 15 6 17 12)	(1 15 6 17 12)
10	(5 18 17 11 20)	(5 18 17 11 20)
11	(4 7 20 10 17)	(4 7 20 10 17)
12	(1 9 17 18 13)	(1 9 17 18 13)
13	(1 12 18 5 19 14 2)	(1 12 18 5 19 14 2)
14	(1 2 13 19 5 20 7 3 16 21)	(1 2 13 19 5 20 7 3 16)
15	(1 16 8 3 7 4 17 6 9)	(1 16 8 3 7 4 17 6 9)
16	(1 21 14 3 8 15)	(1 14 3 8 15)
17	(4 11 10 18 12 9 6 15)	(4 11 10 18 12 9 6 15)
18	(5 13 12 17 10)	(5 13 12 17 10)
19	(5 14 13)	(5 14 13)
20	(5 10 11 7 14)	(5 10 11 7 14)
21	(1 14 16)	nil

Table 3.2: Combinatorial Structures

polytopes with similarly oriented faces.

An advantage of this minimization formulation is its indifference to the adjacency relations in the polytope. A correct adjacency structure is guaranteed, in principle, by Minkowski's original argument. The difference in the second example between the combinatorial structures of the original O and the reconstructed polytope R is due to the implementation of the method, and is not due to the method itself. It is not an error so much as an omission. The missing face is very small in area, and any step to improve the objective function will only step in its direction a small amount. Very early in the minimization, the face is moved outside the polytope, at a distance 0.10 from its correct placement. Since the component of the step in the i^{th} coordinate is proportional to A_i , the corrections are small. At termination, the face has not moved in far enough. Faces can be eliminated and re-introduced by the minimization, as implemented. This face, however, was

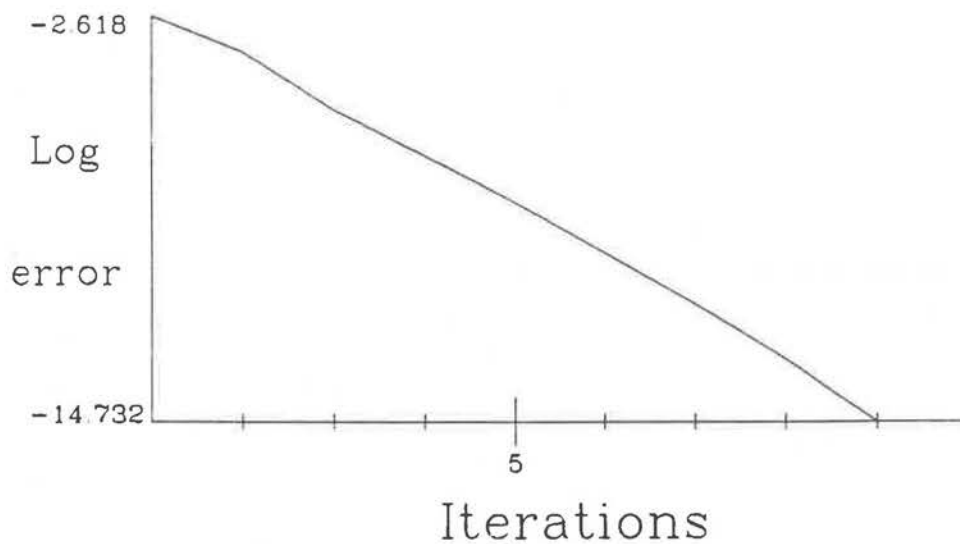


Figure 3.16: Log of Error vs. Iteration for Example 1

not re-introduced into the polytope. A straightforward modification to the implemented method brings each face which has zero area, in the estimate, inwards until it becomes tangent with the estimated polytope. Since all faces are known to exist in the reconstruction, this is sound. A naive method for finding tangency could take $O(n)$ operations.

Errors can arise from several sources: for example, incorrect calculation of the adjacency structure of the polytopes, leading to incorrect estimation of the volume and its gradient. This may slow down convergence, and may be introduced by rounding error during computation of any of these quantities.

Pictured in Figures 3.16 and 3.17 are graphs of the natural logarithm of the truncation error in the mixed volume versus the number of iterations. Both are approximately linear which is in agreement with the expected behaviour of the reduced gradient method. The ratio of successive errors (γ in Equation 3.9) is approximately 4.0 for the first example and 1.3 for the second, which is reduced to 1.02 by the later iterations. The larger number of iterations in the second case can be explained by the small angles between some of the randomly oriented faces causing large changes in A by small changes in H , so that in a sense the problem is ill-conditioned.

3.8 Reconstruction from Partial Information

It is possible to reconstruct the shape of the visible portion of an object, P , given the visible hemisphere of its EGI and the occluding contour, so that the faces in the reconstruction are identical to the visible portion of the object.

The complete EGI is a set of vectors $N = \{n_i | 1 \leq i \leq k\}$. Assume that the partial EGI consists of the first m of these vectors. The occluding contour may be specified as the contour

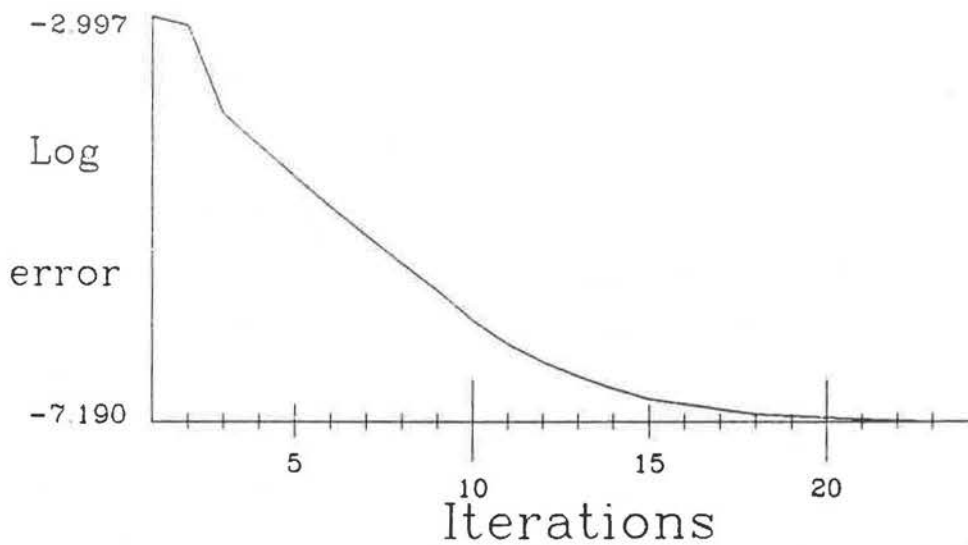


Figure 3.17: Log of Error vs. Iteration for Example 2

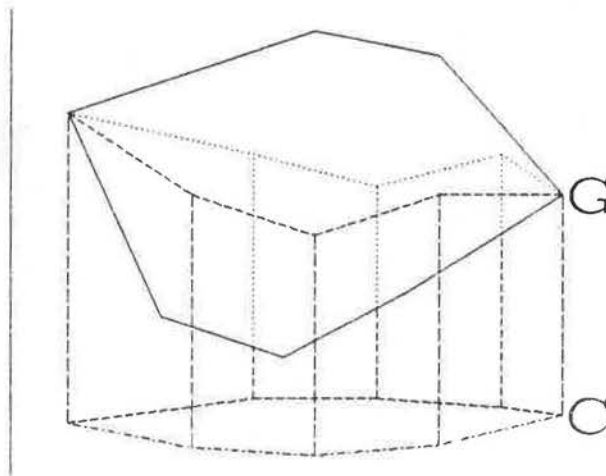


Figure 3.18: Visible Hemisphere of a Polytope, and the Contour Generator

generator, G , a path in R^3 , containing c vertices, or as the 2-d projection, C , of G .

The procedure for reconstruction will be described as if G itself is given; later, a method for generating G from C will be specified. To reconstruct shape from the visible hemisphere and the contour generator G , producing a new polytope D , proceed as follows (making reference to Figure 3.18):

1. Project G onto a plane normal to the viewing direction, incident upon the lowest vertex of G in that direction, producing C , as described above.
2. Connect each vertex in G with its projection in C . This produces a set of new faces whose orientations are normal to the viewing direction and the segment in C .
3. Compute the areas and normals of these faces, as well as the area of the face enclosed by C . Its normal is taken to be the opposite of the viewing direction.

4. Add the products of these areas and normals to the set of vectors comprising the partial EGI, giving a new EGI, N' with $m + c + 1$ elements.

This N' is equilibrated, since it is the EGI of an object which is convex, by the construction. Since it is equilibrated, by Minkowski's theorem, it uniquely corresponds to a polytope D . The portion of D corresponding to the partial EGI is shown to be the same as the polytope P which gave rise to N' by the following argument.

If the same sequence of operations is performed on P , namely, projection of G , giving C , and creation of the c faces connecting G with its projection C , a polytope Q is constructed. Its faces have the same orientation as those of P in the visible hemisphere and the vectors in its EGI correspond one-to-one with the vectors in the new EGI N' . By the uniqueness of reconstruction, this polytope Q is the polytope reconstructed from N' . By the construction, it has the same faces as P within G , hence the faces in D correspond to the faces of P within G .

Given only the projection C and the viewing direction z , proceed as follows. There are c edges in C , and m faces in the partial EGI. Consider assigning a face i to edge j ; this determines the slope of the line in R^3 since one component of the direction of the line is fixed by its projected direction - the second comes from the intersection of its assigned face (i) with a plane passing through the edge normal to the viewing direction z . The assignment of a face to an edge provides a direction in R^3 to the projected edge. Continue in this fashion assigning faces to edges. If each face is assigned to only one edge the number of possible assignments is $(m - c)!$; since a face may be assigned to more than one edge, the number is larger but is bounded above by m^c . The assignment is feasible if the path in R^3 generated by following the (un)projected edges closes. At least one assignment will generate a closed path, but there may be several which are feasible. For each feasible assignment, compute the reconstruction using the procedure described above. In this way a projected contour C can be transformed into a generator G , allowing the use of the previous reconstruction method. Its results are reliable when there is a unique back-projection (assignment of faces to edges), otherwise, the method can fail, in the sense that the reconstructed shape may not be unique.

Chapter 4

Support and Area Functions

4.1 Introduction

The algorithm described in the previous Chapter shows how the support vector for a polytope can be determined from its area vector. In order to facilitate arguments about mixed volumes in the succeeding Chapter, the continuous support function is described in terms of the area function for curves, permitting a qualitative comparison of the smoothness of the two functions. The next Section describes a method of curve segmentation based on extrema of the support function. Because the support function varies with translation, its behaviour under translation is studied, leading to proposals for several candidate centres of support. The study of these ideas suggests some interesting problems in computational geometry.

4.2 The Relation Between Support and Area Functions

For polytopes, the support function $\mathcal{H}(\omega)$ is a continuous function of ω , but the area function $\mathcal{A}(\omega)$ is not. To study the relation between these two functions of orientation one must move to the area of continuous descriptions; to describe their smoothness their Fourier descriptions are used.

4.2.1 Continuous Curves and Surfaces

A curve (a function $c : R \rightarrow R^2$) will be parametrized by arc length, i.e., the length of the tangent vector is unity [DoCarmo,1976]. A parametrized differentiable curve $c(s)$ is regular if and only if $c'(s)$ never vanishes. For any such curve the normal vector $n(s)$ is a unit vector perpendicular to

$c'(s)$, and

$$k(s)n(s) = c''(s)$$

$k(s)$ is the curvature. For the purposes of discussion in this section, orientation will be identified by the angle ω made by the normal vector with the x-axis. So, at any point on the curve,

$$c'(s) = dc/ds = (-\sin(\omega), \cos(\omega)) \quad (4.1)$$

Consider the curve parametrized by ω ; this is an invertible map if the curvature is strictly positive. Then

$$c(\omega) = \int_0^\omega c'(\psi)d\psi + \text{a constant} \quad (4.2)$$

by the fundamental theorem of calculus. By moving the point $c(0)$ to the origin, the constant of integration becomes 0. Now

$$dc/d\omega = \frac{dc/ds}{d\omega/ds}$$

At the point parametrized by s , whose orientation is ω ,

$$d\omega/ds = k(s)$$

where $k(s)$ is the curvature. Let

$$A(\omega) = 1/k(s) \quad (4.3)$$

by identifying the points on the curve. In R^2 the area function is the radius of curvature. Therefore, from Equation 4.1 and Equation 4.3,

$$c(\omega) = \int_0^\omega A(\psi)(-\sin(\psi), \cos(\psi))d\psi \quad (4.4)$$

The definition of the support function is

$$\mathcal{H}(\omega) = \langle c(\omega), (\cos(\omega), \sin(\omega)) \rangle \quad (4.5)$$

which can be expanded to

$$\mathcal{H}(\omega) = \cos(\omega) \int_0^\omega A(\psi)(-\sin(\psi))d\psi + \sin(\omega) \int_0^\omega A(\psi)\cos(\psi)d\psi \quad (4.6)$$

This formulation allows comparison of the Fourier spectrum [Oppenheim et al.,1963] of $A(\omega)$ and the Fourier spectrum of $\mathcal{H}(\omega)$ from which an evaluation of their relative smoothness can be made. In the following analysis, negative frequency components will be ignored and magnitudes will be considered. Let

$$a(f) = \widehat{A}(t)$$

indicate that a is the Fourier transform of A . Then

$$\widehat{A(t)\sin(t)} = a(f-1)$$

where t is in units of 2π and f represents frequency. Then

$$\int \widehat{A(t)\sin(t)} dt = \frac{a(f-1)}{if2\pi}$$

and

$$\cos(t) \int \widehat{A(t)\sin(t)} dt = \frac{a(f-2)}{i(f-1)2\pi}$$

Now the last equation represents the Fourier transform of the first half of the formula for \mathcal{H} in terms of A . The formula in terms of a , the Fourier transform of A , shows that \mathcal{H} is shifted to lower frequencies and its amplitude is reduced by the factor $i2\pi$, so that \mathcal{H} is in fact much smoother. A similar argument holds for the second component of the expression for \mathcal{H} (Equation 4.6). The relative smoothness of \mathcal{H} compared to A will later be used to strengthen, in a qualitative fashion, arguments for the superiority of attitude determination using mixed volumes (Section 5.2.6).

4.3 Curve Segmentation

Mackworth and Mokhtarian[1984] have extended the scale-space approach, first described by Stansfield in 1980, and then by Witkin in 1983, to curves in the plane, for describing the points on a curve as scale varies. This method depends on the existence of inflection points in the curve, i.e. points at which the curvature vanishes, hence on non-convexity in general. Weldon and Horn[1984] discuss smoothings of the extended circular image of curves in the plane.

Segmenting a convex curve can be based on the support function of the curve. Because the support function is not translation invariant it becomes important to find a suitable location in the plane for the curve.

4.3.1 Centre of Support

To study the behaviour of the support function under translation, first consider where \mathcal{H} reaches a local maximum or minimum. At an extremum,

$$\mathcal{H}'(\omega) = 0$$

$$\mathcal{H}(\omega) = \langle \omega, c(\omega) \rangle$$

$$\mathcal{H}'(\omega) = \langle \omega', c(\omega) \rangle + \langle \omega, c'(\omega) \rangle$$

But this last term is 0 because ω , the normal vector, is perpendicular to the tangent of the curve, by definition.

$$\mathcal{H}'(\omega) = 0 = \langle \omega', c(\omega) \rangle = k(\omega) \langle c'(\omega), c(\omega) \rangle$$

and since the curvature is strictly positive

$$0 = \langle c'(\omega), c(\omega) \rangle \tag{4.7}$$

Thus the tangent vector is orthogonal to the position vector, or the normal is collinear with the position vector.

The support function $\mathcal{H}(\omega)$, see Section 2.6.2, varies with translation. In all mixed volume calculations involving the complete figure, translation has no effect; apart from these considerations, what does translation mean? One can consider several candidate points as centres of support, a point which should be moved to the origin to get a suitable form of the support function.

Theorem 4.1:

The centre of the minimal circumscribed circle minimizes the maximum value of \mathcal{H} .

Proof:

From Equation 4.7, at a local maximum of \mathcal{H} , the position vector of the curve and the normal are collinear. The maximum occurs at a vertex having this property. The maximum of \mathcal{H} is then identical to the maximum of the distance function from the centre. So the quantity to be minimized is the same as that for the minimal circumscribed circle.

The minimal circumscribed circle is unique. If there were two, the figure would lie in their intersection. If one circle were entirely contained in the other, then it would be smaller, a contradiction. Hence their intersection forms a *lune*, the region formed by two partially overlapping circles. Its vertices are the endpoints of the chord shared by both circles. The length L of the segment connecting the vertices of the lune is less than the diameters of the circle, since the longest chord in a circle is the diameter. The circle with diameter L whose centre lies at the midpoint of the segment contains the lune, and thus the figure. Its diameter L is less than that of either circle, contradicting the assumption. Hence the minimal circumscribed circle is unique. ■

Theorem 4.2:

The centre of the maximal inscribed circle maximizes the minimum value of \mathcal{H} .

Proof:

The maximal inscribed circle will be tangent to the curve at no less than two points. There the normal of the curve and the circle coincide, so the position vector and the normal are collinear

and \mathcal{H} is a local minimum, by Equation 4.7. The global minimum of H occurs at these points also, since all other points on the curve are outside the circle. The global minimum is maximized since there is no larger inscribed circle, by definition.

The maximal inscribed circle is not unique. Consider a rectangle; the maximal inscribed circle has diameter equal to the shorter of the two sides of the rectangle and its centre can lie along a segment parallel to and halfway between the longer sides.

One can also consider determining points which minimize various integrals of derivatives of \mathcal{H} . First consider the integral:

$$\int_{\Omega} \mathcal{H}(\omega) d\omega \quad (4.8)$$

The domain of integration Ω is the unit circle. Recall the definition of \mathcal{H} : considering c as a function of ω , where ω is an orientation, i.e. a unit vector pointing in the direction of the normal to c , then

$$\mathcal{H}(\omega) = \langle \omega, c(\omega) \rangle \quad (4.9)$$

After any translation t , the above integral becomes

$$\int_{\Omega} \langle \omega, c(\omega) + t \rangle d\omega \quad (4.10)$$

which equals

$$\int_{\Omega} \mathcal{H}(\omega) d\omega + \int_{\Omega} \langle \omega, t \rangle d\omega \quad (4.11)$$

Now the second integral vanishes because the curve is closed, so the integral of \mathcal{H} over orientation is invariant under translation. However, the integral

$$\int_{\Omega} \langle \omega, c(\omega) + t \rangle^2 d\omega \quad (4.12)$$

does vary with translation; in fact, it increases without bound as $|t| \rightarrow \infty$. It has a minimum, since it is bounded below by zero, and is a monotonically increasing function of t . The minimum is achieved when

$$t = \left(\frac{1}{\pi} \int_{\Omega} \mathcal{H}(\omega) \cos(\omega) d\omega, \frac{1}{\pi} \int_{\Omega} \mathcal{H}(\omega) \sin(\omega) d\omega \right)$$

The integral of the derivative of the support function:

$$\int_{\Omega} \langle \omega, c(\omega) + t \rangle' d\omega \quad (4.13)$$

is invariant under translation, since the integral evaluates to:

$$\langle \omega, c(\omega) + t \rangle \Big|_0^{2\pi} = 0$$

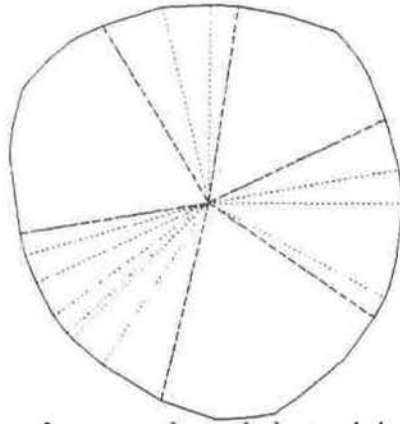


Figure 4.1: Lobes of a curve, bounded at minima of the support function

The integral of the derivative squared

$$\int_{\Omega} ((\omega, c(\omega) + t)')^2 d\omega$$

varies with translation, and is minimized at:

$$t = \left(\frac{1}{\pi} \int_{\Omega} \mathcal{H}'(\omega)(-\sin)(\omega) d\omega, \frac{1}{\pi} \int_{\Omega} \mathcal{H}'(\omega)\cos(\omega) d\omega \right)$$

4.3.2 Segmenting Curves at Extrema of \mathcal{H}

For a polygonal curve, curvature reaches a local maximum at every vertex, so curvature segmentation is at best confusing. Instead local maxima of the support function are examined. Let a region between minima of \mathcal{H} be termed a *lobe*; the size of the lobe is the angular difference between the normals at the enclosing minima. In Figure 4.1, the boundaries of large lobes of the convex curve are bounded by longdashed lines from the centre, and small lobes by dotted lines.

To describe a curve in terms of \mathcal{H} , the best centre must be chosen. The number of extrema of \mathcal{H} depends on the location of the curve in the plane; by translating the curve away from the origin, the number of extrema can be reduced to 2, a maximum in the direction parallel to the translation vector and a minimum in the opposite direction. It becomes interesting to consider where the number of extrema of \mathcal{H} is maximized. There, in a sense, the curve is maximally segmented. A point on a curve is an extremum of \mathcal{H} if condition 4.7 is met, so that the normal of $c(\omega)$ is collinear with the position vector. Minima of \mathcal{H} on a polygonal curve occur on edges. There will a minimum of \mathcal{H} along an edge if the origin lies in a region of the plane bounded by two lines orthogonal to the edge, passing through the endpoints of the edge. Such a region can be called a *slab*. The overlay of all slabs originating on the edges of a polygonal curve subdivides the interior of the closed curve into a set of regions, each of which lies within a varying number of slabs, see Figure 4.2. So to maximize the number of extrema of \mathcal{H} on a polygonal figure, move

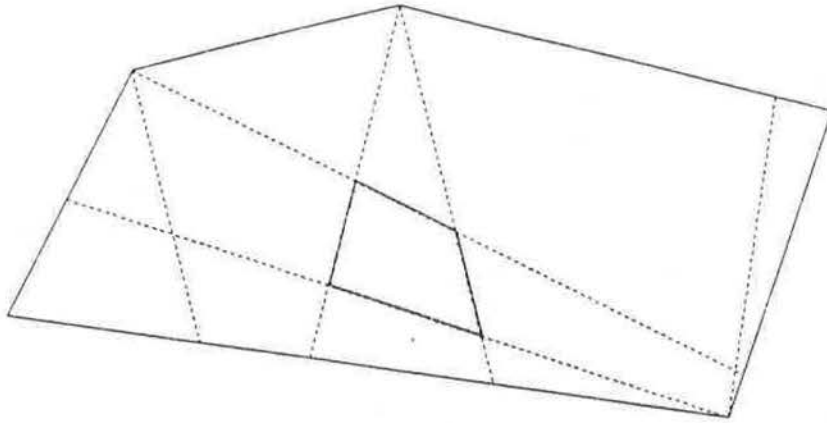


Figure 4.2: Slabs on a polygonal curve; region of maximum overlap is shaded

the curve so that the origin lies in the overlap region which sees the most edges. Since there are $2n$ segments which may each intersect $n/2$ other segments, the intersection of these slabs can take $O(n^2)$ time. Each region generated can be annotated with the number of edges it can see, and the maximum region chosen. Of course, there may be several regions with the same value, so the choice is not unique, as with the maximal inscribed circle. In fact, any point in the region with highest cardinality will produce the same number of extrema.

Chapter 5

Determining Object Attitude from the EGI

5.1 Introduction

The representation of objects by their EGIs leads to a method for determining object attitude. Horn and Ikeuchi[1984] have demonstrated the feasibility of using EGIs for attitude determination. The EGI functions as an area histogram for the object; the sensed portion of an EGI is compared directly with the EGI of a prototype.

The reconstruction method developed in chapter 3 provides a new method for attitude determination, based on the mixed volume. It is shown that this method is both practical and more robust than direct comparison of EGIs. Experiments in determining attitude with a variety of polytopes are presented.

5.2 Determining Attitude

Determining the attitude from the EGI of a known object is equivalent to finding a rotation R , such that

$$R(EGI_{prototype}) = EGI_{sensed} \quad (5.1)$$

Attitude in R^3 can be identified with a rotation $R(\theta, n)$, where θ is the angle by which the object is rotated around the axis n . Alternatively, attitude can be described by a quaternion [Salamin,1979] [Brou,1983], (q_0, q_1, q_2, q_3) , which may be divided into a real part q_0 and a vector part, (q_1, q_2, q_3) . Multiplication, addition and scalar multiplication are defined for quaternions. A pair of unit quaternions, $q(\theta, n), q(-\theta, n)$ are associated with a rotation. Rotation by q is given

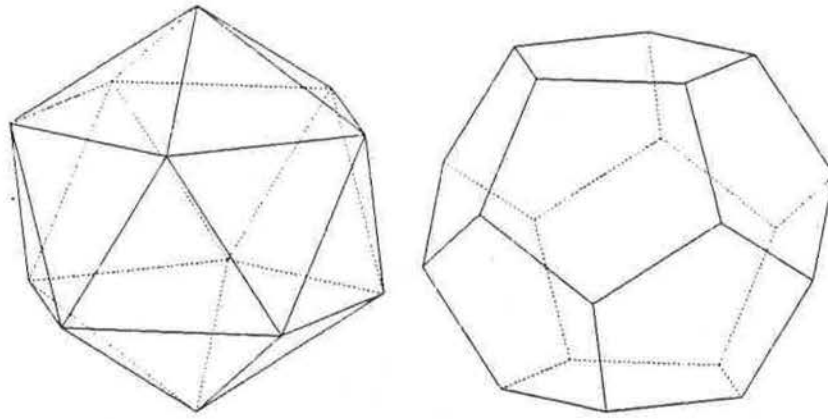


Figure 5.1: Icosahedron and Dodecahedron

by:

$$v' = qvq^{-1}$$

where

$$q^{-1} = \frac{q^*}{|q|}, \text{ and the conjugate of } q, q^* = (q_0, -q_1, -q_2, -q_3)$$

It can be shown that multiplying quaternions is equivalent to composition of rotations. For a detailed discussion of methods for representing and quantizing attitude in R^3 see Brou[1983].

5.2.1 Quantizing Orientation

In order to access an object's area function in a general fashion it is necessary to quantize orientation — U must be tessellated. Brou[1983] has discussed in detail the criteria for tessellating U . Initially, a tessellation of U can be created by projecting a regular polyhedron onto U . The regular polyhedron with the most faces is the icosahedron with 20. The number of facets in a tessellation of U can be increased by subdividing the triangular faces of the icosahedron into smaller triangles. The quantization on U used here is based on subdivisions of the icosahedron, following Brou's use of that polyhedron. Horn and Ikeuchi [1984] use subdivisions of the dodecahedron, to achieve better distribution of angles in the cells. The results presented here should not be affected by the difference in tessellations. Figure 5.1 shows the icosahedron and the dodecahedron used to produce the tessellations. At frequency 1 the faces of the icosahedron are mapped into 1 triangle, and at frequency i , each face is mapped into i^2 triangles. Figure 5.2 shows the frequency 2 and frequency 3 subdivisions of U generated from projection of the icosahedron onto the sphere.

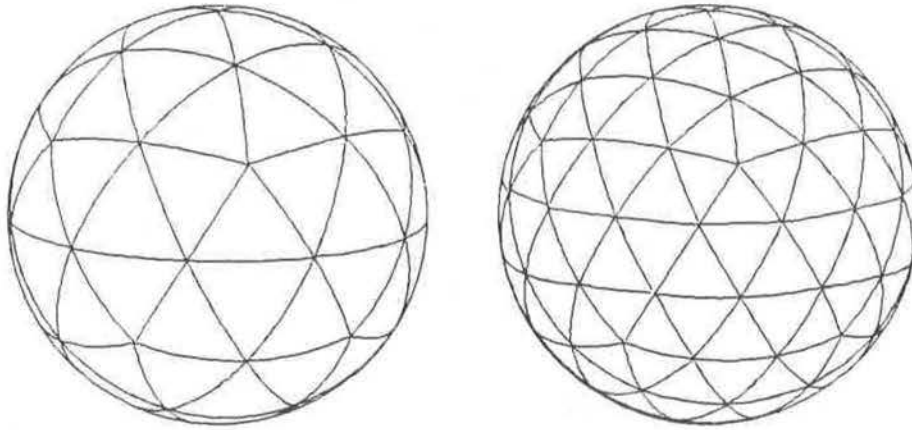


Figure 5.2: Icosahedra Subdivided at Frequency 2 and Frequency 3

5.2.2 Quantizing Attitude

Attitude is quantized using the rotation group Γ_{60} which brings the vertices of the icosahedron into correspondence with each other. One can envision this rotation group by imagining taking each of the 12 vertices of the icosahedron successively into correspondence with one particular vertex. Each vertex on the icosahedron has 5 incident edges. The rotations about the vertex generate 5 attitudes for each vertex mapped to the selected vertex, giving 60 rotations altogether.

The rotations which constitute this group are those at which experiments are done in this chapter. The difference between two attitudes is the angle of the rotation (axis and angle) which takes one into the other. The minimum difference between attitudes in Γ_{60} is 72° .

5.2.3 Attitude Determination

From a suitable closed-form description of a prototype and its sensed area function, it may be possible to determine its attitude analytically. In practice, the description of an object does not have such a succinct representation. An EGI quantizes orientation and records the object area accumulated in each cell on U . First, it is useful to consider the problem of attitude determination without the effects of quantization.

Determining attitude from EGIs can be solved by correlation. A rotation can be considered an offset; for a curve, the EGI represented as a function on the range $0 - 2\pi$ is shifted until it matches itself. A procedure for determining attitude finds an offset (the independent variable in correlation) which is appropriate. For a symmetric object, there will be more than one rotation bringing it into correspondence with itself, so there will be several equally good attitudes. Where the object is not symmetric the uniqueness of the EGI guarantees there will be only one correct

attitude. In the presence of noise both of these assertions are untrue. The solution which maximizes the correlation can be chosen. An exact match may not be found. Variability of the solution with noise is dependent entirely on the object shape. For the EGI, this can be expressed by the autocorrelation function. An autocorrelation has its maximum M at 0 by definition. When the autocorrelation has several other maxima near M in value, it is possible that noise in the data will bring the correlation of the prototype EGI and the sensed EGI above M at incorrect offsets. Again this is dependent on the autocorrelation. Attitude determination is unstable in the presence of noise.

Later in this chapter arguments will be advanced to show why attitude determination with the mixed volume is more effective (stable) in practice than direct correlation. These arguments will depend on an analysis of area and support functions. Both methods select optima of the respective functions. In the presence of noise, both can fail. It will be argued that the mixed volume method is more stable. Finally, experiments with both methods will be examined to support the claims.

5.2.4 Attitude from Graph Matching

Seidel[personal communication,1984] has suggested a procedure for matching complete EGIs in an error-free environment. An EGI under this interpretation is a set of measurements of area at a fixed set of orientations. The procedure is discussed in Appendix B. Attitude can be determined by this procedure in $O(n \log n)$ operations.

Where only some of the orientations are present in the sensed EGI, it is necessary to match a graph derived from the sensed EGI with a subgraph of the prototype graph. Determining whether a particular graph is isomorphic to a subgraph of a graph is NP-complete, in general. Where both graphs are three-connected, with a unique planar embedding, subgraph isomorphism has a polynomial-time solution. A method similar to that of Sugihara [1984] can be used. What prevents the use of a graph-matching method in practice is that sensing error in measured area will alter the order of the sorted area values, invalidating the use of a discrete algorithm based on exact ordering.

5.2.5 Attitude by Comparing Area Functions

To determine the attitude of an object from its EGI, Horn and Ikeuchi [1984] match the prototype EGI and the sensed EGI at a discrete set of attitudes. At each sample attitude a measure M of

matching between the two area distributions on U is computed.

$$\sum_{i=1}^m M(A_i, A_{R(i)}) \quad (5.2)$$

where A_i is the sensed EGI area function evaluated in direction ω_i and $A_{R(i)}$ is the prototype area function evaluated at the direction to which ω_i is transformed by R , and the matching function M is

$$M(A_{R(i)}, A_i) = (A_{R(i)} - A_i)^2 \quad (5.3)$$

This technique and matching function work well when the attitude of the sensed object is close to one of the sample attitudes. Brou [1983,p. 113] admits that "high mismatch values are obtained if the object's orientation is not in the discrete set". The area matching method, at frequency 1, "leads to relatively low values of M_1 even if the object's orientation is slightly different from the ones in the set". M_1 is the match value at frequency 1. "Unfortunately, large M_5 's", the value at frequency 5, "are obtained when the object's orientation is slightly different from the ones in the set." This leads to the suggestion that 5880 different attitudes should be used to compare frequency 5 tessellations.

With increasing resolution the effectiveness of area matching decreases. As the resolution of orientation increases, the number of empty cells increases. A polytope with m faces will always fill at most m cells on U . If the difference in attitude between the prototype and the sensed object is near the cell resolution, the faces of the prototype and the object will not lie in the same cells, even at the correct attitude. Cells with values will be compared with empty cells. The maximum matching error increases and the minimum value will not necessarily indicate the correct attitude.

5.2.6 Determining Attitude with Mixed Volumes

Reconstruction selects a polytope whose area function fits the EGI; in attitude determination an attitude is sought which rotates the sensed EGI into correspondence with the prototype EGI. Both seek an appropriate area function; for attitude determination the choice is restricted to rotated versions of the prototype. In Section 3.4.4 it is shown that, of all polytopes P with fixed volume, the polytope whose support function H minimizes the mixed volume $\langle H_P, A_Q \rangle$ is homothetic to the polytope with the given area function A_Q . The relevant equations for 3-polytopes are described in Section 3.4.2:

$$(1/3 \sum_{\Omega} H_P A_Q)^3 = V_3(P, Q)^3 \geq V(P)V(Q)^2 \quad (5.4)$$

P and Q are the prototype polytope and the completed sensed polytope. This equation leads to

$$(1/3 \sum_{\Omega} H_P A_Q)^3 / V(P) \geq V(Q)^2 \quad (5.5)$$

The mixed volume cubed divided by the volume of P is minimized when P is homothetic to Q .

Rotating a polytope preserves volume, so among all P' which are rotated versions of P that P' which has the same attitude as P minimizes the mixed volume. So, to determine object attitude, minimize:

$$\sum_{i=1}^m (H_{R(i)} * A_i)$$

H_i is the prototype support function, A_i the sensed area function, and $R(i)$ is the position at which i occurs in rotation R .

For polytopes, the area function $A(\omega)$ is discontinuous; in contrast, the support function $\mathcal{H}(\omega)$ varies smoothly on U . In fact, the area function of a polygon, for example, consists of a finite set of non-zero points in the interval $0 - 2\pi$. Its autocorrelation is 0 almost-everywhere. Not surprisingly, when the attitude of the sensed object is slightly different from the prototype (or any attitude in the sample set), the value of the area matching is 0. Discretizing A by sampling U is helpful, as Brou remarks, in introducing a smoothing effect, widening each of the pulses to the resolution. With discretization, small changes in attitude do not affect the area function, until the size of the attitude difference exceeds the resolution on U , when the errors recur.

The support function is an integral transform of A and is much smoother. A more complete discussion is in Section 4.2.1. Because A varies rapidly, the prototype, under a small rotation, is very different in A from the actual attitude. Mismatches then occur. The support function varies less and the error rate is correspondingly lower. For a polygonal curve, the support function is piecewise sinusoidal, (Equation 2.5). The mixed volume of a polygonal curve and itself is also piecewise sinusoidal, since it is the convolution of a set of impulses of height $A(\omega_i)$ with the support function $\mathcal{H}(\omega)$ which is piecewise sinusoidal.

5.3 Experiments in Determining Attitude for Complete EGIs

There are two variables in these experiments: first, the resolution at which orientation is quantized, and, second, the difference between the attitude of the prototype and that of the test object. For each prototype, several experiments were performed. First, the attitude is held constant (usually at 0°), and resolution varies. Second, at a specified resolution, the difference between the attitude of the object and the prototype changes.

The first series of experiments uses several different polytopes and compares the effectiveness of the area and mixed volume methods at determining attitude from a complete EGI, at varying resolutions on U . Both area matching and mixed volume methods succeeded. The theory of mixed volumes predicts that minimizing mixed volume will perform correctly when given the exact support function and the exact areas. Quantization effects can introduce error. That the increase in resolution of models and sensed values produces correct behaviour agrees with the theory.

The results in the following tables describe two values: first, the number of degrees between the attitude selected by the matching method (α_A or α_{MV}) and the correct attitude (α_0), and, second, the rank at which the correct attitude is placed by the method ($Rank(\alpha_0)$). In both methods, the values of matching at all test attitudes are computed and sorted. The minima identify an attitude, which is ranked 0. The correct attitude is the nearest attitude in Γ_{60} to the actual attitude of the prototype. When $Rank(\alpha_0)$ is not 0, $r \neq 0$, the correct attitude occurs in the $r + 1^{th}$ position in the sorted list. The higher $Rank(\alpha_0)$ is, the more poorly attitude has been determined. Each box is subdivided vertically into the measurements for area comparison and mixed volume. Each element in the tables has the form described in Figure 5.3.

The experiments with varying resolution on U , complete EGIs, used 4 different polytopes. The first, with 21 faces, is in Figure C.1, the second, with 40 faces, in Figure 5.5, the third with 20 faces, in Figure C.3 and the last in Figure C.4. The results of these experiments are shown in Table 5.1. These experiments demonstrate that the mixed volume method at frequency level 2 performs as well as area comparison. At resolution level 1 the coarseness of the quantization prevents the mixed volume method from working on some objects.

The second form of experiment varies the difference in attitude between the prototype and the test object. The results of that experiment, using the complex polytope (having 21 faces) in Figure C.1, are shown in Table 5.2. The area matching method fails when the angle between the attitude of the prototype and the test object is as little as 15.0° . Even more serious failure occurs in this set of rotations (Table 5.3). Also, the mixed volume method fails in this example,

Area	Mixed Volume
$\alpha_A - \alpha_0$	$\alpha_{MV} - \alpha_0$
$Rank(\alpha_0)$	$Rank(\alpha_0)$

Figure 5.3: Form of entries in the comparison tables

Object	Frequency 1		Frequency 2	
	0°	144°	0°	0°
21 faces at random distances	0	5	0	0
40 faces on sphere	0°	180°	0°	0°
	0	6	0	0
20 faces on ellipsoid	0°	0°	0°	0°
	0	0	0	0
80 faces ellipsoid	0°	0°	0°	0°
	0	0	0	0

Table 5.1: Errors with varying resolution of orientation

Axis	Frequency 2							
	5°		10°		15°		20°	
1,0,0	0°	0°	0°	0°	144°	0°	72°	0°
	0	0	0	0	2	0	4	0

Table 5.2: Errors at frequency 2 with varying attitude, axis (1,0,0)

Axis	Frequency 2							
	5°		10°		15°		20°	
0,1,0	0°	0°	0°	0°	72°	0°	120°	72°
	0	0	0	0	6	0	21	2

Table 5.3: Errors at frequency 2 with varying attitude, axis (0,1,0)

Frequency 3								
Axis	5°		10°		15°		20°	
1,0,0	0°	0°	72°	0°	144°	0°	72°	0°
	0	0	9	0	8	0	17	0
0,1,0	0°	0°	0°	0°	0°	0°	72°	180°
	0	0	0	0	0	0	11	2
0,0,1	0°	0°	0°	0°	72°	0°	72°	144°
	0	0	0	0	21	0	23	1

Table 5.4: Errors at frequency 3 with varying attitude

but its failure is less severe in two respects: the incorrect attitude selected is closer to the correct attitude and the number of incorrect attitudes lower in rank is smaller. Both methods are correct for all rotations about axis (0, 0, 1).

Table 5.4 shows the results for all rotations at frequency 3. The results for the 20° rotation are shown in another form in Figure 5.4. The graph shows the values of area matching, in a dotted line, and mixed volume, in a solid line, on the vertical; the horizontal axis shows the attitudes (from Γ_{60}) at which the matchings were evaluated. The horizontal axis is separated by dashed lines into regions which are 0, 72, 120, 144 and 180° from the prototype attitude. In this figure, point A is the minimum of the mixed volume method, which is also the global minimum and correctly identifies the object attitude. B, the value of the area matching method for the correct attitude, is rather high. There are 21 attitudes with lower values. The area function achieves a minimum (point C) at an attitude in the 72° set, while the mixed volume goes to a minimum at 0°. The minimum difference between rotations in Γ_{60} is 72°.

In general, the area matching method performs worse at a finer resolution: out of 12 cases, it is worse in 5, better in 2 and the same in 5. The mixed volume is worse in 2 cases, better in 1 and the same in 9. The MV method is in error in only 2 out of 12 cases, while the area matching method fails for 6 out of 12 cases, at frequency 3.

Finally, when the two methods are compared at frequency 5 subdivisions on U , the number of errors increases, as expected, from 6 out of 12 for area matching to 8, while the mixed volume method fails in two cases (see Table 5.5).

The behaviour of the two methods under changing resolution of the tessellation on U agrees with the analysis; at frequency 3 the magnitude of the rotation is approximately the same as the size of the cells of U (15° vs. a range of 6 to 14° on U) [Brou,1983,p. 116]. There the smoothing

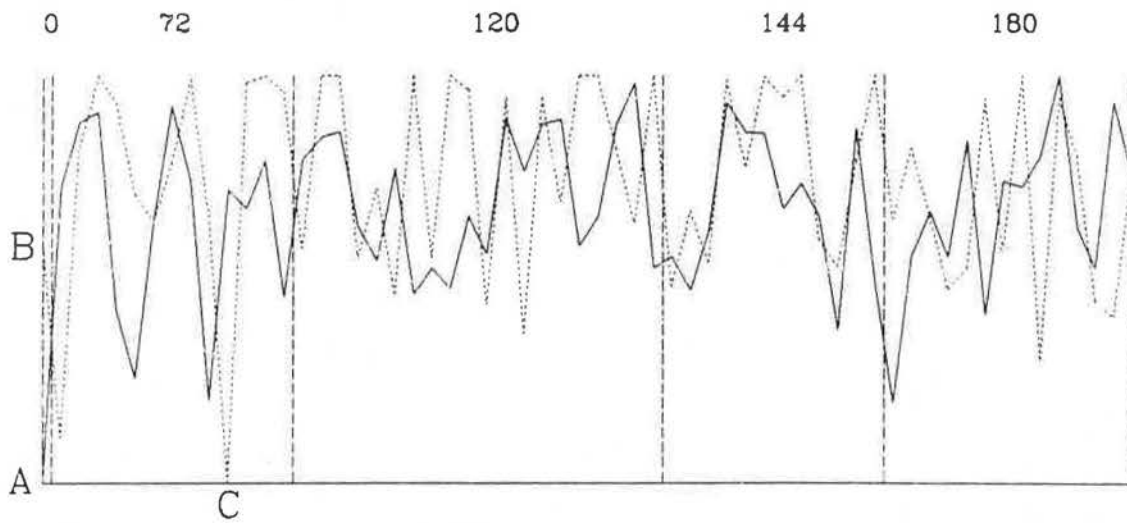


Figure 5.4: Mixed volume (solid) vs. area (dotted), 21 faces, $f=3$, 20 deg.(1,0,0)

Frequency 5								
Axis	5°		10°		15°		20°	
1,0,0	0°	0°	180°	0°	144°	0°	72°	180°
	0	0	1	0	20	0	16	1
0,1,0	0°	0°	0°	0°	120°	0°	120°	0°
	0	0	0	0	5	0	21	0
0,0,1	0°	0°	120°	0°	72°	0°	72°	144°
	0	0	17	0	26	0	17	1

Table 5.5: Errors at frequency 5 with varying attitude

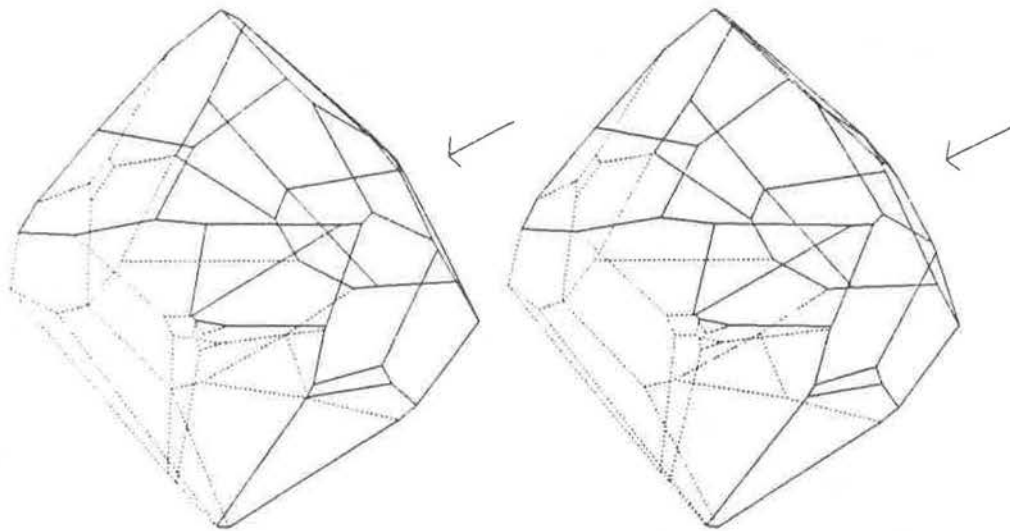


Figure 5.5: Polytope (40 faces) visible edges solid, invisible dotted

effects on A at lower resolution are significantly reduced. At frequency 5 on U the errors are larger because the cells on U are even smaller (a range of 4 to 9° on U). The mixed volume method does err in several cases for rotations of 20°, but since the rank of the correct attitude is second or third, the error can be attributed to a quantization effect on H . The analysis supports these results; since H is smoother, the effects of the difference in attitude from the prototype are significantly smaller for the mixed volume method. It should be possible then to trade off resolution on U against the number of test attitudes, so that a significant improvement in the accuracy of the attitude determination could be achieved.

5.3.1 Using Visible Hemispheres

The previous section demonstrates the mixed volume method can be effective in determining attitudes from complete EGIs. In practice, without fortuitously situated mirrors or multiple cameras, only the portion of a surface corresponding to a single visible hemisphere of the EGI is sensed. In Figure 5.5, the polytope used in the example (with 40 faces on sphere) is shown in stereo; a vector is drawn showing the view vector, and edges not visible from that view direction are dotted. Can the MV method be effective with only partial information? The MV method will only utilize the support function and areas of visible faces, as will the area method. In the following tables the MV technique is contrasted with the area method for several objects. In these example the area matching is:

$$\sum_{\Omega_v} (A_i, A_{R(i)})^2$$

8 face polytope at resolution level 3								
Axis	5°		10°		15°		20°	
1,0,0	0°	0°	0°	0°	180°	0°	120°	120°
	0	0	0	0	4	0	19	5
0,1,0	0°	0°	0°	0°	72°	0°	120°	72°
	0	0	0	0	14	0	16	3
0,0,1	0°	0°	72°	0°	72°	0°	120°	0°
	0	0	5	0	5	0	7	0

Table 5.6: Errors with subsets at frequency 3 with varying attitude (8 faces)

40 face polytope at frequency 3								
Axis	5°		10°		15°		20°	
1,0,0	0°	0°	0°	0°	180°	0°	144°	0°
	0	0	0	0	1	0	2	0
0,1,0	0°	0°	0°	0°	180°	0°	72°	0°
	0	0	0	0	1	0	3	0
0,0,1	0°	0°	0°	0°	72°	0°	72°	72°
	0	0	0	0	2	0	17	4

Table 5.7: Errors with subsets at frequency 3 with varying attitude (40 faces)

and the mixed volume matching is:

$$\sum_{\Omega_v} A_i H_{R(i)}$$

where Ω_v is the set of visible cells on Ω at the given viewing orientation; only a visible hemisphere affects the results.

These results show the MV method performing as well as and better than area methods for visible hemispheres.

The MV method integrates around the entire object the product of the support function and the area function. Using MV for attitude determination depends on the fact that volume remains constant (see Equation 5.5). The volume of the visible hemisphere of an object, however, does not remain constant. Thus, the volume of the subset of the prototype being compared must be estimated for scaling the mixed volume. Also, the area of the invisible face(s) of the polytope must be estimated as well as the support function for the direction(s) away from the viewing direction.

21 face polytope at frequency 3								
Axis	5°		10°		15°		20°	
1,0,0	0°	0°	0°	0°	0°	0°	180°	0°
	0	0	0	0	0	0	10	0
0,1,0	0°	0°	0°	0°	0°	0°	0°	0°
	0	0	0	0	0	0	0	0
0,0,1	0°	0°	72°	120°	72°	120°	72°	120°
	0	0	5	1	7	1	21	2

Table 5.8: Errors with subsets at frequency 3 with varying attitude (21 faces)

80 face polytope at frequency 3								
Axis	5°		10°		15°		20°	
1,0,0	0°	0°	0°	0°	180°	0°	144°	0°
	0	0	0	0	2	0	1	0
0,1,0	0°	0°	0°	0°	0°	0°	72°	0°
	0	0	0	0	0	0	1	0
0,0,1	0°	0°	0°	0°	72°	0°	72°	72°
	0	0	0	0	1	0	6	1

Table 5.9: Errors with subsets at frequency 3 with varying attitude (80 faces)

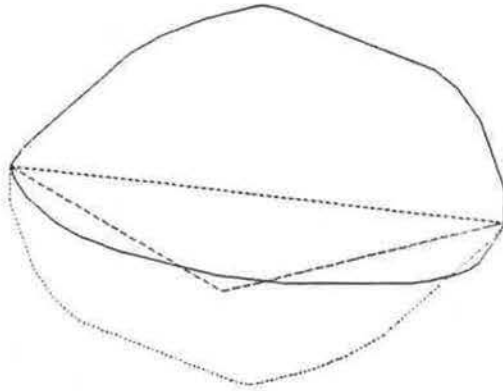


Figure 5.6: Polygon Completion: shortdashed=direct, dashed=spread,dotted=reflection

There are many methods of completing a subset of the faces of a polytope. The orientations and areas for the introduced faces must cause the polytope to close, i.e., if Ω_v are the visible face orientations and Ω_i the invisible orientations, then

$$\sum_{\Omega_v} A\omega = - \sum_{\Omega_i} A\omega \quad (5.6)$$

There is some freedom in selecting appropriate orientations for the introduced face(s). Early experiments were done on polygonal figures with fixed orientations of the faces, so that completion of the invisible portion naturally meant selecting some of the invisible orientations for faces. The area of the invisible portion must satisfy Equation 5.6, so that the solution in 2-d is determined for two orientations; for any more the system is underdetermined and some optimality criterion must be used. The spread completion, as this is called, for a figure is shown in dashed lines in the Figure 5.6. If the orientation of the added face need not arise from a specific set then the direct completion (in shortdashed lines) can be used. Finally, a simple solution in 2-d is to reflect the figure about the line connecting the endpoints of the visible portion, so that areas and orientations from the visible portion can be used. All of these methods work equally well in 2-d, and because a completed polygon is used, the support, and area functions are complete for the entire circle and volume is computed, so that the mixed volume can be appropriately scaled.

The discussion above is situated in R^2 for two reasons: clarity, and because completion is simple in R^2 , retaining the same two points at the boundary of the visible and invisible portion of the polytope. The reconstruction method for partial information described in Chapter 3 will complete a polytope from a visible hemisphere and a bounding 3-d contour. The area of the face appended to the visible portion is equal to the sum of the projected areas of the visible faces, so that is easy to acquire, but the support function of that face is dependent on the coordinates of the boundary contour. In general, the 3-d shape of the bounding contour is not known, and is

difficult to determine, without performing reconstruction. In 2-d the shape and support function can be easily be reconstructed. In 3-d the three completion methods can be implemented. Spread completion requires 3 normals; direct completion uses the closure condition in Equation 5.6 to define a new face, and reflection can be accomplished by using the opposite of the normals for all visible faces. Thus, the reflection method generates a symmetric polytope. None of these methods will necessarily generate the same bounding contour as the real polytope; however, the areas of the visible surfaces will correspond with the real polytope.

The justification for the mixed volume method depends on area functions and support functions over the entire Gaussian Sphere, but it does work well for subsets in practice. If the area and support for the invisible face(s) of the completed polytope were available, the product of the estimated area of the completing face for the sensed EGI would be multiplied by the support value of the prototype in that direction (with completion). This back face lies near the centroid of the object (the origin), so its support value is small. Thus the contribution of the invisible face would be negligible. What about estimating the volume of the visible half of the prototype? As mentioned before, the mixed volume to be minimized is:

$$(1/3 \sum_{\Omega} H_P A_Q)^3 / V(P) \geq V(Q)^2$$

so the quantity to be minimized is:

$$(1/3 \sum_{\Omega} H_P A_Q) / V(P)^{\frac{1}{3}}$$

The cube root of the volume should vary linearly with the support function so this factor is minimal.

Returning to completion methods, the least expensive would be direct completion since for a prototype only one additional number is stored at each cell on U , indicating the support function value of the invisible face when that point is closest to the view direction. Thus storage doubles. For spread completion three values are required. Symmetric reflection can be accomplished using the reconstructed polytope - the polytope is symmetric along any line through the centre of gravity, so its support function in any direction ω is the same as in the opposite direction.

5.4 Non-Convexity and Attitude

Brou[1983] conjectured that one can compare the area function of a non-convex body with the area function of its convex hull. This approach has many difficulties, not the least of which may be that the convex hull must be calculated for each image. One might conjecture that the support

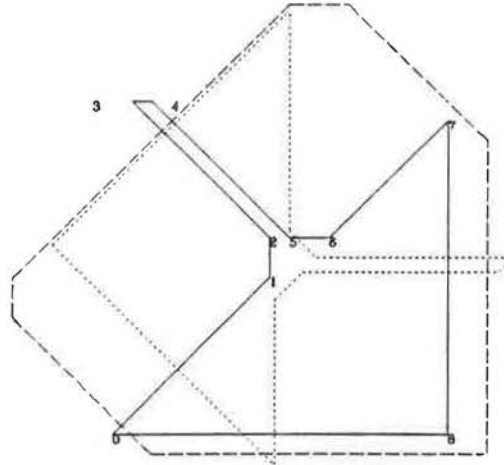


Figure 5.7: Non-convex object

function of the convex hull should be used and the mixed volume method applied. The convex hull of the prototype need only be computed once and stored. This has been tried in R^2 , using the complete EGI, for several objects, and is ineffective. In Figure 5.7 such a polygon is shown. The solid lines describe the object, the dotted lines delimit the object in the position at which the mixed volume (using the convex hull) is minimized.

Dashed lines in this figure delimit the convex polygon reconstructed from the area function of this non-convex object. A mixed volume method can use the support function of this reconstructed object and the area function of the sensed object. Using complete EGIs and ignoring the effects of self-occlusion, this mixed volume method correctly determines attitude for a set of sample polygons for which using the convex hull fails. Lyusternik[1963] mentions that the Brunn-Minkowski inequality (volume of mixture is greater than or equal to mixture of volumes) holds even for non-convex figures. This holds the promise that the mixed volume method can be extended to even non-convex objects, for attitude determination, especially where the effects of self-occlusion are small.

Chapter 6

Non-convex Objects

6.1 Introduction

Mathematical results about support functions, area functions, and EGIs have mainly concerned convex bodies. In practice, these form only a proper subset of the objects to which EGIs have been applied for recognition and attitude determination. The method for reconstructing shape from an EGI has only been described, in this thesis, for convex bodies. This chapter collects known results on the behaviour of the Gauss map for non-convex bodies, and investigates several of the possible ways in which EGIs might be applied to non-convex bodies. It is shown how simple generalizations of EGIs to non-convex bodies fail to allow shape reconstruction. Statements about shape reconstruction for non-convex bodies depend, in general, on the existence of (unique) solutions of systems of partial differential equations, a question in differential geometry which is beyond the scope of this thesis. Instead, an analysis of the requirements of a general solution is provided.

EGIs have been applied to reconstruction, attitude determination and recognition. Even cursory examination of these applications indicates that some form of segmentation of the surfaces will be necessary. For any map to be invertible, it must be one-one, and, so, for reconstruction from an EGI, which is an inverse problem, the Gauss map must be one-one. On a non-convex body, at least one point on the Gaussian sphere will have a multiple inverse image under the Gauss map. The surfaces of non-convex bodies must be subdivided to allow reconstruction. Surfaces of revolution form a class of objects in which segmentation by the sign of the Gaussian curvature can ensure a one-one Gauss map. Although estimating derivatives is not numerically well-behaved, such segmentation can be done at a coarse scale.

It must be mentioned at the start that other object representation schemes, such as polyhedral

models, or decomposition into generic primitives, (as described in Chapter 2) do not have such inherent difficulty in handling non-convex objects. It is nevertheless useful to ask the question "How hard is it to apply the EGI to non-convex objects?"

6.2 EGIs for Non-Convex Bodies

In practice, an EGI is computed from an orientation map by adding the area of the surface subtended by a pixel to the appropriate sum for the orientation of the pixel. At each pixel the subtended area is calculated by dividing the area of the pixel on the image by the cosine of the angle between the viewing direction and the surface orientation. As mentioned in Chapter 2, curvature is inversely related to area. Horn[1983] notes that this method of computing the EGI is effectively adding the absolute value of the inverse of curvature, ignoring possible negative curvature regions. This allows an extension of EGIs to non-convex objects, which Ikeuchi et al.[1983] have utilized in a system for manipulating objects.

Under this definition of EGIs for non-convex objects, the Gauss map is no longer an injection. Multiple points have the same orientation on a non-convex object, and all are similarly mapped. Horn provides an illustrative example. The EGI of a torus, produced by summing $|1/K|$ for all points, is equivalent to the EGI of a convex body whose silhouette is the curve of least energy [Horn,1982]. This interpretation of the EGI cannot discriminate between these two objects and is ineffective for recognition, except where, as in an industrial environment, conditions can be controlled so that no two objects with similar EGIs occur together. This example, the torus, shows that this variant of an EGI is not invertible. The use of this form of the EGI in attitude determination [Ikeuchi et al., 1983] shows its value in practice.

Ikeuchi[1983] has suggested using multiple EGIs of a non-convex object, where a separate EGI is maintained for each possible attitude, incurring an enormous penalty in storage.

6.2.1 Curvature and the Gauss Map

All previous uses of the EGI described in this dissertation have depended on the fact that the Gauss map is one-one. For a strictly convex body, a body strictly containing the line between any two points in the body, Gaussian curvature is everywhere positive, and there exists only one point on its surface having any particular orientation. A line parallel to the axis of a cylinder and tangent to the cylinder will lie entirely in the cylinder; all points on the line will be mapped to the same point by the Gauss map. To prevent multiplicity under the Gauss map, curvature must be strictly positive. For polytopes, conditions are slightly different; since all points on a

face of a polytope have the same orientation, convexity ensures that no two faces have the same orientation.

The Gaussian curvature, K , of a surface is a real-valued function on that surface, the determinant of the Jacobian of the Gauss map. Curvature dissects a surface into regions by the sign of K : *elliptic*, where curvature is positive, the principal curvatures are either both positive, denoted $(++)$ or both negative $(--)$; *hyperbolic*, curvature is negative, the signs of the principal curvatures are different $(+-)$; *parabolic*, at least one of the principal curvatures is zero (0) . Along parabolic lines the Gaussian curvature vanishes; these lines separate elliptic and hyperbolic regions. Refer to Section 2.3 for a discussion of the Gauss map, and curvature. On regular C^∞ surfaces [DoCarmo,1976], the Gauss map is continuous, so parabolic lines form smooth closed loops [Koenderink and van Doorn, 1980]. Koenderink has studied parabolic lines, both in the context of their photometric invariance[1982], and as characteristic lines of the surface, following Klein[in Hilbert and Cohn-Vossen,1952]. Stevens [1981] discusses the analysis of surface shape into regions of constant sign of Gaussian curvature.

When EGI methods are applied to non-convex objects, the underlying powerful theoretical results do not all generalize. The Gauss map is no longer an injection on the surface of a non-convex body. To use the EGI for reconstructing non-convex objects, the mapping from the object to the Gaussian sphere must be segmented in some fashion so that portions of the surface with the same orientation are not confused.

6.2.2 Segmenting Surfaces

Hoffman and Richards [1983] offer a theory for decomposing objects into parts as the human visual system. They cite the property of *transversality regularity*: two interpenetrating objects meet along a concave discontinuity of curvature. Based on this geometric principle, their segmentation scheme can account for reversals in perceived shape. This scheme is: segment bodies at minima of curvature along lines of curvature. While this scheme may account for some human perceptual behaviour, it does not lead to a simple decomposition scheme for machine vision based on orientation. A simple example shows (Figure 6.1) shows that the Gauss map is not one-one for a segmentation of a planar curve, using minima of curvature.

Differential geometry supplies a wealth of global results on convex bodies; it is much more difficult to discover relevant material for non-convex bodies. Here several of the known results are collected.

It would be very useful to know to what extent the curvature of a portion of a surface

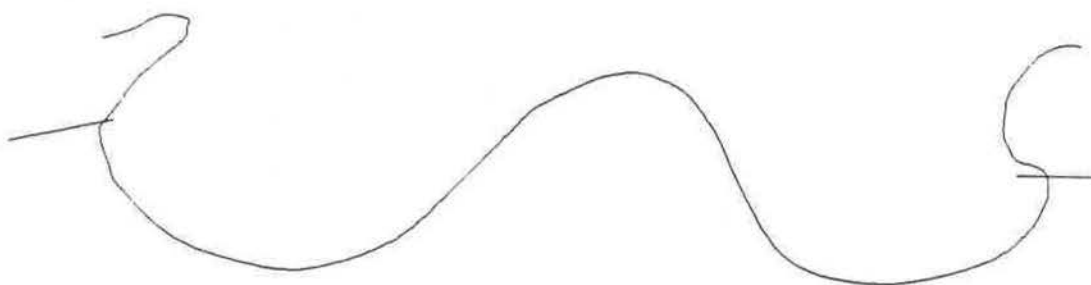


Figure 6.1: Plane Curve, Hoffman decomposition, not 1-1 Gauss map

constrains the multiplicity of the Gauss map. The degree of the Gauss map ($\text{deg}(G)$), the “signed” (by curvature) sum of the number of points in $G^{-1}(p)$ for any regular value p on U is:

$$\text{deg}(G) = X(M)/2 = 1 - g \quad (6.1)$$

where $X(M)$ is the Euler characteristic, $g = \text{genus} = \text{number of holes}$ [Spivak, 1970, V.3, p.414]. At a regular point the determinant of the Jacobian does not vanish. The determinant of the Jacobian for the normal map is the Gaussian curvature, so $\text{deg}(G)$ is not defined for parabolic points. $\text{deg}(G)$ is constant for all other points on S .

For example, the degree of the Gauss map on a torus is zero, since the torus has one hole ($g = 1$); there are two points mapping into any point on U , one at positive curvature and the other negative. All points on parabolic lines are excluded, which are the points aligned with the axes of the torus.

The function on U , the Gaussian sphere, denoted $\#_p$, is the number of points in $G^{-1}(p)$, for $p \in M$, a 2-manifold:

$$\int_M |K| dA = \int_{U-c} \#_p da = \text{total absolute curvature}$$

The integration is on U except for c , the set of critical points in G , but c is of measure 0, by Sard’s theorem. The total curvature is:

$$\int_M K dA$$

It can be shown that:

$$\int_{\{p \in M : K(p) > 0\}} K dA = 4\pi \iff G \text{ is one-one on } \{p \in M : K(p) > 0\} \quad (6.2)$$

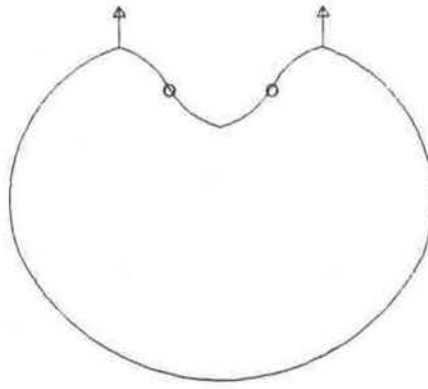


Figure 6.2: Plane curve and Curvature Regions Segmented at Vertical Normals

Spivak [1970, V.3, pp.414-419] shows that on a surface in R^3 for which the condition 6.2 holds, the regions of positive curvature form a subset of the boundary of a convex body. The boundaries of this subset of the body are convex plane curves, which are just the parabolic lines of the surface. This is a general characteristic of surfaces satisfying condition 6.2. The torus satisfies condition 6.2 so the positive curvature portion of the torus has a one-one Gauss map. Since the degree of the Gauss map is 0 for the torus, the region of negative curvature has one-one Gauss map. The torus is well-behaved; the sectioning of the torus at parabolic line gives two complete sheets each covering the entire Gaussian sphere.

6.3 Reconstruction

Consider accumulating in an EGI at each orientation, the total surface area which bears that orientation, but instead of summing the absolute value of the inverse of curvature, separating the different parts of the surface. The Gaussian sphere would then be multiply covered. Covering the Gaussian sphere with multiple sheets cannot, in general, be a unique object representation. In R^2 the separation into sheets is sufficient, since the topology (adjacency) is unambiguous. The simple example in Figure 6.2 shows plane curve where segmentation at points of inflection ($K = 0$) (the o's in the figure) does not produce one-one regions under the Gauss map. Segmentation at the points drawn with vertical normal vectors will divide the curve into one-one regions. Once a correct segmentation is done, reconstruction is possible. In R^3 the familiar topological difficulty recurs, so that the class of objects which are equivalent under a multiple sheet representation includes those in which translation of segmented elements in faces is allowed (see Figure 6.3). Parts A and B have the same EGI. Translating these parts in the supporting planar face will not affect any multiple sheet EGI. The effectiveness of the EGI for attitude determination and

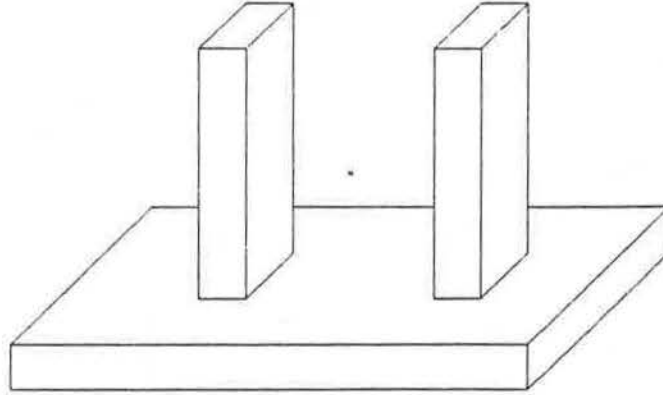


Figure 6.3: An object for which multiple-map EGI is not uniquely invertible

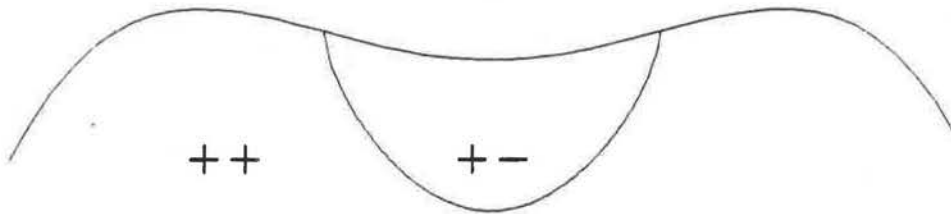


Figure 6.4: Elliptic parabolic and hyperbolic regions

recognition is its insensitivity to local coordinates. To capture fully the structure of the object in Figure 6.3, not only the topology of the regions but also the relative coordinate axes are necessary. An augmentation of the EGI with some structure to relate local coordinate axes of parts may be necessary, but is, in a sense, contrary to the spirit of the representation.

Reconstruction of general surfaces meets serious difficulties, because most surfaces will not subdivide into easily described regions for which the Gauss map is invertible. In general, non-convex bodies can have multiple inverse images of the Gauss map, even after sectioning by parabolic lines. An indication of the difficulties in segmentation for non-convex bodies can be found by examining a counterexample by Brou[1983,p.98], a spiralling figure, showing a large connected region of positive curvature in which the Gauss map covers U many times. A region of elliptic curvature ($++$) can be defined in which the Gauss map is not one-one (see Figure 6.4). This surface, an example from Koenderink[1982], has an elliptic($++$) region surrounding a hyperbolic($+ -$) region; in the elliptic region there are two points with the same normal. Higher degrees of covering are possible, by several methods, which are detailed by Koenderink. His analysis does not, however, suggest segmentations. The covering of U by the surface in Figure 6.4 can be segmented into regions for which the Gauss map is one-one. The boundaries

of these regions, however, are not defined by local measures such as curvature. Other criteria can be used to segment, such as tracking the inverse image (under of the Gauss map) of the parabolic lines on other parts of the surface.

6.3.1 Surfaces of revolution

It is shown here that a surface of revolution can be sectioned at parabolic lines into regions in which the Gauss map is one-one. A surface of revolution can be parametrized by θ , the angle about the axis of revolution, and $r(z)$, a function of z , which is taken to be the axis of revolution of the solid (following [Horn, 1984]). The derivation shows:

$$K = \frac{r_{zz}}{r(1+r_z^2)^2} \quad (6.3)$$

The curvature vanishes when r_{zz} does, so that the parabolic lines on the surface of revolution coincide with the revolution of points of inflection in $r(z)$. Because $r(z)$ is a plane curve and a one-one function of z , between the points of inflection the normal map on the unit circle is unique. The normal map on the surface is parametrized by θ around the surface, and by z vertically, so the Gauss map is unique between parabolic lines.

For example, the description of the curvature of the torus, parametrized by coordinates on the Gaussian sphere (as in [Horn,1984]) is, for portions of positive curvature:

$$K_+(\zeta, \eta) = \frac{1}{\rho} \frac{\cos(\eta)}{R + \rho \cos(\eta)}$$

Consider two tori T_0 and T_1 which have the same curvature function. By identifying the two functions,

$$\frac{1}{\rho_0} \frac{\cos(\eta)}{R_0 + \rho_0 \cos(\eta)} = \frac{1}{\rho_1} \frac{\cos(\eta)}{R_1 + \rho_1 \cos(\eta)}$$

Solving for R_1 , by eliminating $\cos(\eta) \neq 0$:

$$R_1 = \frac{\rho_0 R_0 + (\rho_0^2 - \rho_1^2) \cos(\eta)}{\rho_1}$$

Since R_1 is independent of $\cos(\eta)$, the coefficient of $\cos(\eta)$ must vanish, so:

$$\rho_0 = \rho_1$$

Substituting in the earlier expression, this shows that

$$R_0 = R_1$$

A similar arguments holds for negative Gaussian curvature. This demonstrates that the torus is uniquely specified by the EGI on the two portions of negative and positive curvature. From

two values of the curvature within a region of positive curvature, the two radii of the torus can be identified, completing the inversion of the EGI. This shows that inversion is possible for some non-convex bodies, specifically tori. For arbitrary surfaces of revolution, inversion requires solution of the differential Equation 6.3, which may be difficult.

6.4 Attitude Determination

The area function of a non-convex object N is equilibrated, since the surface is closed. The area function will correspond to some convex object C under shape reconstruction as described in Chapter 3. The visible area function of N may be identical to that of C from some viewpoints. For these viewpoints, the attitude determination method using the mixed volume of the support function of C and the area function of N will behave as if C in fact were the object, and correctly determine attitude. The situations are indistinguishable. The effects of self-occlusion play an important role in non-convex objects. The visible portions of a non-convex object are not necessarily those with orientations toward the viewing direction; self-occlusion obscures portions of the surface. Attitude determination by the method of mixed volumes requires well-behaved variation of the visible-area function with changing viewpoint.

If a suitable segmentation method can be described, then convex portions of a non-convex object can be treated in the same way as visible portions of convex objects. The attitude determined the portion of the object can be found in the same way as the attitude of the visible portion of a convex object is found. Should such results be unreliable in practice, the attitudes of portions hypothesized to belong to a single object could be combined into a more reliable global estimate of attitude.

6.5 Curvature Graphs and Recognition

Consider the connectivity graph of the regions generated by segmenting along parabolic lines. This graph inherits the invariance of curvature under Euclidean motion, and thus is an appropriate descriptor for recognition tasks. To facilitate intuition, one can consider the graph of an object such as a stylized baseball bat (Figure 6.5). The bat has three curvature regions: the tip of the bat (A) (elliptic $(++)$), the shaft (B) $(+-)$ and the knob at the end (C) $(++)$. Its graph consists of three nodes:

$$A_{++} \longleftrightarrow B_{+-} \longleftrightarrow C_{++}$$

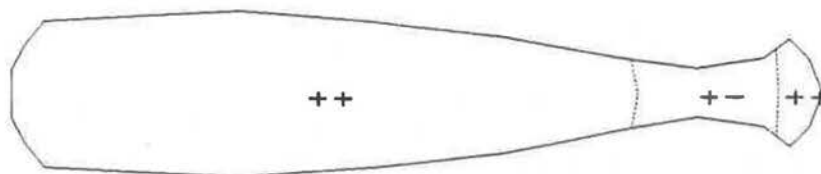


Figure 6.5: Curvature Regions on a bat

In an intuitive sense, this captures the notion of shape of a bat. Note that one can distinguish a bottle from a bat under this description - the bottle has two more regions: one hyperbolic and the other elliptic ($--$), both at the bottle opening. Koenderink and van Doorn [1980] describe some of the local structure of these graphs. Nackman and Pizer [1985] suggest similar curvature regions in their analysis of the use of the symmetric axis transform. They also mention slope districts which depend on the underlying axes generated by the transform. Curvature regions, of course, can be described without reference to these axes. The work of Asada and Brady [1983] on smoothed local symmetries generates decompositions of boundaries of regions at significant changes in curvature.

The powerful results involved in the fingerprint theorems for zero-crossings [Yuille and Poggio, 1984] apply to the analysis of parabolic lines, as they are zero-crossings of curvature. Curvature patches [Brady, 1984] segment surfaces into domains of curvature, bounded by parabolic lines and supplied with the coordinates curves by lines of principal curvatures. Such further decomposition of surfaces may be of significant use for recognition.

6.6 Estimating Curvature

Estimating curvature requires calculating derivatives, which is not numerically well-behaved. It may be possible to estimate the average derivative effectively for a region. Spacing is important, since a set of data may support a reliable estimate of orientation at one scale, and an estimate of curvature with similar reliability at a coarser scale, as higher derivatives are required. Terzopoulos [1983] gives examples of curvature estimation, using multi-resolution techniques. For the reconstruction results described above, the estimate of curvature must be accurate for segmentation. Of course, only the sign of curvature is necessary, and it may be possible to divide a surface roughly into elliptic, hyperbolic and (almost) parabolic regions.

Chapter 7

Conclusions and Open Questions

The constructions in Minkowski's proof of the uniqueness and existence of an inverse to an Extended Gaussian Image lead to a useful iterative algorithm for inversion. The novel geometric computation, the mixed volume, is central to this procedure. To reconstruct shape from an EGI, the mixed volume of the area function from the EGI and the support function of the reconstructed object is minimized. The development of this procedure answers the open question of the inversion of Extended Gaussian Images corresponding to polytopes.

This same minimization forms the core of a method for determining the attitude of a known convex object. The experiments described herein demonstrate the effectiveness of attitude determination by mixed volume and its insensitivity to small changes in attitude which affect other methods.

Non-convex objects require modification of the concept of EGI. Segmentation into curvature regions does not ensure uniqueness of the Gauss map, except for surfaces of revolution, which form a class of well-behaved non-convex objects. Generally, the Gauss map is not unique, however, even for convex regions of non-convex objects.

7.1 Open Questions

Several open questions arise from this analysis and implementation. The mixed volume measures shape similarity. How can it best be used to compare shapes? The mixed volume maps two polytopes at particular attitudes into a scalar value. What transformation of the two polytopes will leave the mixed volume unchanged? Does the mixed volume distinguish recognizable differences between polytopes?

It is possible to reconstruct the shape of a tetrahedron analytically from its EGI. No reduction

was found to show the possibility or impossibility of solving the set of area equations generated in reconstructing shape from the EGI analytically. It would be informative to find such a reduction. Another issue is the realizability of combinatorial types with faces of fixed orientations.

What is the expected behaviour of the convex hull step in the reconstruction, given the previous hull? The analysis of support functions suggested a new method for segmenting convex curves. Can this be generalized to surfaces? Finding centres of support provided several questions. Can the overlap regions for a polygonal curve be computed in less than $O(n^2)$ time, say in the size of the output? How long does this computation take for polytopes in R^3 ?

Recovering shape from an EGI can be posed as a discretization of a continuous non-linear problem. Pogorelov [1952] discusses general solutions to the so-called Minkowski problem. The differential equation to be solved is nonlinear, of the Monge-Ampere type. Assuming the surface is given by a function of two coordinates $F(x, y)$, then the equation is:

$$F_{xx}F_{yy} - F_{xy}^2 = G(F_x, F_y)$$

G is a function on the unit sphere. In a discussion of the regularity of the solution to this equation, Cheng and Yau[1976] state:

“Minkowski solved the problem for the category of polyhedrons. Then A.D. Aleksandrov and others solved the problem in general. However, this last solution does not provide any information about the regularity of the (unique) convex hypersurface even if we assume K is smooth.

In the two-dimensional case, H. Lewy[1938] was the first one who proved that if K is analytic, then the solution to the Minkowski problem is also analytic. Around 1953, A.V. Pogorelov and L. Nirenberg solved the regularity problem in the smooth category independently.”

For a discrete solution of the continuous problem, a regular mesh on U could be used. Is there a way of employing relaxation methods, perhaps even multigrid methods [Brandt, 1977], to the solution of this version of the problem? When reconstructing a portion of a polytope from partial information (see Section 3.8), using the projection of the occluding contour C , a number of different assignments of faces to edges may result in a closed path. It is an interesting combinatorial problem to determine bounds for the number of feasible assignments. Even more challenging is to give a method for determining this number exactly.

7.2 Future Work

Because minimizing the mixed volume is at the root of attitude determination method, any investigation of shape recovery may be beneficial for attitude determination. For a practical implementation of the reconstruction procedure, the convergence rate can be improved. It may be possible to take advantage of the ready availability of the Hessian matrix. The procedures for determining attitude for convex bodies are shown in experiments to work satisfactorily, without completing the EGI to the full sphere, but completion, which requires reconstruction, improves its reliability. So practical reconstruction implementations could become useful for attitude determination.

It has been pointed out [Mackworth, personal communication, 1985] that the optimization used in shape reconstruction can be accomplished by staying on the constraint surface $\langle A, H \rangle = c$, a hyperplane, and maximizing the volume $V(H)$. Since the constraint and objective function are convex, it is possible to switch their roles in this way. In either method, the volume and the gradient of volume must be computed, but in the proposed scheme, less scaling is performed. The performance of this method may be less subject to accumulated rounding error. Near the optimum, of course, the tangent hyperplane and the plane $\langle A, H \rangle = c$ are nearly parallel.

The minimization converges slowly when many small, almost-parallel faces occur in the EGI. Simplifying the EGI by removing these small faces, and redistributing the excess area to close the EGI, may speed convergence. In fact, the overall performance may be improved significantly by this technique, since the number of faces can be reduced by some constant ratio, leading to a hierarchical method. Each step would be less expensive and there would be fewer steps overall.

Attitude determination for non-convex objects may be solvable by methods similar to those for convex objects, provided the effects of self-occlusion are small. A better characterization of the limits of these uses and the behaviour of orientation-based representations under occlusion will be valuable.

The EGI can serve as a description of surfaces useful for finding significant features and for analyzing surfaces. Features on EGIs could be employed in recognition or attitude determination. These features could be interpreted topographically, as peaks and pits. On an EGI of a convex object, a pit corresponds to a large, slowly curving region. If the curvature of the EGI surface is high, then this slowly curving region is surrounded by edges in the object. A pit on the EGI corresponds to a vertex of the body, a point of high curvature.

Segmentation into curvature regions is not a general method for segmenting objects into regions in which the Gauss map is invertible. However, characterizing objects by the connectivity

graphs of curvature regions can be useful in preliminary processing for both recognition and attitude determination.

These representation schemes open up a large class of problems and provide new and interesting interpretations of problems in vision. That the solution of a problem of converting between object representations has suggested practical techniques for attitude determination is a happy result of the process of research.

Bibliography

- [1] R. Bajcsy, "Three-Dimensional Scene Analysis", *Proc. of 5th Int. Conf. on Pattern Recognition*, pp. 1064-1074, Miami Beach, Dec. 1980.
- [2] R. Mohr and R. Bajcsy, "Packing Volumes by Spheres", *Pattern Analysis and Machine Intelligence*, vol. 5, no. 1, pp. 111-116, 1983.
- [3] K. Sugihara, H.H. Baker, and T.O. Binford, "Depth From Edge and Intensity Based Stereo", *Proceedings of the Seventh IJCAI*, pp. 631-636, Vancouver, 1981.
- [4] H.G. Barrow and J.M. Tenenbaum, "Recovering Intrinsic Scene Characteristics From Images", *Computer Vision Systems*, ed. A.R. Hanson and E.M. Riseman, pp. 3-26, Academic Press, New York, 1978.
- [5] T.O. Binford, "Visual Perception by Computer", *IEEE Conference on Systems and Control*, Miami, Dec. 1971.
- [6] H. Blum, "A Transformation for Extracting New Descriptions of Shape", *Models for the Perception of Speech and Visual Form*, ed. W. Dunn, pp. 362-380, MIT Press, Cambridge, 1967.
- [7] H. Blum, "3-D Objects: Canonical Description and Quasi-Topologies via Symmetric Axis Coordinates", *Workshop on the Representation of 3-D Objects*, Philadelphia, May 1-2, 1979.
- [8] R.C. Bolles, P. Horaud, and M.J. Hannah, "3DPO: A Three-Dimensional Part Orientation System", *Proceedings of the Eighth IJCAI*, pp. 1116-1120, 1983.
- [9] M. Brady and H. Asada, "Smoothed Local Symmetries and Their Implementation", *The First International Symposium on Robotics Research*, Cambridge, Mass., 1983.
- [10] M. Brady, "Criteria for shape representations", *Human and Machine Vision*, ed. J. Beck and A. Rosenfeld, Academic Press, New York, 1983.

- [11] A. Brandt, "Multi-level adaptive solutions to boundary-value problems", *Mathematics of Computation*, vol. 31, pp. 333-390, 1977.
- [12] R.A. Brooks, "Symbolic Reasoning Among 3-D Models and 2-D Images", *Artificial Intelligence*, vol. 17, no. 1-3, pp. 285-348, 1981.
- [13] P. Brou, "Finding Objects in Depth Maps", Ph.D. Thesis, M.I.T. Department of Electrical Engineering and Computer Science, 1983.
- [14] P. Brou, "Using the Gaussian Image to Find the Orientation of Objects", *The International Journal of Robotics Research*, vol. 3, no. 4, pp. 89-125, Winter 1984.
- [15] C.M. Brown and H.B. Voelcker, "The PADL-2 project", *Proceedings of 7th NSF Conference on Production Research and Technology*, pp. F1-F6, September, 1979.
- [16] K.Q. Brown, "Fast Intersection of Half Spaces", CMU Technical Report CMU-CS-78-129, 1978.
- [17] K.Q. Brown, "Geometric Transforms for Fast Geometric Algorithms", CMU Technical Report CMU-CS-80-101, 1980.
- [18] S.Y. Cheng and S.T. Yau, "On the Regularity of the Solution of the n-Dimensional Minkowski Problem", *Communications on Pure and Applied Mathematics*, pp. 459-516, 1976.
- [19] Manfredo DoCarmo, "Differential Geometry of Curves and Surfaces", Prentice-Hall, Englewood Cliffs, NJ, 1976.
- [20] S.W. Draper, "The use of gradient and dual space in line-drawing interpretation", *Artificial Intelligence*, vol. 17, pp. 461-508, 1981.
- [21] W. Fenchel and B. Jessen, "Mengenfunktionen und konvexe Körper", Det Kgl. Danske Videnskab. Selskab. Math.-fys. Medd., 1938.
- [22] P.E. Gill, W. Murray, and M.H. Wright, "Practical Optimization", Academic Press, New York, New York, 1981.
- [23] W.E.L. Grimson, "From Images to Surfaces: A Computational Study of the Human Early Visual System", MIT Press, Cambridge, Mass, 1981.
- [24] B. Grünbaum, "Convex Polytopes", John Wiley and Sons, Ltd., London and New York, 1967.

- [25] L. Guibas, L. Ramshaw, and J. Stolfi, "A Kinetic Framework for Computational Geometry", *Proc. of the 24th Annual IEEE Symposium on Foundations of Computer Science*, 1983.
- [26] A. Guzman, "Decomposition of a Visual Scene Into Three-Dimensional Bodies", *Proc. AFIPS 1968 Fall Joint Computer Conference*, pp. 291-304, 1968.
- [27] D. Hilbert and Cohn-Vossen, "Geometry and the Imagination", Chelsea Publishers, 1952.
- [28] D.D. Hoffman and W.A. Richards, "Parts of Recognition", AI Memo 732, Artificial Intelligence Lab., Massachusetts Institute of Technology, 1983.
- [29] D.A. Huffman, "A duality concept for the analysis of polyhedral scenes", *Machine Intelligence*, ed. B. Meltzer and D. Michie, Edinburgh Univ. Press, Edinburgh, U.K., 1971.
- [30] B.K.P. Horn, "Obtaining Shape From Shading Information", *The Psychology of Computer Vision*, ed. P.H. Winston, pp. 115-155, McGraw-Hill, New York, 1975.
- [31] B.K.P. Horn, "Sequins and Quills - a representation for surface topography", *Representation of 3-dimensional Objects*, ed. R. Bajcsy, Springer-Verlag, Berlin and New York, 1982.
- [32] B.K.P. Horn, "The Curve of Least Energy", *ACM Transactions on Mathematical Software*, vol. 9, no. 4, December 1983.
- [33] B.K.P. Horn, "Extended Gaussian Images", *Proceedings of the IEEE*, pp. 1671-1686, December, 1984.
- [34] B.K.P. Horn and K.I. Ikeuchi, "The Mechanical Manipulation of Randomly Oriented Parts", *Scientific American*, August 1984.
- [35] K.I. Ikeuchi, "Recognition of 3-D Objects Using the Extended Gaussian Image", *Proceedings of the Seventh IJCAI*, pp. 595-600, 1981.
- [36] K.I. Ikeuchi, "Determining Attitude of Object From Needle Map Using Extended Gaussian Image", AI Memo 714, Artificial Intelligence Lab., Massachusetts Institute of Technology, 1983.
- [37] K.I. Ikeuchi and B.K.P. Horn, "Numerical Shape from Shading and Occluding Boundaries", *Artificial Intelligence*, vol. 17, pp. 141-185, 1981.
- [38] K.I. Ikeuchi, B.K.P. Horn, S. Nagata, T. Callahan, and O. Feingold, "Picking up an object from a pile of objects", AI Memo 726, Artificial Intelligence Lab., Massachusetts Institute of Technology, 1983.

- [39] T. Kanade, "Recovery of the Three Dimensional Shape of an Object from a Single View", *Artificial Intelligence*, vol. 17, pp. 409-461, 1981.
- [40] J.R. Kender, "Shape From Texture : an Aggregation Transform That Maps a Class of Textures Into Surface Orientation", *Proceeding of the Sixth International Joint Conference on Artificial Intelligence*, pp. 475-480, Tokyo, 1979.
- [41] D.E. Knuth, "Big Omicron and Big Omega and Big Theta", *SIGACT News*, April-June, 1976.
- [42] J.J. Koenderink and A.J. van Doorn, "Photometric invariants related to solid shape", *Acta Optica*, vol. 27, no. 7, pp. 918-996, 1980.
- [43] J.J. Koenderink and A.J. van Doorn, "The shape of smooth objects and the way contours end", *Perception*, vol. 11, pp. 129-137, 1982.
- [44] H. Lewy, "On differential geometry in the large I (Minkowski's problem)", *Trans. American Math. Society*, vol. 43, no. 2, pp. 258-270, 1938.
- [45] J.J. Little, "Automatic Registration of Landsat MSS Images to Digital Elevation Models", *Proc. Workshop on Computer Vision: Representation and Control*, pp. 178-184, Rindge, New Hampshire, 1982.
- [46] J.J. Little, "An Iterative Method for Reconstructing Convex Polyhedra from Extended Gaussian Image", *Proceedings of AAAI-83*, pp. 247-250, 1983.
- [47] D.G. Lowe, "Perceptual Organization and Visual Recognition", Stanford CS Dept. Report, Sept 1984 .
- [48] L.A. Lyusternik, "Convex Figures and Polyhedra", Dover Publications, New York, 1963.
- [49] A.K. Mackworth, "Interpreting Pictures of Polyhedral Scenes", *Artificial Intelligence*, vol. 4, no. 2, pp. 121-137, 1973.
- [50] Alan Mackworth and Farzin Mokhtarian, "Scale-Based Descriptions of Planar Curves", *submitted to CGIP*.
- [51] D. Marr, "Early Processing of Visual Information", *Phil. Trans. Royal Society of London*, vol. 275B, no. 942, pp. 483-524, 1976.

- [52] D. Marr, "Analysis of Occluding Contour", *Proc. Royal Soc. London*, vol. B, no. 197, pp. 441-475, 1977.
- [53] David Marr and H.K. Nishihara, "Representation and Recognition of the Spatial Organization of Three Dimensional Structure", *Proc. Royal Soc. London*, vol. B, no. 200, pp. 269-294, 1978.
- [54] D. Marr, "Vision: A Computational Investigation into the Human Representation and Processing of Visual Information", W.H. Freeman, San Francisco, 1982.
- [55] Herman Minkowski, "Allgemeine Lehrsätze über die konvexe Polyeder", *Nachr. Ges. Wiss. Göttingen*, pp. 198-219, 1897.
- [56] A.P. Morgan, "A Method for Computing All Solutions to a System of Polynomial Equations", Report GMR-3651, Mathematics Department, GM Research Labs, January 1982.
- [57] J.A. Mulder, "Representing Ambiguity and Hypotheticality in Visual Knowledge", Ph.D. Thesis, U.B.C. Department of Computer Science, 1985.
- [58] L.R. Nackman, "Three-Dimensional Shape Description Using the Symmetric Axis Transform", Ph.D. dissertation, Dept. Comput. Sci., Univ. North Carolina, 1982.
- [59] L.R. Nackman and S.M. Pizer, "Three-Dimensional Shape Description Using the Symmetric Axis Transform I: Theory", *Pattern Analysis and Machine Intelligence*, vol. 7, no. 2, pp. 187-203, 1985.
- [60] Louis Nirenberg, "The Weyl and Minkowski Problems in Differential Geometry in the Large", *Communications on Pure and Applied Mathematics*, vol. VI, pp. 337-394, 1953.
- [61] H.K. Nishihara, "Intensity, Visible-Surface, and Volumetric Representations", *Artificial Intelligence*, vol. 17, pp. 265-284, 1981.
- [62] A.V. Oppenheim, A.S. Willsky, and I.T. Young, "Signals and Systems", Prentice Hall, Englewood Cliffs, NJ, 1983.
- [63] C.H. Papadimitriou and K. Steiglitz, "Combinatorial Optimization: Algorithms and Complexity", Prentice Hall, Englewood Cliffs, NJ, 1982.
- [64] A.V. Pogorolev, "The Minkowski Multidimensional Problem", Winston and Sons, 1978.

- [65] F.P. Preparata and S.J. Hong , "Convex Hulls of Finite Sets of Points in Two and Three Dimensions", *CACM* , vol. 20, pp. 87-93, 1977.
- [66] A.A.G. Requicha, "Representations for Rigid Solids: Theory, Methods, and Systems", *ACM Computing Surveys*, vol. 12, no. 4, pp. 437-464, Dec. 1980.
- [67] E. Salamin, "Applications of Quaternions to computations with rotations", Stanford Artificial Intelligence Laboratory, Internal Working Paper, 1979.
- [68] R. Seidel, "A Method for Proving Lower Bounds for Certain Geometric Problems", TR 84-592, Feb. 1984.
- [69] D.A. Smith, "Using Enhanced Spherical Images", AI Memo 530, Artificial Intelligence Lab., Massachusetts Institute of Technology, May, 1979.
- [70] M. Spivak, "A Comprehensive Introduction to Differential Geometry, Volume One", Publish or Perish, Inc., Wilmington, DE, 1970 .
- [71] J.L. Stansfield, "Conclusions from the Commodity Expert Project", AI Memo 601, Artificial Intelligence Lab., Massachusetts Institute of Technology, 1980.
- [72] E. Steinitz, "Polyeder und Raumein teilungen", *Enzykl. Math. Wiss. Vol. 3, (Geometrie) Part 3AB12*, pp. 1-139, 1922.
- [73] K. A. Stevens, "The Visual Interpretation of Surface Contours", *Artificial Intelligence*, vol. 17, pp. 47-75, 1981.
- [74] K. Sugihara, "Mathematical Structures of Line Drawings of Polyhedrons - Toward Man-Machine Communication by Means of Line Drawings", *Pattern Analysis and Machine Intelligence*, vol. 4, pp. 458-468, 1982.
- [75] K. Sugihara, "An $n \log n$ Algorithm for Determining the Congruity of Polyhedra", *Journal of Computer and System Sciences*, vol. 29, pp. 36-47, 1984.
- [76] D. Terzopoulos, "The Role of Constraints and Discontinuities in Visible-Surface Reconstruction", *Proceedings of the Eighth IJCAI*, pp. 1073-1077, 1983.
- [77] W.T. Tutte, "A Census of Planar Triangulations", *Canadian Journal of Math.* , vol. 14, pp. 21-38, 1962.

- [78] D.L. Waltz, "Understanding Line Drawings of Scenes with Shadows", *The Psychology of Computer Vision*, ed. P.H. Winston, pp. 19-91, McGraw-Hill, New York, 1972.
- [79] E.J. Weldon, Jr. and B.K.P. Horn, "Filtering Closed Curves", in press, 1984.
- [80] A.P. Witkin, "Recovering Surface Shape and Orientation from Texture", *Artificial Intelligence*, vol. 17, pp. 17-47, 1981.
- [81] A.P. Witkin, "Scale-space Filtering", *Proceedings of the Eighth IJCAI*, pp. 1019-1023, 1983.
- [82] R.J. Woodham, "Photometric Method for Determining Surface Orientation from Multiple Images", *Optical Engineering*, vol. 19, pp. 139-144, 1980.
- [83] A.L. Yuille and T. Poggio, "Fingerprint Theorems for Zero Crossings", AI Memo 730, Artificial Intelligence Lab., Massachusetts Institute of Technology, 1983.

Appendix A

Volumes of Mixtures and Mixed Volumes

In this appendix it is proved that the volume (area) of the mixture of two polygons P and Q can be expressed in terms of λ , the mixing parameter, and in terms of the volumes of the two polygons and a term called the "mixed volume", which expresses the relation between the shapes of the two polygons. The treatment of the proof that follows is from Lyusternik[1963]. Translate a convex polygon P so that the origin is an interior point of P . Its area can be expressed in terms of the triangles formed by the origin and the endpoints of the sides of P . The area of each triangle is the product of the length of its base, the side A_{P_i} , and its height, which is the perpendicular distance from the origin to the line containing the side. If ω_i is the normal to side i , the height of any triangle is H_{P_i} , the value of the support function $\mathcal{H}_P(\omega_i)$.

The area of P , termed $V(P)$, is

$$V(P) = 1/2 \sum_{i=1}^n A_{P_i} * H_{P_i} \quad (A.1)$$

Consider two polygons P and Q whose sides are ordered to correspond in direction. By means of introducing sides of zero length any two arbitrary polygons can be made to satisfy this condition. Denote the vertices of P by p_i and the corresponding vertices of Q by q_i . Lengths of corresponding sides are denoted by A_{P_i} and A_{Q_i} . Let the support functions for the sides of P and Q be H_{P_i} and H_{Q_i} . Then the areas of P and Q are

$$V(P) = 1/2 \sum_{i=1}^n A_{P_i} * H_{P_i} \quad (A.2)$$

$$V(Q) = 1/2 \sum_{i=1}^n A_{Q_i} * H_{Q_i} \quad (A.3)$$

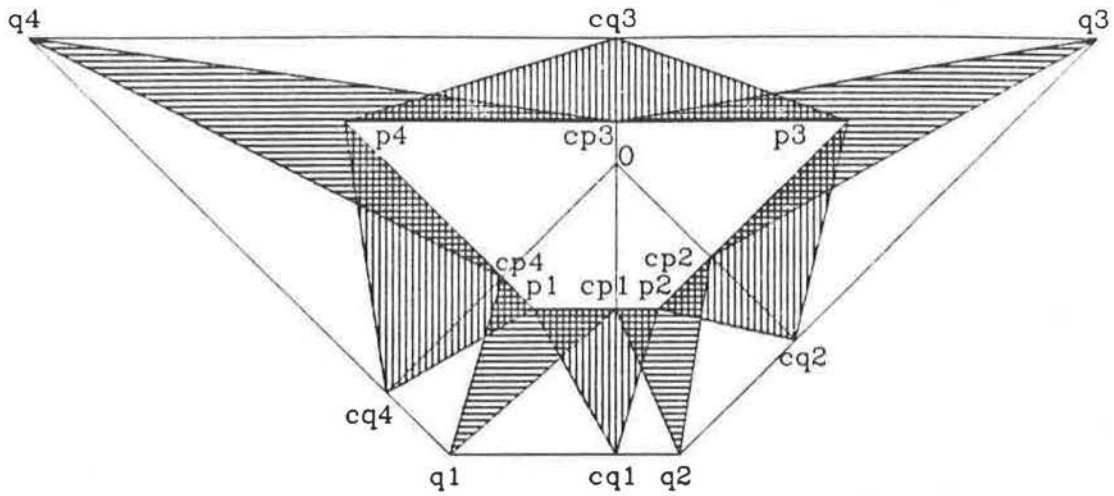


Figure A.1: Construction showing mixed volumes are equal

Let $R = \lambda P + (1 - \lambda)Q$. The support functions for the sides of R are

$$H_{R_i} = \lambda * H_{P_i} + (1 - \lambda) * H_{Q_i} \quad (A.4)$$

The area of R , from Equations 3.2, A.4, and A.1 is

$$V(R) = 1/2 \sum_{i=1}^n A_{R_i} * H_{R_i} = 1/2 \sum_{i=1}^n (\lambda * A_{P_i} + (\lambda - 1) * A_{Q_i}) * (\lambda * H_{P_i} + (\lambda - 1) * H_{Q_i}) \quad (A.5)$$

Combining terms, one obtains a new expression for $V(R)$:

$$\lambda^2 [1/2 \sum_{i=1}^n A_{P_i} * H_{P_i}] + 2\lambda * (\lambda - 1) [1/2 \sum_{i=1}^n A_{P_i} * H_{Q_i} + 1/2 \sum_{i=1}^n A_{Q_i} * H_{P_i}] + (\lambda - 1)^2 [1/2 \sum_{i=1}^n A_{Q_i} * H_{Q_i}] \quad (A.6)$$

By Equations A.2 and A.3 the coefficients of λ^2 and $(\lambda - 1)^2$ in Equation A.6 are respectively equal to $V(P)$ and $V(Q)$.

Consider the two sums in the coefficient of $\lambda * (\lambda - 1)$. From the origin O drop perpendiculars to the sides of P and Q , denoting the feet of the perpendiculars by cp_i and cq_i . The lengths of $O cp_i$ and $O cq_i$ are H_{P_i} and H_{Q_i} respectively. Connect each point cq_i with the two corresponding vertices p_i and p_{i+1} of P (all indices are modulo n). The resulting polygon is shown in Figure A.1, as the outer border of the area shaded with vertical lines. Its area can be calculated as the sum of the quadrilaterals $Op_i cq_i p_{i+1}$, each of whose area is one half the product of the lengths of its diagonals, since they are perpendicular, and is $1/2 A_{P_i} * H_{Q_i}$. Its area is

$$1/2 \sum_{i=1}^n A_{P_i} * H_{Q_i} \quad (A.7)$$

Similarly construct the polygon $q_1 cp_1 q_2 cp_2 \dots q_n cp_n q_1$, shown as the outer border of the area shaded with horizontal lines in Figure A.1. It is composed of the triangles $Ocp_i cq_i$ and $Oq_{i+1} cp_i$.

Each triangle has base Ocp_i and altitude equal to the length of the segment $q_{i+1}cq_i$ (denoted Q_{i2}), - its area is $1/2H_{P_i}Q_{i2}$; also the triangle $Oq_i cp_i$ has the same base and its altitude is the length of the segment $cq_i q_i$ (denoted Q_{i1}) - its area is then $1/2H_{P_i}Q_{i1}$. Since $Q_{i1} + Q_{i2} = A_{Q_i}$, the sum of these areas equals $1/2A_{Q_i}H_{P_i}$; the sum of the areas of all pairs of such triangles equals the area of the polygon $q_1cp_1q_2cp_2 \dots q_ncp_nq_1$.

$$1/2 \sum_{i=1}^n A_{Q_i} * H_{P_i} \quad (A.8)$$

The polygon $p_1cq_1p_2cq_2 \dots p_ncq_n p_1$ consists of the polygon P plus the $2n$ triangles (shaded with vertical lines) $p_1cq_1cp_1, cp_1cq_1p_2 \dots p_ncq_ncp_n, cp_ncq_n p_1$. The polygon $q_1cp_1q_2cp_2 \dots q_ncp_nq_1$ comprises the polygon P plus the $2n$ triangles (shaded with horizontal lines) $p_1q_1cp_1, cp_1q_2p_2 \dots p_nq_ncp_n, cp_nq_1p_1$.

The triangles horizontally and vertically shaded are pairwise of equal area in the order in which they are written. For example, $cp_1q_1p_1$ and $p_1cq_1cp_1$ have the common base p_1cp_1 and the vertices opposite them, q_1 and cq_1 , lie on a line parallel to the base, so they have the same altitude. Thus the triangles have the same area. The areas of the two polygons, $q_1cp_1q_2cp_2 \dots q_ncp_nq_1$ and $p_1cq_1p_2cq_2 \dots p_ncq_n p_1$, must therefore be equal. It follows that

$$1/2 \sum_{i=1}^n A_{P_i}H_{Q_i} = 1/2 \sum_{i=1}^n A_{Q_i}H_{P_i}$$

The expression

$$1/2 \sum_{i=1}^n A_{P_i}H_{Q_i} = 1/2 \sum_{i=1}^n A_{Q_i}H_{P_i}$$

is called the mixed volume of polygons P and Q , and is denoted ${}_2V(P, Q)$. Equation A.6 may then be written in the form

$$V(R) = \lambda^2 V(P) + 2\lambda * (\lambda - 1)V_2(P, Q) + (\lambda - 1)^2 V(Q) \quad (A.9)$$

When the two polygons are homothetic, i.e., related by translation and scaling, the mixed volume is related to the volumes of P and Q by the appropriate scaling.

Minkowski's proof for d-polytopes relies on a similar decomposition into d-simplices (tetrahedra for $d = 3$).

Appendix B

Matching EGIs Discretely

Orientations are a set of points on the unit sphere U ; form a tessellation of these points on U . The area measurements are sorted, so they can be considered as integers from 1 to n , the number of orientations. The tessellation of points on the unit sphere forms a planar graph. Use an algorithm for planar graph isomorphism [Hopcroft and Wong, 1974] which takes linear time, to determine an isomorphism between the EGI of a prototype and a sensed EGI. This isomorphism describes a one-to-one mapping between the two graphs from which one can determine the rotation taking the sensed EGI into the prototype EGI. The tessellation graphs of the EGIs are identical. But Seidel has suggested annotating each node of the graph with a small planar graph indicating the order of that particular node's area measurement in the sorted list of areas. Each graph would be composed of a chain of $\log n$ nodes, from each of which is attached either a single node, indicating a 0 in that bit position, or a chain of two nodes, indicating a 1 in that bit position. These small graphs are planar, there are n different such graphs, and each contains at most $2 * \log n$ nodes, so the annotated tessellation graph contains at most $2 * n * \log n$ nodes. Determining graph isomorphism on such a graph can be performed in $O(n * \log n)$ time, which is required in any case for sorting the areas.

Appendix C

Example Polytopes

Contained herein are the figures showing the polytopes used in many of the examples in the text.

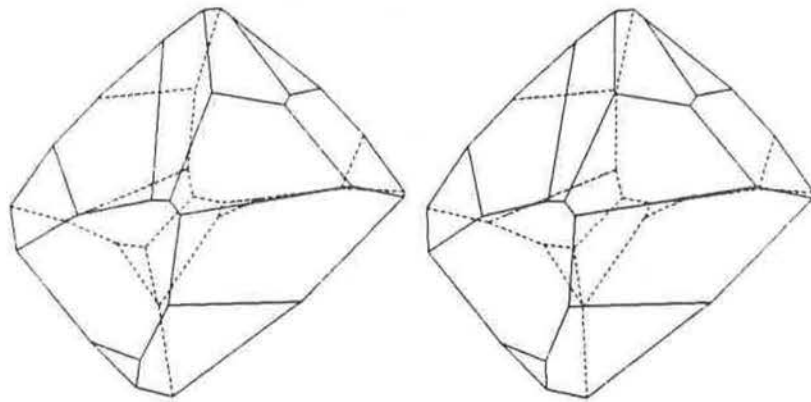


Figure C.1: Polytope with 21 faces

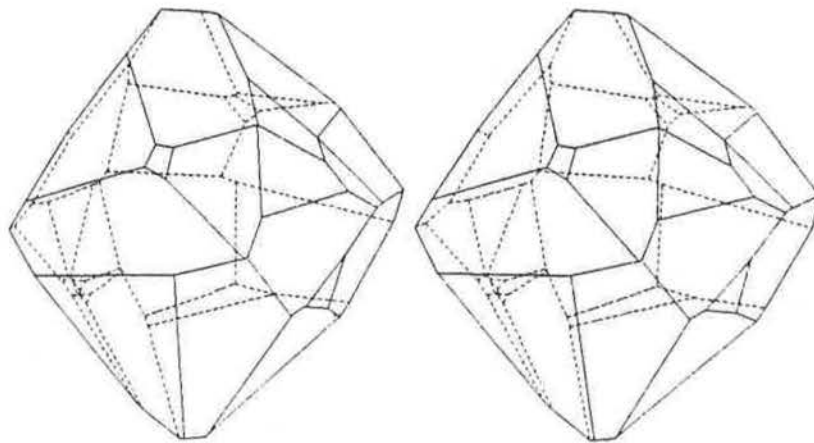


Figure C.2: Polytope with 40 faces

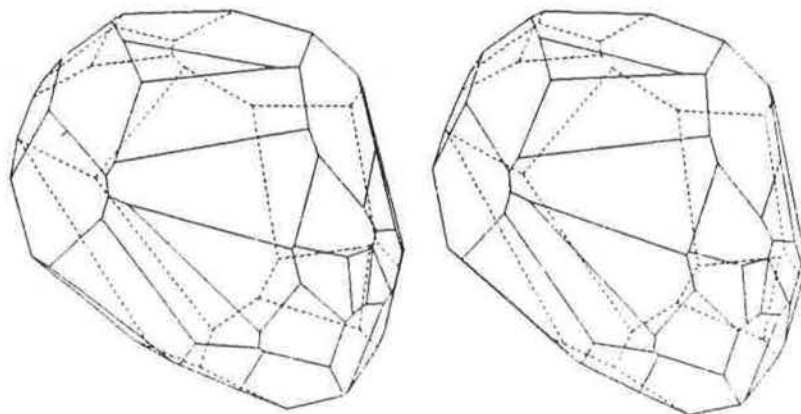


Figure C.3: Polytope with 40 faces on ellipsoid

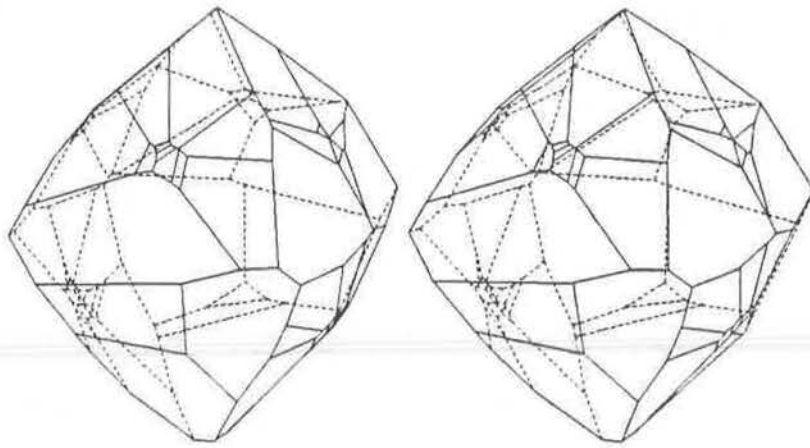


Figure C.4: Polytope with 80 faces on ellipsoid