Recognizing Shapes
Via Random Chord Samplings

R. Hashim
and
W.N. Martin

Computer Science Report No. TR-86-07
May 14, 1986

This work will be presented at the Computer Vision and Pattern Recognition Conference, Miami, Flordia, June 22-26, 1986, in a paper by the same title. This extended form, however, is taken from the Master of Science thesis entitled, Athena: A system for Shape Recognition, by Raza Hashim.

ABSTRACT

A brief introduction to various methods of two dimensional shape recognition is given. A procedure for image-plane rotation and translation invariant object recognition is described. The method is based on sampling random chords from the greylevel image of an object and using the texture information associated with these chords to recognize the object. A voting scheme based on a generalized Hough transform performs the recognition. Schemes for organizing random chord shape descriptions and for matching the textural information associated with the chords are presented in terms of the overall voting scheme. Examples are given which show the effectiveness of this technique. Finally, directions of future research are discussed.

TABLE OF CONTENTS

1. Introduction 1

1.1. The Problem of Recognizing Shapes 1

1.2. Some Methods for Shape Recognition 2

1.2.1. Template Matching 2

1.2.2. Fourier Descriptors 4

1.2.3. Moment Invariants 6

1.3. Organization of the Mammalian Visual Pathway 7

1.4. Connectionist Models for Shape Recognition 9

1.5. Chord Distributions For Recognizing Shapes 10

1.6. Proposal for a New System for Recognizing Shapes 10

2. Background 12

2.1. The Hough Transform 12

1.2.1. Hough Transforms for Detecting Analytic Shapes 12

2.1.2. Generalization of the Hough Transform 13

2.2. Chord Distribution Methods 15

2.3. A Possible Connectionist Implementation of Athena 17

3. Design and Implementation of Athena 19

3.1. System Design 19

3.2. Organization of Shape Description and Chord Matching 25

3.3. Reference Point Hough Transform 33

3.4. Detecting Object Orientation 36

4. Examples Using Athena 45

4.1. Progression of Chords 45

4.2. Example 1 45

4.3. Example 2 50

4.4. Example 3 53

4.5. Example 4 53

4.6. Example 5 58

5. Conclusion 63

5.1 Summary 63

5.2. Future Research 63

References 65

Appendix A 70

Appendix B 73

TABLE OF FIGURES

Figure 1. Overall design of Athena 20

Figure 2.a. AquireImage component of Athena 22

Figure 2.b. GenerateLibrary component of Athena 23

Figure 2.c. RecognizeShape component of Athena 24

Figure 3.a.1. I-function 27

Figure 3.a.2. Non-matching I-function for Figure 3.a.1. 28

Figure 3.b.1. I-function 29

Figure 3.b.2. Matching I-function of Figure 3.b.1. 30

Figure 3.c.1. I-function 31

Figure 3.c.2. I-function that matches Figure 3.c.1. under reflection 32

Figure 4. Chords from an object description 35

Figure 5. Chords from the library and the viewframe 37

Figure 6. Contents of the voting array for Figure 5. 37

Figure 7. Physically rotated object 39

Figure 8. Chords from the library and viewframe 41

Figure 9. Chords from the library and viewframe 41

Figure 10. Chords from the library and viewframe 43

Figure 11. Plot of the voting array for Figure 10. 43

Figure 12. Example data set used in the experiments 46

Figure 13. Progression of chords from D 46

Figure 14. Setup for example 1 48

Figure 15. Contents of the voting array for Figure 14. 49

Figure 16. Setup for example 2 51

Figure 17. Contents of the voting array for Figure 16. 52

Figure 18. Setup for example 3 54

Figure 19. Contents of the voting array for Figure 18. 55

Figure 20. Setup for example 4 56

Figure 21. Contents of the voting array for Figure 20. 57

Figure 22. Setup for example 5 59

Figure 23. Plot of the voting array for example 5 60

CHAPTER 1

INTRODUCTION

1.1. The Problem of Recognizing Shapes

Numerous methods exist for the analysis of two dimensional shapes. Several comprehensive surveys of the major techniques for the analysis of silhouettes of shapes [1,2,3] have been made, providing classifications of these techniques under criteria such as : information-preserving vs. non-information-preserving, boundary tracing vs. interior examining, scalar transform vs. space domain transforms, etc..

The shape recognition problem can be posed as : given a set of known shapes, i.e., a library, with a description of each shape in a frame of reference local to the shape, and given an unknown viewframe of shape features, find a correspondence between the features in the viewframe and the features in the known shape descriptions, if such a correspondence exists. In other words, the shape recognizer considers the set of features in the viewframe, and decides whether that set of features can be explained as one of the shapes in the library.

There are a number of factors that effect the shape recognition task. Among such factors we can include occlusion, segmentation of the shapes from textured backgrounds, knowledge about lighting conditions and reflectance properties of objects, noise in the image, information about the boundary of the shape, and knowledge about the rotation, translation and size of the shapes. The generality of a shape recognition algorithm can be measured by what the algorithm assumes regarding the above mentioned factors and how the behavior of the algorithm is altered if these assumptions are violated.

This study is an attempt to develop a new image plane rotation and translation invariant shape recognition system which utilizes some ideas from existing shape recognition algorithms, taking into consideration the architecture of biological vision systems. By "shape" we will mean not only the boundary, but more importantly the texture inclosed by that boundary. After a brief survey of some of the existing algorithms for shape recognition, and a discussion of the neurological evidence for connectionist processing in natural vision systems, the design of a new system is discussed. Finally, the implementation details and some sample runs are reported. This study concludes with a discussion of possible extensions of the system to solve related problems in computer vision, such as dynamic scene analysis.

1.2. Some Methods for Shape Recognition

This section is a brief review of three methods for shape recognition : template matching, Fourier descriptors and moment invariants. A brief analysis of the related advantages and disadvantages of each technique is given. Template matching is discussed for historical reasons, as it is one of the earliest attempts at solving the problem of recognizing shapes. Fourier descriptors and moments are reviewed because these methods have addressed image plane rotation and translation invariant shape recognition — the central topic of this study.

1.2.1. Template Matching

Template matching was one of the first techniques used for shape recognition. This procedure can be viewed as passing a window, $W(p,q)$, containing a specific pattern over an arbitrary image, $I(i,j)$, in order to search for an occurrence of $W(p,q)$ in $I(i,j)$. The window is passed over the image by positioning the window at different locations in the image, and computing a similarity measure for each position. The position corresponding to the highest value of the similarity measure is declared to be the position of the template in the unknown image.

Some studies have explained template matching as a dot product of vectors [4], or as cross-correlation [5] between $W(p,q)$ and $I(i,j)$. Let the image $I(i,j)$ be of dimension $M \times N$, with i in $[1,M]$ and j in $[1,N]$, and let $W(p,q)$ be of dimension $(2m+1) \times (2n+1)$, with p in $[-m,m]$ and q in $[-n,n]$. Generally m and n are chosen such that $m << M$ and $n << N$. For the cross-correlation method [5], $W(p,q)$ and $I(i,j)$ are compared with $W(p,q)$ located at (k,l) as follows :

$$\rho(k,l) = \sum_{s=-m}^{m} \sum_{t=-n}^{n} W(s,t)\ I(k+s,l+t) \tag{1.1}$$

where

$$m \leqslant k \leqslant (M-m)$$

$$n \leqslant l \leqslant (N-n)$$

In Equation (1.1) $\rho(k,l)$ represents the value of discrete cross-correlation between $W(p,q)$ and $I(i,j)$ at location (k,l). This process can be imagined as shifting the template across the image to different offsets, and multiplying the superimposed values of the template and image at each offset, and summing the products to give the cross-correlation value at a particular offset.

The value of $\rho(k,l)$ can be adversely affected if there are large values of random noise in the array $I(i,j)$. In this case, the value of $\rho(k,l)$ will be very high. In order to minimize the effect of noise in $I(i,j)$ on the value of ρ, a normalized cross-correlation between $W(p,q)$ and $I(i,j)$ has been proposed [5], and is given by :

$$\rho'(k,l) = \frac{\rho(k,l)}{\left[\sum_{s=-m}^{m} \sum_{t=-n}^{n} W^2(s,t)\right]^{\frac{1}{2}} \left[\sum_{s=-m}^{m} \sum_{t=-n}^{n} I^2(k+s,l+t)\right]^{\frac{1}{2}}} \tag{1.2}$$

where,

$$m \leqslant k \leqslant (M-m)$$
$$n \leqslant l \leqslant (N-n)$$

An analysis of the template matching technique points out some disadvantages in using this method. Template matching is difficult to use in presence of occlusion or noise in the unknown image. To detect randomly oriented objects in the unknown image, a procedure can be adopted that looks at the cross-correlation between the input image and windows containing the object in different orientations. The orientation of the object in the unknown scene, then, would be the orientation of the template yielding the highest correlation value. Similarly, the change in scale can be detected by computing the cross-correlation between the input image and templates containing different sized instances of the same object. The size of the object in the unknown scene will correspond to the size of the template that has the highest correlation value. All these modifications to the template matching procedure, however, are computationally expensive.

Some special purpose architectures [6] have also been proposed to provide hardware support for template matching. These architectures try to overcome some of the disadvantages of template matching by providing powerful processing elements, and using parallel processing techniques.

1.2.2. Fourier Descriptors

Fourier descriptors of the boundary of shapes have also been used for shape recognition [7,8]. This method has been successfully used for such diverse tasks as recognizing hand written characters [8], automatic inspection of machine parts [9] and three dimensional aircraft recognition [10]. This method assumes that the structural information found in the boundary of a shape is sufficient to recognize the shape. Hence, information about the texture of the object is not kept. One study has described the Fourier transform of the boundary of a shape in terms of the tangent angle versus the arc length along a curve [8]. In this method a starting point on the boundary of the curve is selected and a function, $\theta(l)$, is defined which measures the angular direction of the tangent to the curve as a function of arc length from the selected starting point. This periodic function is expanded in a Fourier series after it has been normalized with regard to starting point selection, length and orientation, and the first few terms of this expansion are used as the shape description.

Let γ be a clockwise oriented, simple, closed curve with parametric representation, $(x(l),y(l)) = Z(l)$, where l is the arc length along the curve from a selected starting point, and $0 \leqslant l \leqslant L$. Let $\theta(l)$ denote the angular direction of the tangent to γ at point l, and let $\phi(l)$ denote the angle between the tangent direction at the starting point, $l = 0$, and at the point, l. A "normalized cumulative angular bend" function, ϕ^*, which is invariant to translation, rotation and changes in scale, i.e., length of perimeter L, has been defined [7] as :

$$\phi^*(t) = \phi\left(\frac{Lt}{2\pi}\right) + t \tag{1.3}$$

where,

$$0 \leqslant t \leqslant 2\pi$$

$\phi^*(t)$ can be expanded in a Fourier series and the first few terms of this series can be used as shape descriptors. The Fourier expansion of $\phi^*(t)$ is given by :

$$\phi^*(t) = \mu_0 + \sum_{k=1}^{\infty} A_k \cos(kt - \alpha_k) \tag{1.4}$$

It has been pointed out that some sequences $\{A_k, \alpha_k\}$ do not describe closed curves. Moreover, it is difficult to handle discontinuities in the boundary of the shape. Reconstruction of the boundary of the object using the Fourier descriptors requires numerical integration. For a more detailed analysis of the advantages and disadvantages of specifying Fourier descriptors as cumulative angular functions, see [11]. A mathematical treatment of Fourier descriptors as complex periodic functions of arc length can be found in [7].

It should be noted that Fourier descriptors are global descriptors of the boundary of the object. Local variations in the boundary are not localized in the Fourier descriptors of the boundary. As a result, the characteristics of the Fourier descriptors of shapes change drastically with occlusion. This characteristic of the Fourier descriptors makes recognition of partially hidden objects difficult, as the local boundary features cannot be extracted from the global boundary description provided by the Fourier descriptors.

1.2.3. Moment Invariants

Some studies have used measures invariant to rotation and translation based on moments of the boundary or the silhouette of an object to classify shapes. Moments have been used in many domains including problems such as the recognition of printed characters and numerals [12], chest X-ray analysis [13], aircraft identification [14], and ship recognition [15]. Let $f(x,y)$ be an $(m \times n)$ image. The two dimensional $(p+q)th$ order moment of a digital image, $f(x,y)$, which is insensitive to translation, is defined as [5]:

$$\mu_{pq} = \sum_{x=1}^{m} \sum_{y=1}^{n} (x-\overline{x})^p (y-\overline{y})^q f(x,y) \tag{1.5}$$

where

$$\overline{x} = \frac{\sum_{x=1}^{m} \sum_{y=1}^{n} x f(x,y)}{\sum_{x=1}^{m} \sum_{y=1}^{n} f(x,y)} \tag{1.6}$$

$$\bar{y} = \frac{\sum_{x=1}^{m} \sum_{y=1}^{n} y\, f(x,y)}{\sum_{x=1}^{m} \sum_{y=1}^{n} f(x,y)} \tag{1.7}$$

For boundaries and silhouettes the image is a binary valued function, i.e., $f(x,y) = 0$ *or* 1. Using μ_{pq}, one study has derived a set of seven second and third order moments that are insensitive to image plane rotation [16]. In another study [17], these rotation and translation invariant moments, along with a circular template matching scheme, have been used to identify shapes. A two stage matching procedure is employed by that scheme to provide an inexpensive measure (the zeroth-order moment) in the first stage to determine the likely position in $f(x,y)$ where the template might exist. The scheme then uses a more accurate and computationally complex measure (second and third order moments) in the second stage to determine the best match among the likely ones.

Moments, like Fourier descriptors, are information preserving in nature, and provide a global description of the shape. Therefore, methods based on moments face many of the same problems with occlusion and partial views. A special purpose architecture has also been proposed for methods based on moments [18].

1.3. Organization of the Mammalian Visual Pathway

A study of the organization of the visual pathway shows that the output from each eye is conveyed to the brain by the optic nerve, which is a bundle of the axons of the ganglion cells found in the retina. A large fraction of the optic nerve fibers pass uninterrupted to the lateral geniculate nuclei, which in turn have axons that extend to the primary visual cortex.

The distribution of photoreceptors, e.g. rods and cones, in the retina is not uniform. As we move away from the fovea, the concentration of rods and cones per unit area decreases. The visual cortex, on the other hand, has a remarkably uniform structure. A region having a high concentration of photo receptors in the retina maps to more cortical area than a region with a lower concentration of photoreceptors [19]. Experiments done with the visual cortex of monkeys [20] have given fundamental insights into the organization of the visual cortex.

The ganglion cells found in the retina, and the lateral geniculate nuclei, respond best to a roughly circular spot of light of a particular size in a particular part of the visual field. Each cell has a center-surround receptive field, with the center being excitatory and surround being inhibitory, or exactly the reverse configuration. The neurons at this level of visual processing are concerned with making a comparison between the light level in a very small area of the visual scene and the average illumination of the immediate surround [20].

Some studies have shown [20] that the operations the cortex performs on the information it receives are strictly local. The details of neural connections between inputs and outputs differ from one area to the next, but within a given area the neural connections seem to be rather stereotyped.

Experimental evidence suggests that neurons in the cortex are organized in a hierarchy, with each level in the hierarchy responding to simple stimuli and feeding higher levels. For example, neurons in the lowest level of the hierarchy, which receive input from the lateral geniculate nuclei, respond best to specifically oriented line segments in a particular position in the visual field. As we go higher in this hierarchy, we find "complex" neurons that are less particular about the exact position of a line.

Thus the overall organization of the visual cortex is that of a hierarchical collection of neurons, with each level of neurons doing local and piecemeal analysis of the input from lower levels. In this hierarchy neurons doing computations related to simple features feed input to neurons further along the visual pathway doing computations related to more general or complex features.

1.4. Connectionist Models for Shape Recognition

Neurological experiments have shown that the typical response time of sensory organs is approximately one hundred milliseconds. The time it takes for a neuron to transmit a signal is on the order of one millisecond. These facts when taken together place an upper limit of a hundred steps for information processing. Some studies have proposed a connectionist model [21] for the visual system which takes into consideration the facts that neurons can only transfer simple signals, and that there exists a hundred-computational-step constraint on information processing in natural systems.

Connectionist models are based on the premise that the responses of neurons depend on the relative strength of their synaptic connections, and that complex information is represented by appropriate connections of a large number of similar units rather than by transmission of complex information through a single connection. Some studies, have presented theories of high, intermediate, and low level vision based on the connectionist paradigm [21,22]. For a detailed discussion of the properties of the connectionist model see [23].

Taking the organization of the visual cortex into consideration, and using the connectionist paradigm, one study has developed a scheme for the recognition of Origami objects [24]. This study addresses the question of building a hierarchical description of shapes from analytical features that are detected using Hough transforms [25]. In that study, Hough transforms are used to detect simple features, i.e. edges, in the input frame. These simple features in turn vote for the occurrence of possible corners in the viewframe. Using these corners, the procedure tries to hypothesize possible polygonal faces in the input image, which are later used to build a shape description. At each level of this hierarchy, a generalized Hough transform scheme is used to detect the possible occurrence of features. Simple features found at lower levels in the hierarchy, increase the confidence of complex features at higher levels.

1.5. Chord Distributions For Recognizing Shapes

Complete chord sets selected from the interior of shapes, have also been used for recognition [26]. In this method, distributions of various chord measures, such as length and orientation are used as shape features. A more detailed discussion of this method will be given in Chapter Two, while the relevant mathematical foundations can be found in the area of integral geometry [27].

1.6. Proposal for a New System for Recognizing Shapes

Analysis of some of the existing shape recognition systems indicates that no single method exists that has solved all problems related to shape recognition. Techniques like Fourier descriptors, and moments tend to give a global description of shape, with local changes in the boundary encoded in the global description. Consequently, these techniques are sensitive to occlusion. Template matching relies on raster scans of the viewframe, and hence is sensitive to orientation of the object in the viewframe. It can be argued that, partly due to the complexity of the shape recognition task, and partly due to the fact that existing shape recognition methods look at a small number of shape features, which can turn out to be insufficient in cases where assumed constraints of the method are relaxed, no general solution to the shape recognition has been found.

Biological evidence suggests that nature has provided not only special purpose hardware for solving the problem of shape recognition, but has employed a local, hierarchical, and piecemeal analysis of the visual input. This evidence, in turn, has suggested a connectionist approach to the shape recognition task. In the connectionist models, the Hough transform has become a primary computational paradigm.

This thesis specifies a model for image plane rotation and translation invariant shape recognition using Hough transforms on features derived from the concepts from integral geometry [27]. A system, referred to as Athena, has been implemented, to incorporate this model. The method is based on sampling random chords from the greylevel image of an object and organizing a shape description to include the intensity functions associated with these chords. This intensity information is termed the I-function of a chord. Athena combines structural information of random chords, e.g., the length, etc., with texture information associated with the I-functions to develop a shape recognition procedure that is invariant under image plane rotation and translation.

The next chapter describes Hough transforms, and the chord length distribution method which provide the computational framework for this model. A possible connectionist implementation for this method will also be discussed.

CHAPTER 2

BACKGROUND

This chapter will discuss a generalized Hough transform [25], and a normalized chord length distribution method [26]. Hough transforms are insensitive to certain degrees of occlusion and noise in the input image, hence they provide a robust basis for shape recognition algorithms applicable in general situations. The chord length distribution method, which is based on the concepts of integral geometry [27], provides a good mathematical framework to deal with the task of shape recognition.

2.1. The Hough Transform

Hough transforms [25,28] have been used in a wide variety of problems in shape recognition that include recognition of printed characters [29], image data compression [30], detection of storage tanks in aerial images [31], and detection of ribs in chest X-ray images [32]. One study has established the equivalence of analytic Hough transforms to template matching while indicating the computational savings possible with the Hough transforms [33]. Another study [34] deals with problems with the Hough transform procedure, e.g., the scatter of peaks in the Hough transform space, etc., and suggests strategies for reducing the problems due to the discretization of the parameter space required in the process.

2.1.1. Hough Transforms for Detecting Analytic Shapes

The Hough transform for detecting analytic shapes, as described in [35], forms parameter spaces from the defining parameters of analytically describable shapes. For example, an algorithm for detecting lines in the input image starts with quantizing the ρ - θ space between appropriate minimum and maximum values. An accumulator array is formed that is indexed by the possible pairs, (ρ,θ). A line at a distance ρ from the origin of the image plane, and at an angle $\theta+\pi$, under this scheme, gets transformed to a point (ρ,θ) in the parameter space, and the corresponding value in the accumulator is incremented. After this processes has been repeated for all points in the input image, the index of the accumulator element having the maximum value corresponds to the parameters specifying the line that best fits the set of points in the image input.

Note that in order to detect any parametric curve, e.g., lines, circles, ellipses, etc., a parameter space needs to be formed that corresponds to the space of parameters of the given curve. For example, in case of a circle, which has the general equation :

$$(x-a)^2+(y-b)^2=r^2$$

an accumulator array needs to be constructed that is indexed by possible values of (a,b,r). It can be seen that the size of the parameter space grows exponentially in the number of parameters describing a shape. Moreover, since most shapes cannot be described analytically, this procedure needs to be generalized.

2.1.2. Generalization of the Hough Transform

A generalization of the Hough transform is from a parameterized analytical shape descriptor to a list structure description of a general shape. The generalization is achieved by picking a reference point for the silhouette of the object and for each boundary point of the shape calculating the tangent direction at that boundary point. The location of the selected reference point is determined relative to the position and tangent direction of each boundary point and a table is constructed containing the relative location vectors of this reference point. The table is organized by tangent directions, so that the vectors from boundary points with the same tangent direction are grouped into one entry. For example, if the points, $p_1, p_2, \ldots, p_n$, have the tangent direction, ϕ, then the table entry for ϕ will have the list of corresponding vectors, $v_1, v_2, \ldots, v_n$, to the reference point.

This table is a generalized template suitable for Hough transform shape detection. The gradient information in the unknown viewframe provides tangent directions which index into the generalized template and vote for possible reference points in the viewframe via a Hough transform scheme. The generalized Hough transform algorithm can be given as [36] :

Algorithm 2.1 : The Generalized Hough Transform.

Step 0. Make a generalized template of the shape to be located by selecting a reference point in the shape, and constructing a voting vector from each boundary point to this reference point. Organize the generalized template indexed by the gradient direction of points on the boundary.

Step 1. Form an accumulator array of possible reference point locations initialized to zero.

Step 2. For each edge point in the viewframe do the following :

(1) Compute the gradient direction, ϕ, at the point.

(2) For each table entry for ϕ, calculate the hypothesized locations for the reference point, and increment the contents of the accumulator array at each location.

Step 3. Possible locations for the shape are given by maxima in the accumulator array.

A key problem in shape recognition arises from the fact that the frame of reference of an object in an unknown view is different from the frame of reference for the representations in the system library. Given a library shape description and an unknown view of a shape, the Hough transform tries to find if the current view supports a particular transformation of a library object description.

Several extensions have been proposed to the generalized Hough transform algorithm [37,38] that deal with possible strategies for incrementing the accumulator array and using more complex features as voting elements. One study has proposed a hierarchical generalization of the Hough transform [37]. In that study, a hierarchy in feature space is specified by a grammar. A composite shape in the image is detected by detecting individual components of the shape, and grouping these components together, using the shape grammar, to form a composite shape. This study also proposes the use of line segments as features in a generalized Hough transform, instead of points in the input image.

In another study, see [38], a generalized Hough transform is used with several boundary points grouped together as a single feature. That study introduces, for object features, the concept of "saliency" a measure of the extent to which a boundary segment distinguishs the object to which it belongs from other objects present in the library.

2.2. Chord Distribution Methods

Techniques based on integral geometry have also been used to characterize shapes [39]. For a shape with n boundary points, $\frac{n(n-1)}{2}$ chords can be drawn by joining each point with the remaining $(n-1)$ points. An empirical chord length distribution can be calculated and used as a rotation and translation invariant feature of a shape. One study [40] has looked at the statistics of histograms of chord lengths, and angles. Peaks in these histograms have been used to indicate the presence of structure in a shape. For a discussion of the advantages and disadvantages of the chord distribution methods, and a comparison of its performance to techniques such as the Fourier descriptor method [7,8] and the moment invariant method [14], see [41].

It has been shown [26] that using a normalized chord length distribution for shape matching has the advantage of producing a measure which is invariant to scale, translation, and rotation of shapes, and is stable with respect to noise and distortion in the shape boundary. However, the method is not information preserving nor a unique measure of shape [41]. This implies that given a chord length distribution of a shape, it is not possible to reconstruct that shape. Moreover, two completely different shapes can have the same chord length distributions. The chord length distribution method described in [41] has a computational complexity of $O(n^2 \log n)$ for a shape of n boundary points.

The chord length distribution method suggests a nonraster scan approach to obtain a representative sample of the shape. When raster scans of unknown shapes are analyzed to characterize the shapes, the result tends to be a scale, and rotation sensitive measure. The chord length distribution method has demonstrated that a measure of the complete set of chords is a reliable, although not unique, identifier of shapes. The reliability of this method results from utilizing the redundant information in a complete set of chords.

The above considerations naturally suggest the basis of Athena, which takes advantage of the redundant information in a nonraster scan and utilizes the additional information from the I-functions of the sampled set to supplement the recognizer. Various properties of random chord samplings are discussed in [42]. Athena will use a sampling composed of randomly oriented chords, with the intent of providing the rotational invariance that is obtained by taking a complete sample of all possible chords.

2.3. A Possible Connectionist Implementation of Athena

Chapter Three describes the overall design of Athena. However, it should be noted that this system can also be implemented using a connectionist approach. In a connectionist system, Athena will have units that respond to various features of chords present in the viewframe. These features will be organized in a hirearchical manner to detect individual chords. At the base level of this hirearchy, will be individual units responding to intensity variations between adjacent points along the I-functions of chords. At the same level in this hirearchy, there will be units that respond to chords of a particular length and chords of a particular orientation. A chord recognition unit can be built using input from length, orientation, and greylevel variation detection units. In this case the I-function information will be provided by the greylevel units, while the structural information will be provided by the length and orientation detection units.

The chord units, in turn, can provide input to identification units. The identification units will detect the location and orientation of an object in the unknown scene. Each identification unit will recieve input from several chord units, and vote for a particular orientation and location of a given object. Note that a chord unit can provide input to several identification units, as well.

A shape description unit will recieve input from the identification units that detect the given shape in various orientations and locations in the viewframe, and will select that identification unit that is most supported by chords from the viewframe. Hence, each shape description unit will be connected to a network that performs the shape detection task. At the highest level in this network will be identification units that detect the given shape in particular orientations and locations in the viewframe. An identification unit will recieve input from chord units that support the particular orientation and location of the object represented by this identification unit. The chord units, in turn, will be connected to a network of orientation, length, and greylevel variation units.

Recognition can be performed by a unit recieving input from the various shape description units. This unit will select that shape description unit that is most supported by the chords in the unknown view, as the recognized shape.

Hence, using local and piecemeal analysis of the visual input, Athena might be implemented in a connectionist scheme. This problem, however, needs further research.

CHAPTER 3

DESIGN AND IMPLEMENTATION OF ATHENA

3.1. System Design

This section describes the overall design of Athena as depicted in Figure 1. Athena works with the greylevel image of an unknown shape to construct a viewframe description. Athena allows the unknown view to be a rotated and translated version of any of the shapes present in the library. In order to achieve rotational and translational invariance, the shape matching procedure makes a random selection of chords from the interior of the unknown object. The random scan of chords is insensitive to the location and orientation of the object in the image. Random selection of chords implies a low probability for the exact same chord occurring in two separate samplings. Thus Athena does not require exact matches of I-functions for chords from the viewframe to the chords in the shape descriptions.

The effectiveness of the recognition procedure relies on the fact that although exact resampling of a chord is unlikely, nearby chords will be sampled yielding close matches for the I-functions, thus allowing recognition of similar shapes. Two things should be considered here. First, given a complete set of I-functions, it is not known how to reconstruct the image without additional structural information. Second, the random sampler need not form a complete set of chords, i.e. $\frac{n(n-1)}{2}$ chords, in order to recognize the shape, if it is used in conjunction with a Hough transform voting process.

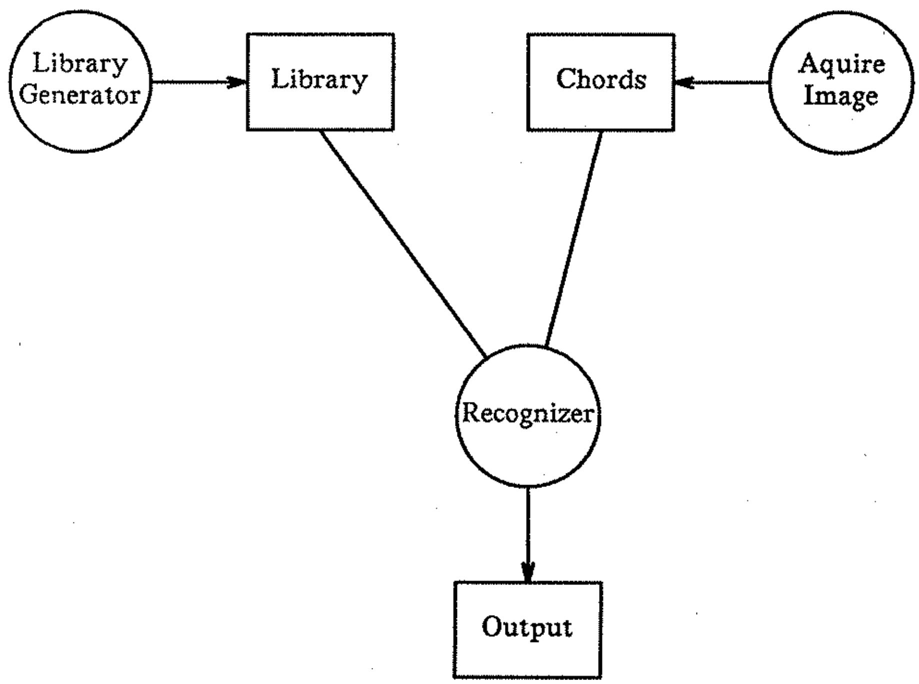


Figure 1. Overall Design of Athena.

A camera is used to form an image of the input scene, transforming the visual data into a 2-D intensity function, $f(x,y)$. This intensity function can be thought of as representing the brightness at all points, (x,y), in the image. The greylevel image is passed to a module which samples random chords from the function, $f(x,y)$. Associated with these chords are two types of information :

a) Structural : Length of chord; start and end points, and orientation; and for a library shape description, a vector to a reference point.

b) Intensity : $f(x,y)$ for each (x,y) of a chord.

The output of the sampler is passed to the recognizer which examines a library of object descriptions and classifies the given view as one object in the library, if possible. The shape descriptions of objects in the library are also made up of random chords drawn from example images of the respective objects.

The GenerateLibrary module of Athena, see Figure 2., constructs a library shape description from randomly scanned chords, using MakeTemplate. MakeTemplate selects a reference point in the shape and creates a vector from the center of each random chord to this reference point.

RecognizeShape sends each object in the library, and the unknown view, to module Compare, that carries out a voting procedure similar to the generalized Hough transform. The pseudocode for module, Compare, can be found in Appendix A.

Module, Identify, selects the object in the library that best matches the unknown scene, by looking at the results of Compare and selecting the object in the library that received the largest number of votes. Module, Identify, outputs the parameters returned by the associated call to Compare. These parameters specify the location and orientation of the unknown object relative to the corresponding object in the library.

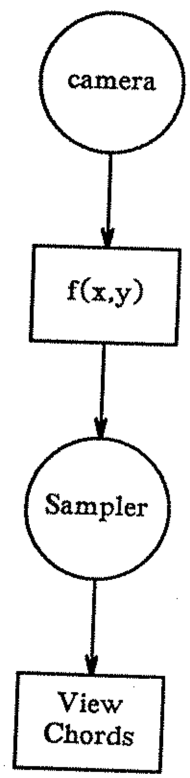


Figure 2.a. AquireImage component of Athena.

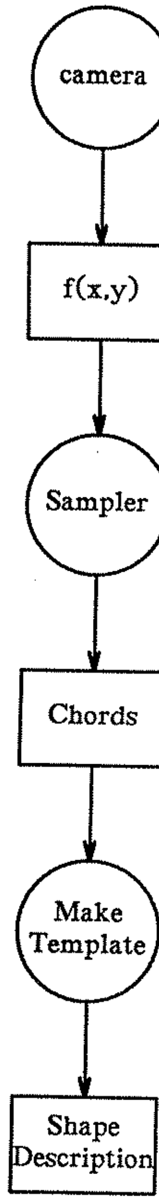


Figure 2.b. GenerateLibrary component of Athena.

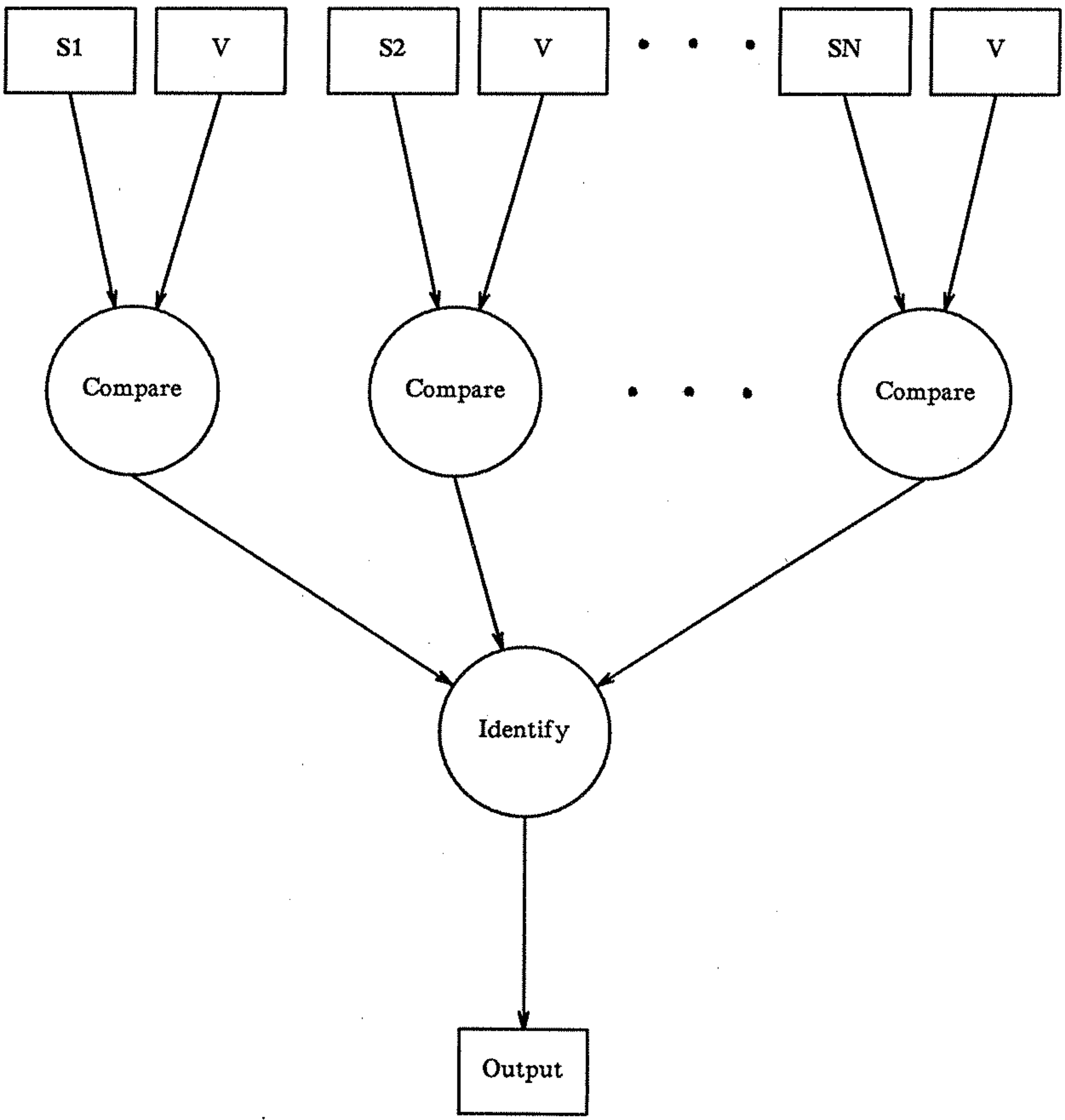


Figure 2.c. RecognizeShape component of Athena.

In order to establish that this technique can detect the location of the object, and its orientation, it has to be shown that :

(1) The peak in the voting array corresponds to the reference point in the image.

(2) Looking at the transformation of chords from the library that match view chords voting for the peak, it is possible to detect the orientation of the object.

3.2. Organization of Shape Description and Chord Matching

The efficiency of the overall system depends upon the computations required to determine the number of chords in the shape description that are similar to a chord in the viewframe. Hence, while organizing the shape description, care should be taken to reduce the number of comparisons of chords to determine similarity.

Let S and V be the set of chords in the shape description and the viewframe, respectively. For each chord, v in V, Athena must find each chord, s in S, such that s and v have matching I-functions. The computational complexity of this search results from the number of times we have to perform the match operation, for each chord in the viewframe. Clearly, in the worst case each chord v in V will be compared with all chords in S, which would mean that the match operation would be performed $|V| \times |S|$ times. Normally there are 300 to 400 chords in the shape description, which make the matching procedure using the above mentioned algorithm very expensive. This suggests that the library shape descriptions need to be organized to reduce the number of chords from the shape description that are compared to each chord in the viewframe.

Athena can partition S according to different criteria of the I-functions associated with the chords in S. Two chords fall in the same partition if they evaluate to be the same under the given criterion. For example, if the mean greylevel is chosen as the partitioning criterion, then all chords that have similar, i.e., absolute difference is less than a preset threshold, mean greylevels for their I-functions will be in the same partition, thereby reducing the size of the search space for chord, v, of V. If a chord, v, of V is given to the matching module, the chord will only be matched to those chords in S that have similar mean greylevels for their I-functions. Hence, the shape description will be an organized set of chords indexed by mean greylevel.

Athena matches chords according to the textural constraints in the I-functions. For two I-functions to match, however, the chord lengths must be similar. Athena reduces the number of full I-function matches that are attempted, and fail, by matching the length first. The reader is reminded that the chord length distribution itself has been shown [26] to discriminate shapes. However, even with a complete chord sampling, the chord length distribution is not a unique identifier of shape. Athena adds discriminatory power through the use of the textural constraints.

A single chord from the viewframe will index into a subset that can contain many chords with similar mean greylevels, but it is not necessary that the view chord will match all chords in this subset. Figure 3a. shows two I-functions that have similar mean greylevels, but the I-functions are sufficiently different that these chords will not match. In Figure 3b. there are two I-functions that not only have similar mean greylevels, but the greylevel variations are similar, hence they match. The situation in 3c. is different in that the two I-functions have similar mean greylevels, while the greylevel variations are reflections of one another. This reflection property might occur from sampling the same chord starting at the opposite end. In other words, the second sampling is different by a 180 degree rotation about the center of the chord. Since our recognition process is to be

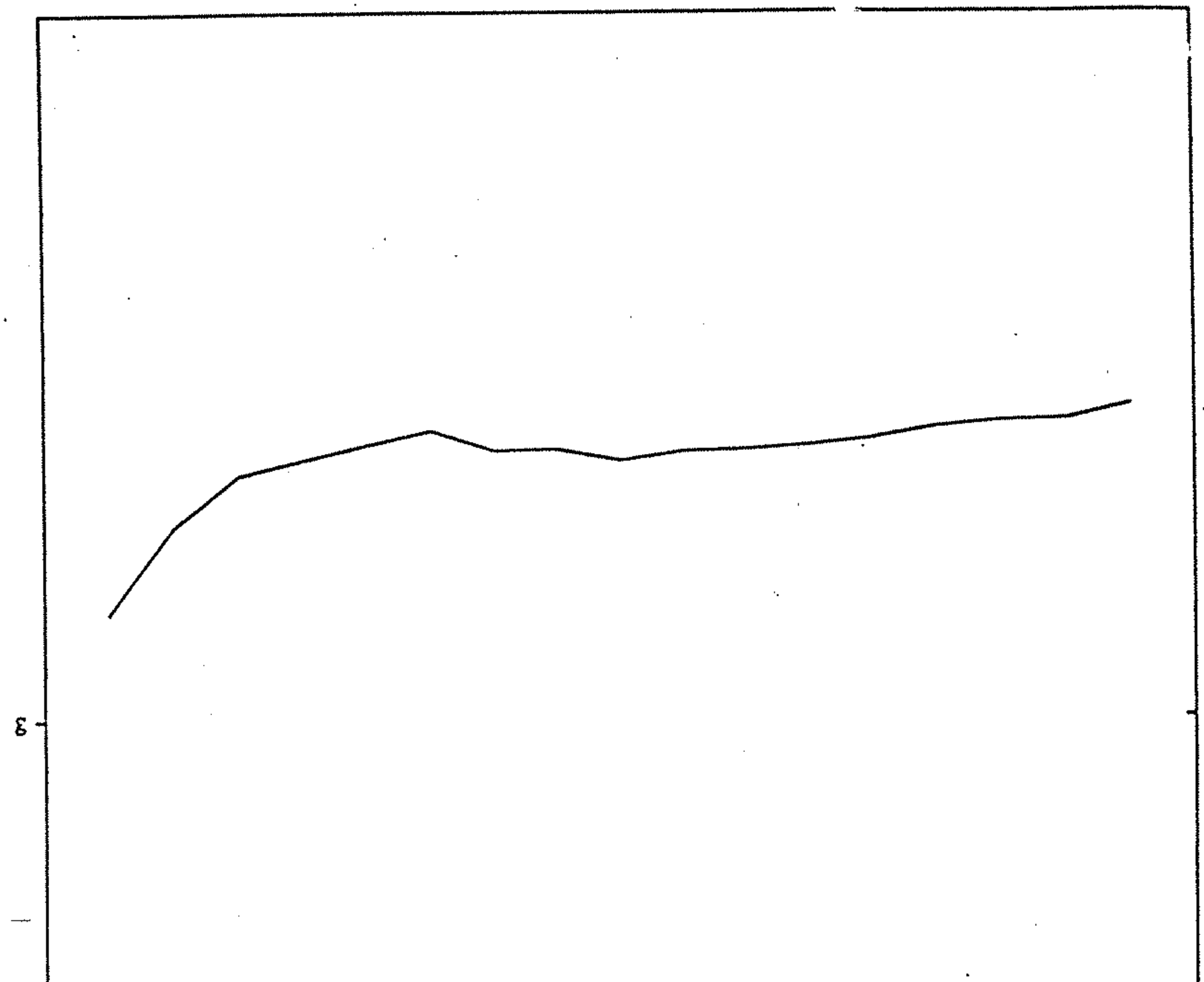


Figure 3.a.2. Non-matching I-function with mean similar to the chord shown in Figure 3.a.1.

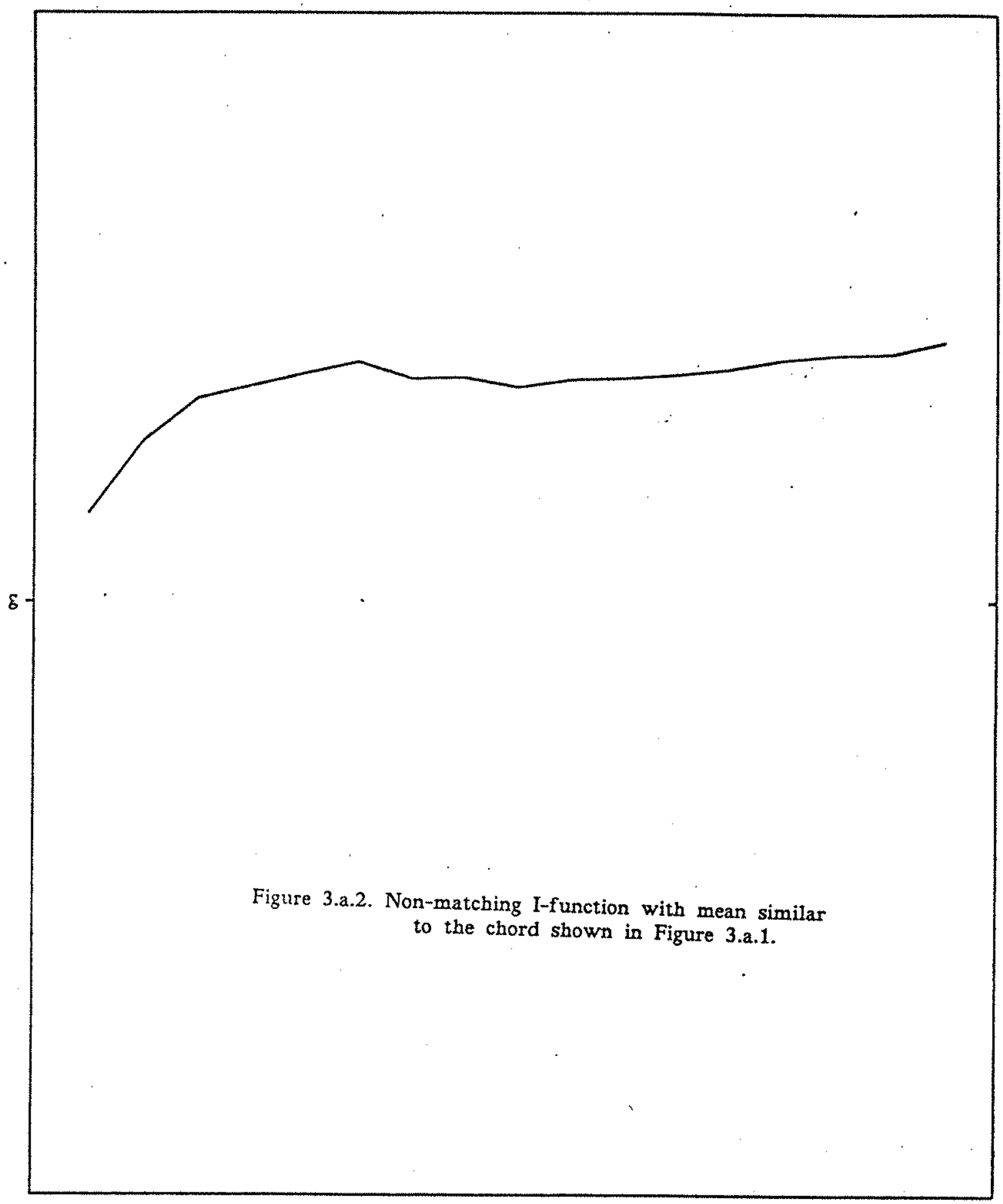


Figure 3.a.2. Non-matching I-function with mean similar to the chord shown in Figure 3.a.1.

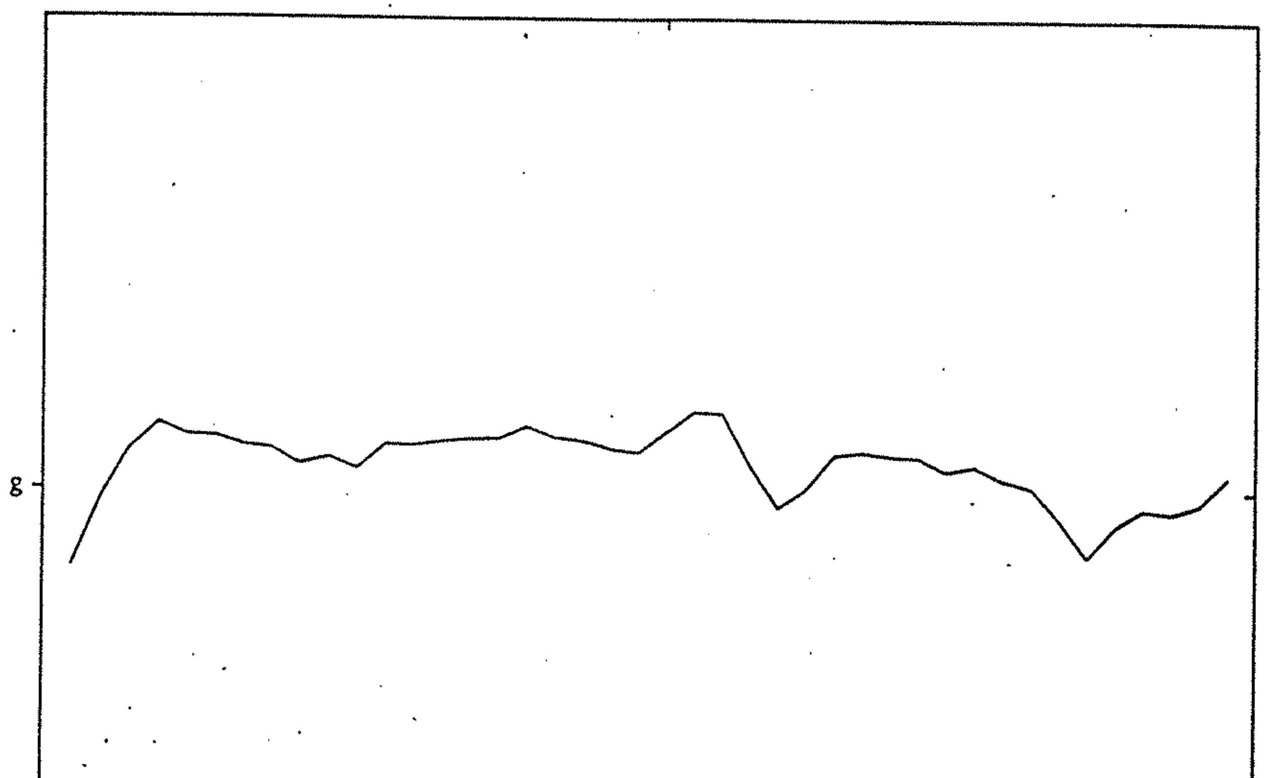


Figure 3.b.1. I-function of an example chord.

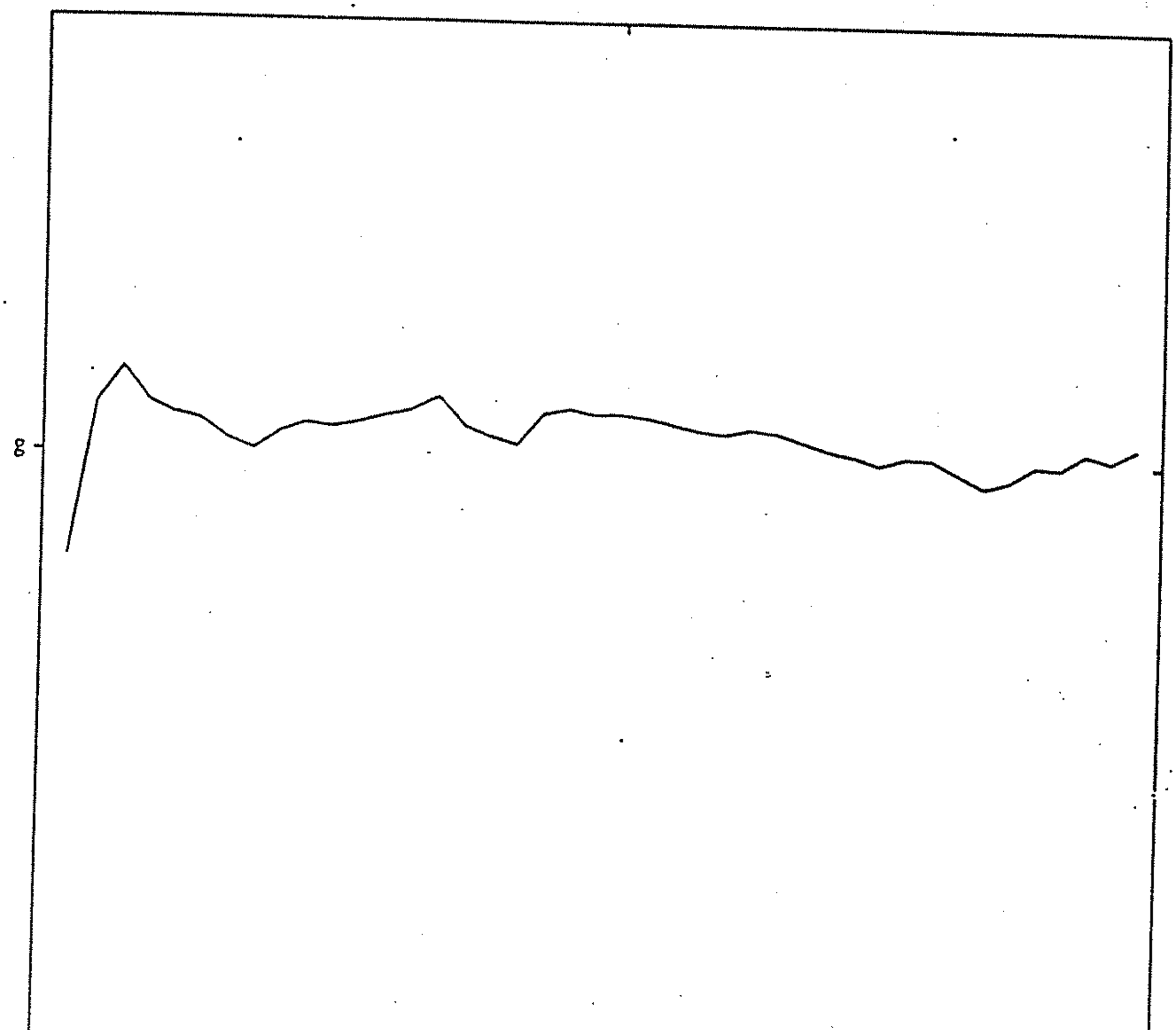


Figure 3.b.2. Matching I-function of the chord shown in Figure 3.b.1.

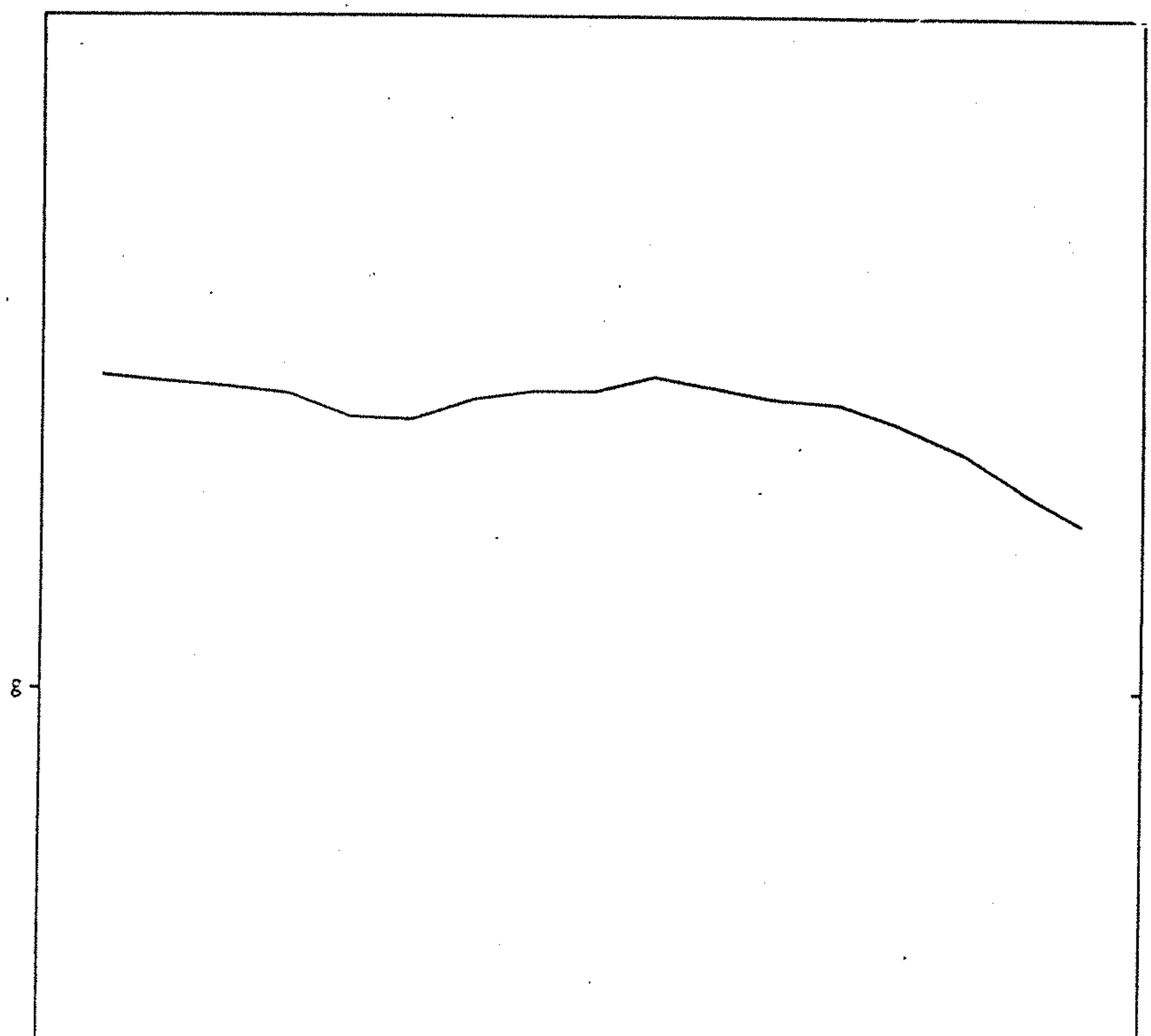


Figure 3.c.1. I-function of an example chord.

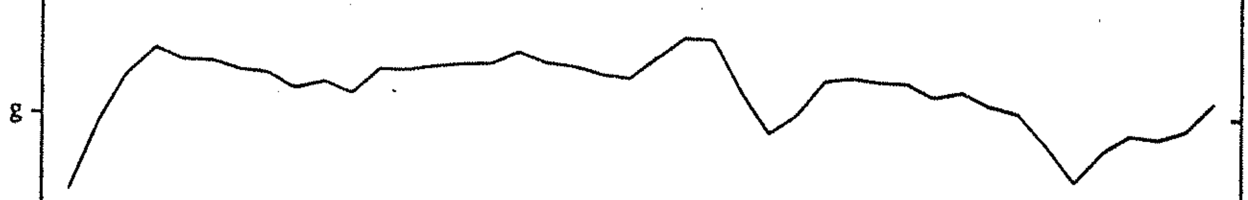


Figure 3.c.2. I-function that matches under reflection to the chord shown in Figure 3.c.1.

rotation invariant, the two I-functions in Figure 3c. will be declared to match. Athena declares two chords to match if the average value of the absolute difference between the I-functions of the two chords fall within a preset threshold.

3.3. Reference Point Hough Transform

Once a chord, v, in V matches a chord, s, in S Athena uses the reference vector of s to construct a voting vector for v. This operation can be thought of as placing s in the viewframe, and translating s to have the same starting point as v. Now s is rotated about its starting point to have same orientation as v. The reference vector of s, transformed along with s, now will point to a hypothesized reference point in the viewframe and the accumulator element for that point will be incremented by one. This procedure is

repeated for all chords in V, and their votes are collected in an accumulator array that represents the viewframe. The absolute maximum in this array corresponds to the reference point in the viewframe most supported by the unknown shape. Note that a chord in the viewframe can match several chords in the shape description, and hence it can vote for many hypothesized reference points. This voting mechanism is similar to that of the generalized Hough transform.

Note that it is possible to locate the reference point in the unknown scene if exactly the same sampling of chords from the shape can be obtained from the viewframe as in the library. In this case, after Hough voting has been carried out, the peak in the voting array will correspond to the location of the reference point in the viewframe. This is because each chord in the viewframe will match exactly at least one chord in the library. Moreover, all chords in the viewframe will be at the same orientation, with respect to one another, as the chords in the corresponding shape description. Hence, each voting vector constructed for a correctly matched view chord will point to the accumulator array element that maps to the original reference point. For each of those vectors the accumulator array element will be incremented yielding a peak at the location of the reference point. However, it is possible that a chord in the viewframe may match more than one chord in the object description. In this case, there will be other votes in the accumulator besides the peak. These extraneous votes will not form significant peaks for general shapes having general textures.

Throughout this section the term chord set is used to mean chords from the shape descriptions in the library, and the unknown view. Unless indicated otherwise, these chord sets will have 200 chords. Figure 4. shows the results of one such experiment. The chord set labeled A in Figure 4. is from an object description in the library. In this case, the unknown view chords are exactly the same chords as the library description, A. Athena correctly identified the unknown view as an instance of the object in the library. The peak found in the voting array was exactly at the same location as the reference point of the object description A.

In actual situations, however, the exact location and orientation of the object is not known. Athena, therefore, takes a random sampling of chords from the viewframe to recognize the object. Figure 5. shows a chord set labeled A that will be matched as the library entry against the unknown view chord set labeled B. Chord set B is a different random scan of the same image from which chord set A was obtained. In this experiment both the library descriptions and the unknown view consisted of 300 chords. The reference point for the object, A, in Figure 5. was analytically determined to correspond to the point (27,32) in the voting array. After the voting was carried out, a significant peak is indeed evident in the accumulator array. Table 1. gives the location of all voting array elements that

Figure 4. Chords from object description of A.

TABLE 1

Number of chords in the library description : 300.

Number of chords in the unknown view : 300.

Coordinates	Votes	% of 300
(26,32)	19	6.33
(26,33)	24	8.00
(27,32)	24	8.00
(27,33)	15	5.00

were more than half the magnitude of the largest peak in the array.

It should be noted that two maxima occur, but they are adjacent and one of them is exactly at the analytically determined position in the voting array, while other locations of high value are very close. Hence this experiment supports the claim that random sampling of chords along with the matching and voting procedure described in this chapter can locate an object in an unknown scene.

It should be noted that the value of the highest peak in this table is 8.0 percent of the total chords present in the unknown scene. Moreover, all the significant peaks in the voting array are very close to the analytically determined peak. This slight scatter can be explained by the fact that the random sampling of chords is very unlikely to get exactly the same chord in two different random samplings of the same object. However, it is possible that chords that are spatially close, and hence very similar in greylevel intensity variations, are sampled. The votes in the voting array, therefore, will be scattered slightly. Figure 6. shows a plot of the contents of the voting array after the comparison of the chord set B with the chord set A shown in Figure 5.

3.4. Detecting Object Orientation

Once the location of the object has been found in an unknown scene by selecting the highest peak in the voting array, Athena performs a second Hough transform using only those chords that voted at the peak location in the reference point Hough transform, to determine the orientation of the object in the unknown scene. The rotation each object chord had to go through in order to be parallel to the view chord it matched, is counted in another accumulator array for possible object rotations. The absolute maximum in this array of possible rotations corresponds to the most likely rotation the object description has to go through to be in the same orientation as the object in the unknown scene.

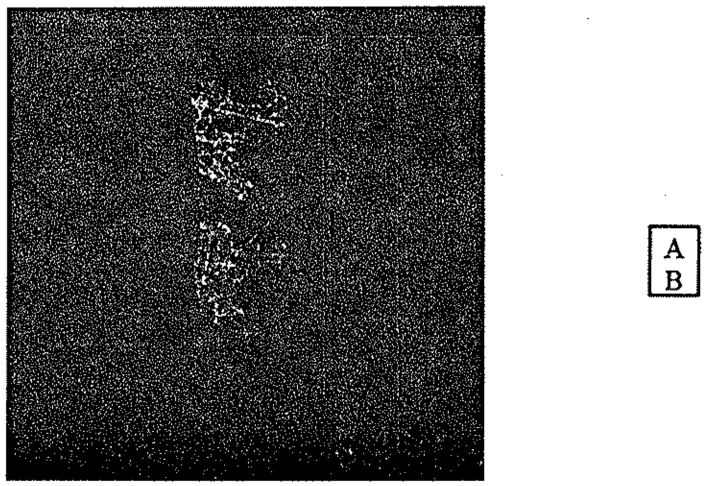


Figure 5. A : Chords from the object description.

B : Chords from the unknown view.

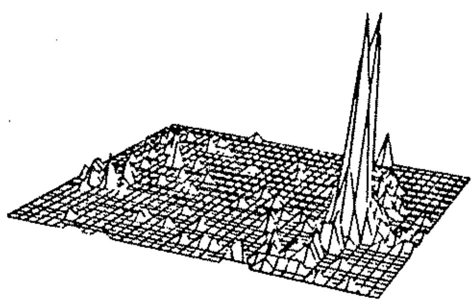


Figure 6. Contents of the voting array for Figure 5.

Figure 7 shows two instances of the same object labeled S1 and S2. S2 is a rotated version of S1 about point O. Let this rotation be, θ. Let A and B be two points on the boundary of S1. Let A' and B' be the corresponding points to A and B, respectively, on S2. Let the the coordinates of A, B, A', B' be (x_1,y_1), (x_2,y_2), (x_3,y_3) and (x_4,y_4), respectively. Let α be the angle AB makes with the horizontal, and β be the angle chord $A'B'$ makes with the horizontal. This implies that :

$$\tan(\alpha) = \frac{(y_2-y_1)}{(x_2-x_1)} \tag{3.1}$$

$$\tan(\beta) = \frac{(y_4-y_3)}{(x_4-x_3)} \tag{3.2}$$

Since A' and B' correspond to points A and B rotated by an angle θ about O :

$$x_3 = x_1\cos(\theta) + y_1\sin(\theta) \tag{3.3}$$

$$y_3 = x_1\sin(\theta) + y_1\cos(\theta) \tag{3.4}$$

Similarly,

$$x_4 = x_2\cos(\theta) - y_2\sin(\theta) \tag{3.5}$$

$$y_4 = x_2\sin(\theta) + y_2\cos(\theta) \tag{3.6}$$

Hence $\tan(\beta)$ can be written as :

$$\tan(\beta) = \frac{(y_2-y_1)\cos(\theta) + (x_2-x_1)\sin(\theta)}{(x_2-x_1)\cos(\theta) - (y_2-y_1)\sin(\theta)} \tag{3.7}$$

But,

$$\tan(\alpha+\beta) = \frac{\tan(\alpha) + \tan(\beta)}{1 - \tan(\alpha)\tan(\beta)} = \frac{(y_2-y_1)\cos(\theta) + (x_2-x_1)\sin(\theta)}{(x_2-x_1)\cos(\theta) - (y_2-y_1)\sin(\theta)} \tag{3.8}$$

Therefore, using (3.7) and (3.8) we get :

$$\tan(\alpha+\theta) = \tan(\beta) \tag{3.9}$$

This, along with the assumption that the object has rotated θ, means that :

$$\theta = \beta - \alpha \tag{3.10}$$

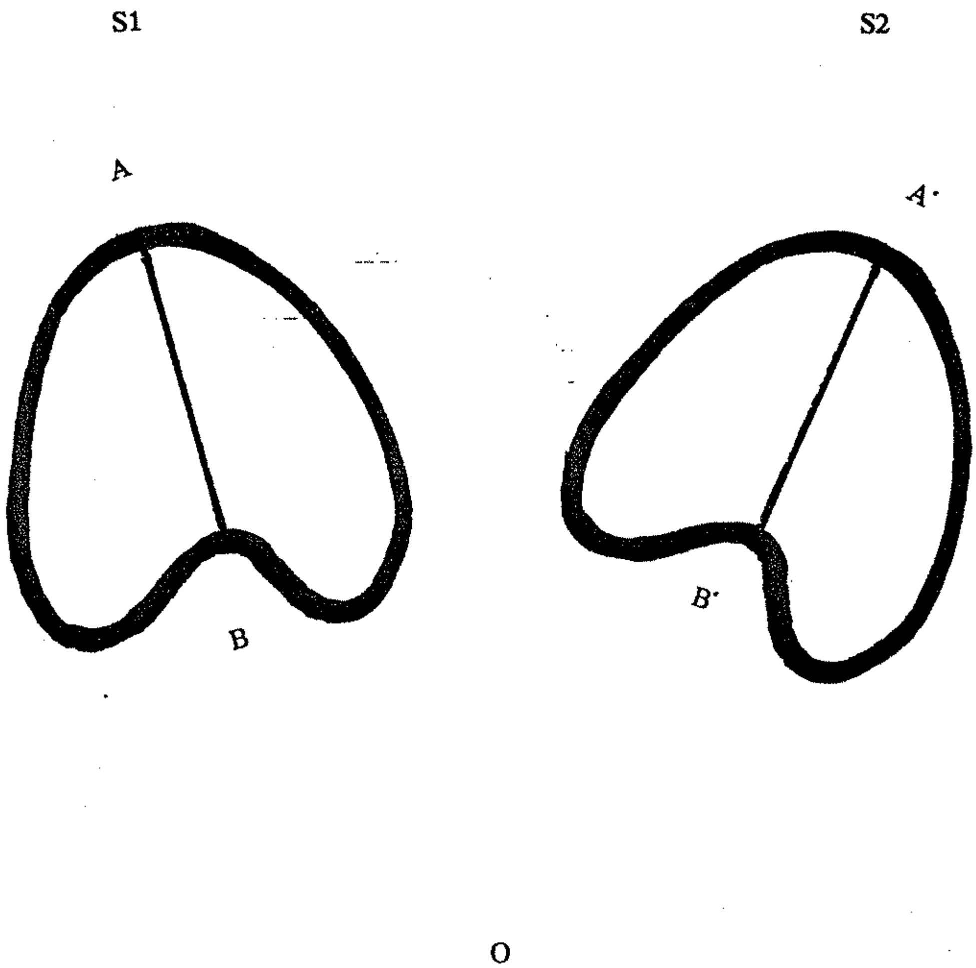
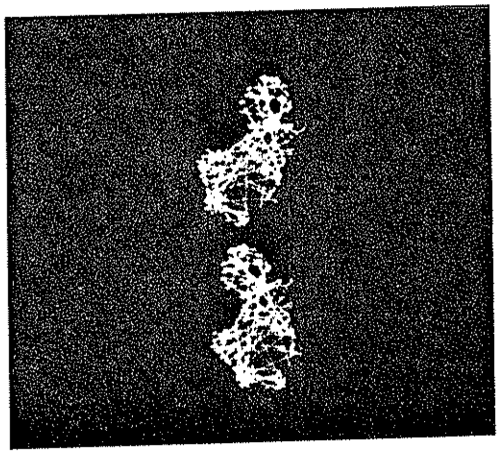


Figure 7. Physically rotated object.

This obvious result can also be concluded from the fact that the object S2 is a rigid body rotation of object S1 about point O. Hence, the relative position of all points on the boundary of S1 with respect to one another remain the same. This means that all points in the object go through the same rotation, θ, about point O. Therefore, if the angle of a chord, AB, was α before the rotation, the angle will be $\alpha+\theta$ after it. If we accumulate a count of the rotation, θ_i, each chord went through in an array, then the maximum in this array will correspond to the rotation of the object description.

To illustrate this process consider an example in which the view chord set is an analytically rotated version of the library description of an object. The chords in the viewframe are rotated by -25 degrees about the origin. Indeed the voting yields a peak in the accumulator for possible rotations exactly at -25 degrees. Figure 8. shows the results of this experiment. The chord set labeled A is from an object description in the library. The chord set labeled B is from the analytically rotated viewframe.

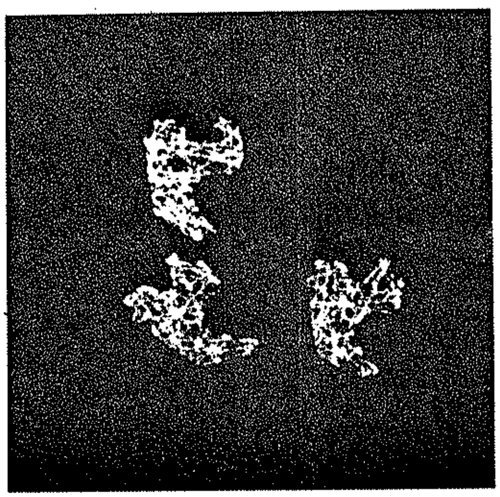
Figure 9. shows the results of an experiment in which the unknown view, labeled B, consists of an analytically rotated version of a rescan of object A. That is, the same image from which A was sampled is resampled with a different selection of chords and the resulting chords are analytically rotated by -45 degrees about the origin to yield B. Athena indeed finds a significant peak in the reference point Hough transform and the chords at that peak indeed indicate that a rotation of -45 degrees about the origin would be needed to bring B into the same orientation as A. In Figure 9., the shape labeled C shows chord set B inversely rotated by that angle. These experiments demonstrate that, for analytically transformed examples, Athena can correctly identify the orientation of an object in the unknown viewframe, after the object has been recognized.



A
B

Figure 8. A : Chords from the object description.

B : Chords from analytically rotated viewframe.



A
B C

Figure 9. A : Chords from the object description.

B : Chords from analytically rotated rescan of A.

In order to see if physically rotating the object in the unknown scene would produce the same results, Athena was presented with an arbitrarily placed object in the viewframe. The results of this experiment are shown in Figure 10. The chord set labeled A contains 400 chords from the object description present in the library, while chord set labeled B contains 400 chords from the arbitrarily placed object in the viewframe. Athena correctly identified the object in the viewframe and indicated that the unknown object can be rotated by 26 degrees in order to be in the same orientation as the object in the library. The chord set labeled C in Figure 10. corresponds to a 26 degree rotation of the chords in B about the origin of the viewframe. Visual inspection shows A and C to be in the same orientation. Figure 11. shows the plot of the voting array for this comparison. Table 2. shows those elements of the voting array that are greater than 50 percent of the highest peak in the voting array.

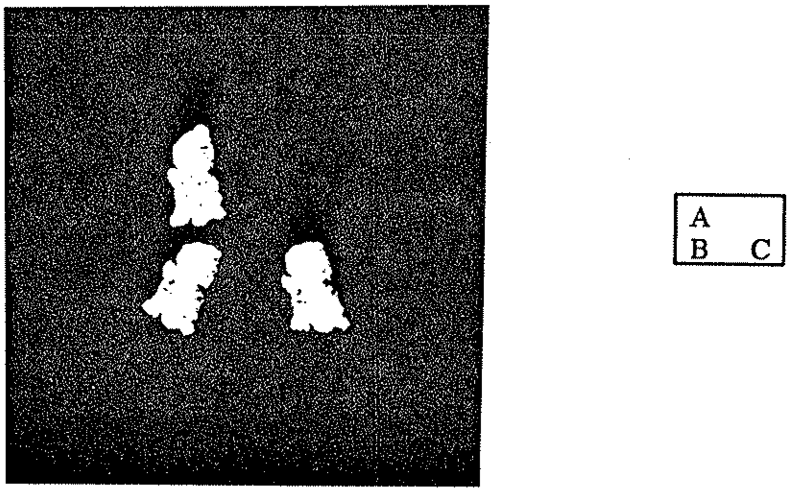


Figure 10. A : Chords from the object description.

B : Chords from a physically rotated version of the object in the library.

C : Analytically adjusted object B to have the same orientation as A.

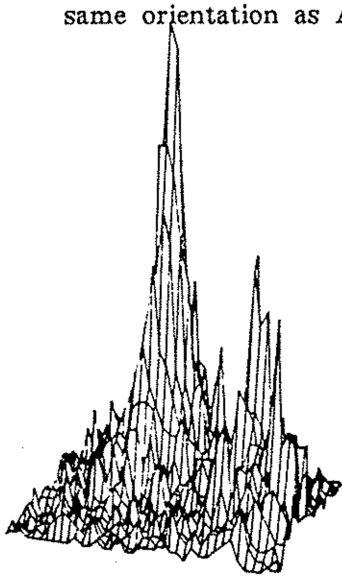


Figure 11. Plot of the voting array for Figure 10.

TABLE 2

Total number of chords in shape description : 400

Total number of chords in unknown view : 400

Coordinates	Votes	% of 400
(17,20)	34	8.5
(19,19)	34	8.5
(20,18)	35	8.7
(20,19)	46	11.5
(21,17)	39	9.7
(21,18)	51	12.7
(21,19)	54	13.5
(21,20)	39	9.7
(22,17)	51	12.7
(22,18)*	68	17.0
(22,19)	65	16.2
(22,20)	38	9.5
(22,21)	34	8.5
(23,17)	53	13.2
(23,18)	53	13.2
(23,19)	49	12.2
(23,20)	40	10.0
(24,17)	36	9.0
(24,18)	44	11.0
(24,19)	40	10.0
(25,17)	37	9.2
(25,31)	39	9.7
(26,31)	37	9.2

CHAPTER 4

EXAMPLES USING ATHENA

Five examples will be discussed in this chapter that were chosen to demonstrate the capabilities of Athena. Figure 12. shows the example data set used in these experiments. The labels on the images in Figure 12. will be used throughout this chapter to refer to the respective chord sets. The original images were either cartoon or human figures. In all experiments described in this chapter the voting array had a dimension of 40 x 40. Randomly sampled chords from the images were restricted to the interior of the object in the viewframe by using an image thresholding scheme. Five examples will be discussed in this chapter that were chosen to demonstrate the capabilities of Athena.

4.1. Progression of Chords

Shown in Figure 13. is a progression of chord samples taken from object, D. The extreme left shape in Figure 13. shows the first 50 chords from D. The remaining shapes to right in Figure 13. are the first 100, 200 and 400 chords, respectively, taken from the same chord set. Note that even with a small sample the structure of the object is evident. As the number of chords in the sample is increased, the chords further cover the interior and complete the boundary of D. It can be seen that the random chord sampling contains redundant information about the boundary and interior of a shape.

4.2. Example 1

This example demonstrates that Athena can correctly identify the unknown shape when the library contains more than one object. Shown in Figure 14., top

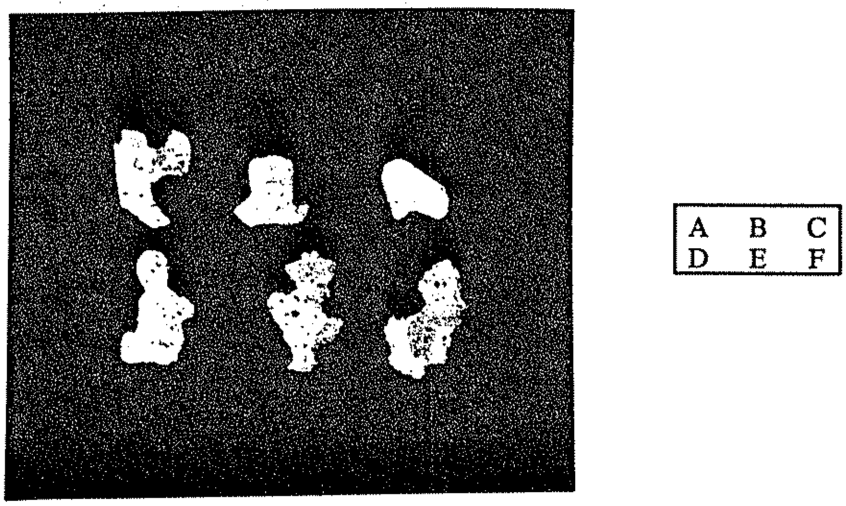


Figure 12. Example data set used in the experiments.

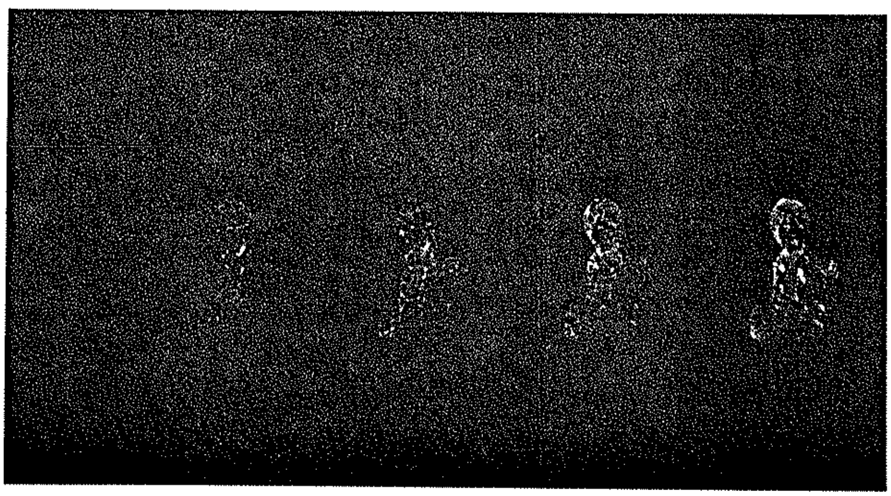


Figure 13. Progression of chords from D.

row, are 200 chords from D, E, F, and A, respectively, from the library, while the bottom row displays 200 chords from the unknown view, D'. The unknown view has been placed under its corresponding library object. The correspondence, in this case, was established by Athena.

The unknown view, D', in Figure 14. was created as a rescan of object, D. Figure 15. shows the plots of the reference point voting array after the unknown view was compared to each object in the library. It can be seen that the peak in Figure 15.d., that corresponds to the plot of the voting array after the comparison of D' with D, is significantly larger than the peaks in Figure 15.a, 15.b, and 15.c. Hence, Athena identifies the unknown view to be an instance of object D. Table 3. shows the size of the highest peaks for the plots shown in Figure 15.

TABLE 3

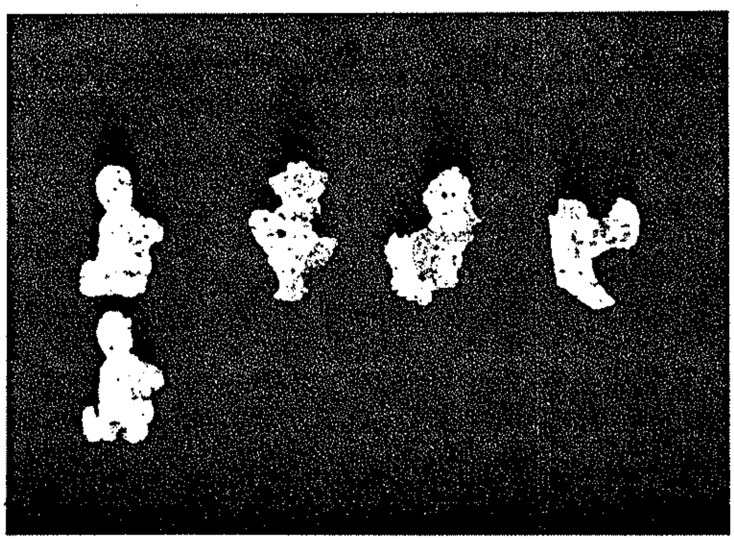
Number of chords in each shape description : 300

Number of chords in the unknown view : 300

Name of Object	Value of Peak	% of 300
D	32	10.02
A	4	1.03
F	13	4.03
E	13	4.03

It should be noted that the sizes of the peaks for the incorrect matches are less than 50 percent of the size of the peak for the correct match.

The reference point for the object description, D, was specified to correspond to the coordinates (37,25) in the voting array. Table 4. shows those elements of the voting array, for D vs. D', that were greater than 50 percent of the highest



D E F A
D'

Figure 14. Setup for example 1.

Figure 15.a. D' vs. A.

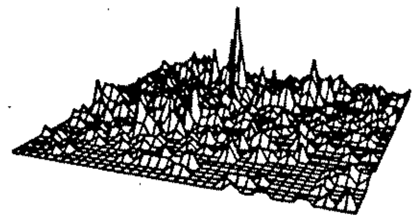


Figure 15.b. D' vs. F.

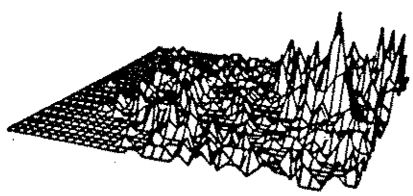


Figure 15.c. D' vs. E.

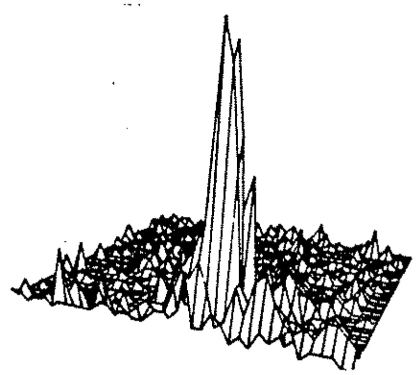


Figure 15.d. D' vs. D.

Figure 15. Contents of the voting array for Figure 14.

peak in that voting array.

TABLE 4

Number of chords in the library : 300

Number of chords in the unknown view : 300

Coordinates	Votes	% of 300
(35,27)	17	5.06
(36,24)	25	8.03
(36,25)	30	10.00
(36,26)	21	7.00
(37,24)*	32	10.06
(37,25)	28	7.03
(37,26)	18	6.00

4.3. Example 2

In this experiment each shape in the library, and the unknown view, consisted of objects of the same texture. The library objects, B and C, and the unknown view, B', consisted of 200 chords each. The shapes labeled, B and C, in Figure 16. are 50 chords each from the library shape description, while the shape labeled, B', is composed of 50 chords from the unknown view, a rescan of object B. Athena correctly identified B' as an instance of B. The highest peak in the voting array for the comparison, B' vs. B, had 33 votes, while the highest peak in the voting array for the comparison, B' vs. C, had 8 votes.

Shown in Figure 17.a is a plot of the contents of the voting array for the comparison, B' vs. B. It can be seen that the peak in this voting array is high, relative to the noise in the array. Figure 17.b. shows a plot of the contents of the voting array for the comparison, B' vs. C. Since the texture of the two shapes was

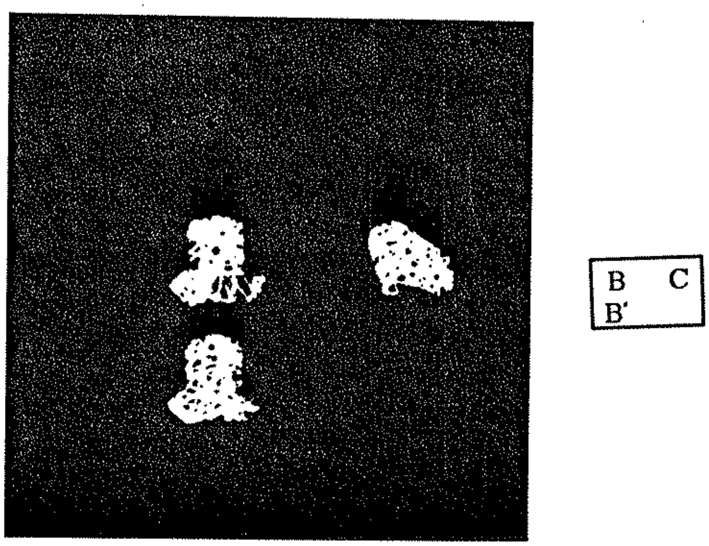


Figure 16. Setup for example 2.

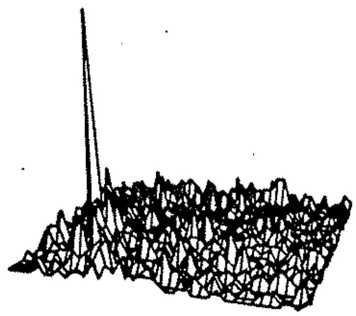


Figure 17.a. Plot of the voting array for B' vs. B.

Figure 17.b. Plot of the voting array for B' vs. C.

exactly the same, the voting array is fairly noisy, indicating that many chords in B' matched chords in C. However, since the two shapes were different, there is no dominant peak. Further examination of the voting array for the comparison, B' vs. C, shows that there are more than fifty elements in this array that had values greater than 50 percent of the highest peak in this array.

4.4. Example 3

In this experiment the library consisted of objects, E and F, while the unknown view, E', was a scan of a physically rotated and translated version of object, E. Each object description in the library had 400 chords in it, while the unknown view was composed of 340 chords. The shapes labeled, E and F, in Figure 18. show 50 chords each from the library shape description and the unknown shape. The shape labeled E'', shows the chords in E' analytically rotated by the negative of the angle determined by Athena, in this case -41 degrees.

Figure 19.a. shows a plot of the contents of the voting array for the comparison, E' vs. E, while Figure 19.b. shows E' vs. F. The highest peak in the voting array for the comparison, E' vs. E, was 46, while for E' vs. F it was 16. Hence, Athena identified the unknown view, E', to be an instance of E.

4.5. Example 4

In this experiment the library consisted of 200 chords from each object, B and C, as shown in the top row of Figure 20. The unknown view, B' (Figure 20. bottom row) had the same texture as C, while its shape was exactly the same as B. The highest peak in the reference point voting array for the comparison, B' vs. C, was 15, while there were no votes in the comparison, B' vs. B. Hence, Athena incorrectly identified B' to be an instance of C. Figure 21. shows the contents of the voting array for the comparison, B' vs. C. Note that the votes in the array

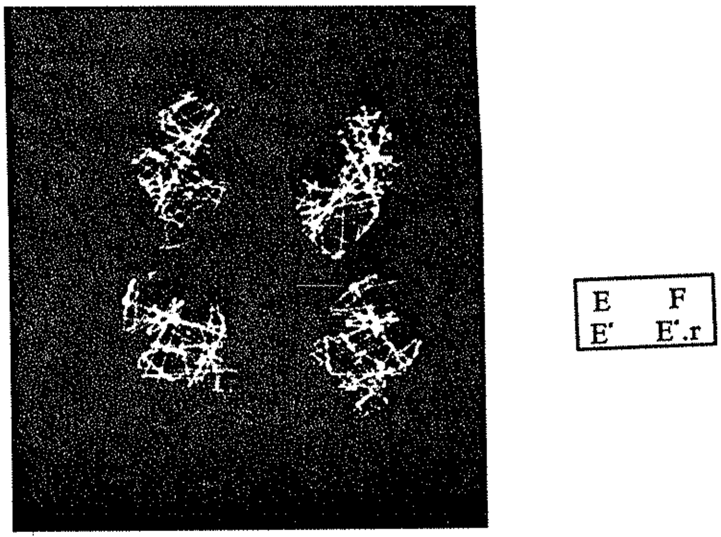


Figure 18. Setup for example 3.

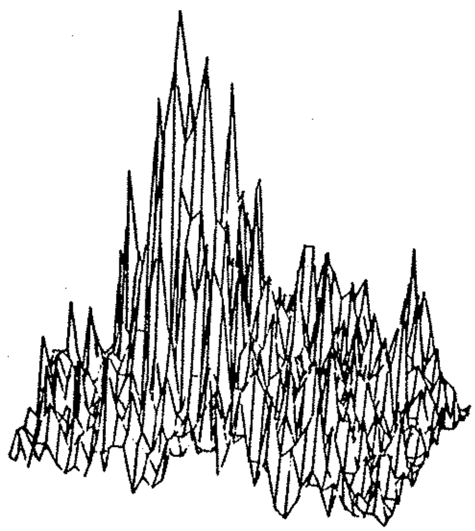


Figure 19.a. Plot for the comparison E' vs. E.

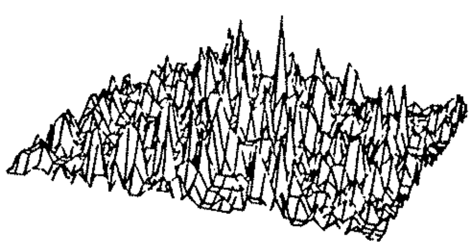
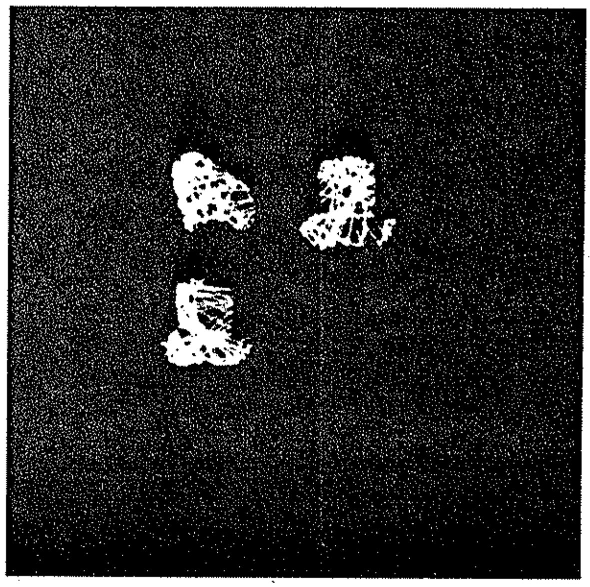


Figure 19.b. Plot for the comparison E' vs. F.



C	B
B'	

Figure 20. Setup for Example 4.

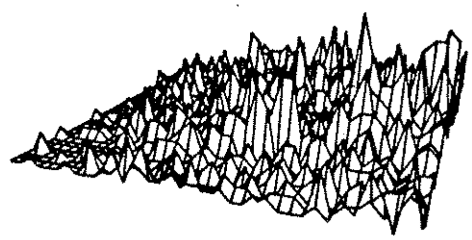


Figure 21. Contents of the voting array for the comparison B' vs. C.

are scattered, and that instead of a single

distinct peak the array contains many similar elements of low magnitude. This example shows the strong influence of the textural features of the random chord samples.

4.6. Example 5

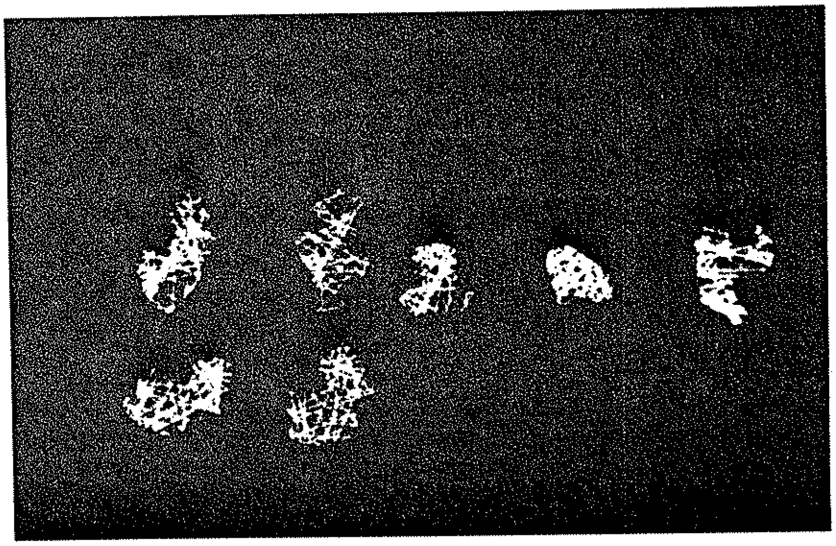
In this experiment the library consists of 400 chords from each of the objects A, B, C, E and F, while the unknown view is a sample of a physically rotated version of the image from which chord set F was obtained. In Figure 22., the top row shows 50 chords from each of the objects in the library. The shape labeled, F', in Figure 22., shows 50 chords from the physically rotated image. The shape labeled, F'.r, in Figure 22. shows 50 chords from the unknown view, analytically rotated, by the negative of the orientation detected by Athena. Figure 23. shows the plots of the contents of the reference point voting array for each of the comparisons carried out by Athena. Table 5. shows the maximum values for each of these reference point voting arrays.

TABLE 5

Number of chords in each object description : 400

Number of chords in the unknown view : 400

Name of Object	Size of Peak	% of 400
E	22	5.50
A	7	1.75
C	28	7.00
B	36	9.00
F	56	14.00



F E B C A
F' F'.r

Figure 22. Setup for example 5.

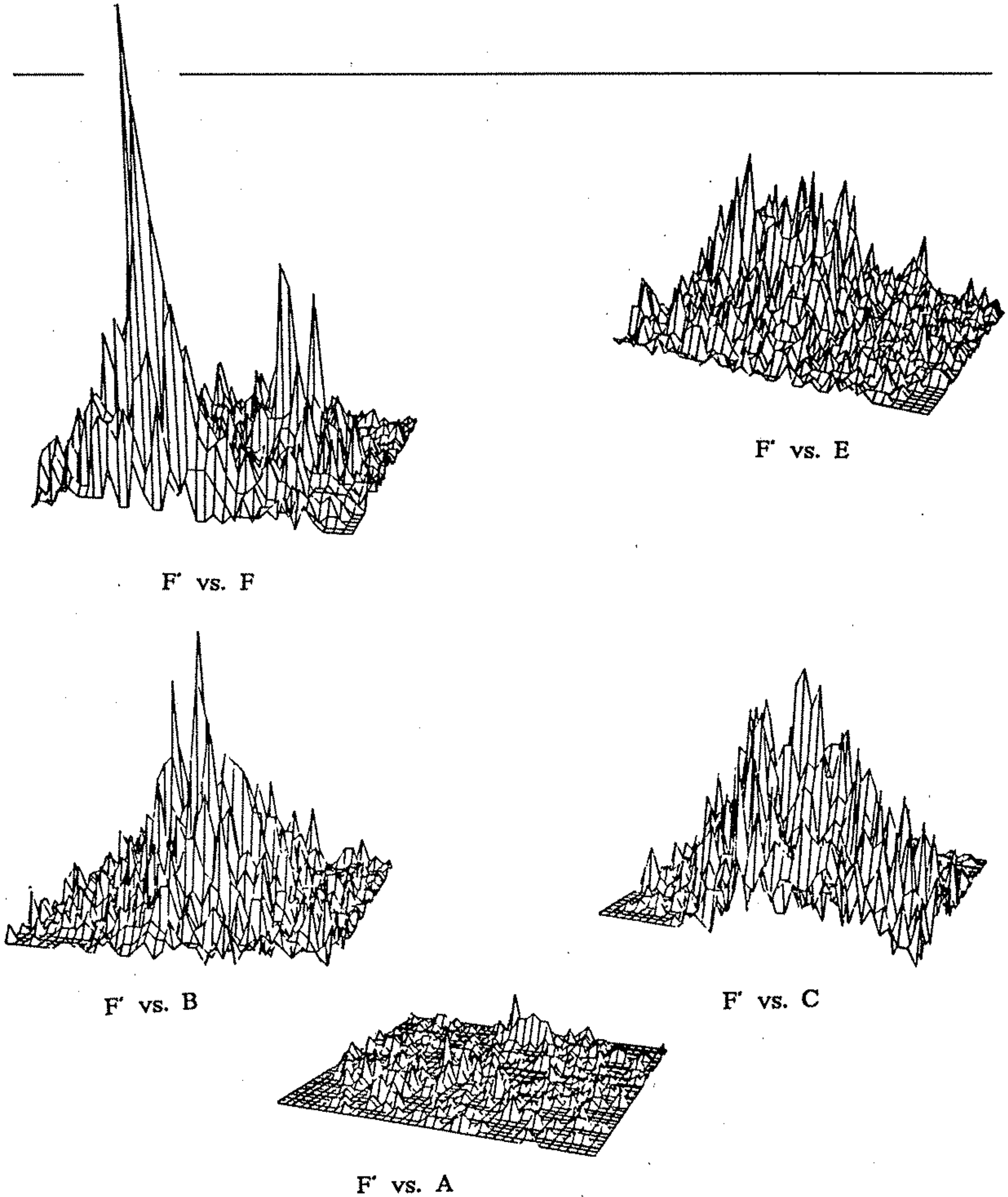


Figure 23. Plots of voting array for example 5.

Table 6. shows peaks in the voting array for the comparison, F' vs F, that were greater than 50% of the highest peak in this comparison. Table 7. shows the contents of voting array for possible orientation values that were greater than 50 percent of the maximum peak found in this array.

TABLE 6

Number of chords in the object description : 400

Number of chords in the viewframe : 400

Coordinates	Size of Peak	% of 400
(34,15)	39	9.90
(35,15)	34	8.50
(35,16)*	56	14.00
(35,17)	33	8.25
(35,18)	28	7.00
(36,16)	33	8.25
(36,17)	47	11.75
(36,18)	35	8.75

TABLE 7

Change in Orientation	Votes
36	4
37	5
38	6
39	3
41	4

Athena detected the orientation angle to be 38 degrees. Visual inspection of Figure 22. indeed confirms that F'.r and F are of the same orientation.

For a more detailed set of examples see Appendix B.

CHAPTER 5

CONCLUSION

5.1. Summary

This study addresses the question of recognizing shapes and has examined a method for two-dimensional shape recognition, which is invariant to image plane rotation and translation. This method utilizes the greylevel information associated with a random sampling of chords from an image of a shape. Schemes for organizing the chord information in order to reduce the computational complexity of the recognition task, thereby increasing the efficiency of the system have been discussed. A system — Athena — has been implemented that is capable of identifying library objects, and determining the location of a reference point and the orientation of the object about that point. Five examples have been discussed that demonstrate the capabilities of Athena.

5.2. Future Research

This study has only looked at a single partitioning criterion for chords. It is possible to have multiple partitions that organize chords under different criteria allowing multiple voting passes for each chord. Such an organization of the shape description will be able to take more features into account during the recognition process.

It is also possible to use pairs of chords in specifying the structural information instead of single chords. If pairs of chords are used, the recognition procedure will have more information about the image than linear chord samplings. The additional structural information, however, will affect adversely the computational complexity of this scheme.

In general, scenes will contain objects that are only partially visible. The occurrence of such occlusion will affect the way the chords are represented and organized to form a shape description. Handling occluded objects will require a matching function that can perform partial matchings of I-functions.

The question of a connectionist implementation of Athena also will be addressed in more detail. A connectionist design for Athena needs to consider important questions regarding organization of the library as a network of shape recognition units.

Finally, a generalized vision system will need to recognize moving objects. Research is under way to modify the design of Athena to allow the system to be applied to dynamic scenes.

REFERENCES

(1) T. Pavilidas, A review of algorithms for shape analysis, Comput. Graphics Image Processing, vol. 7, pp. 243-258, 1978.

(2) T. Pavilidas, Algorithms for shape analysis of contours and waveforms, IEEE Trans. Pattern Analysis and Mach. Intell., vol. PAMI-2, no. 4, pp. 301-312, 1980.

(3) L. G. Shapiro, Recent progress in shape decomposition and analysis, in *Progress in Pattern Recognition 2*, Ed. by L. N. Kanal & A. Rosenfeld, North-Holland (1985).

(4) R. C. Gonzalez & P. Wintz, *Digital Image Processing*, Addison-Wesley, Reading, Mass., (1977).

(5) M. D. Levine, *Vision in Man and Machine*, McGraw-Hill (1985).

(6) P. E. Danielsson & S. Levialdi, Computer architectures for pictorial information systems, Computer, vol. 14, no. 11, pp. 53-67, 1981.

(7) E. Persoon & K. S. Fu, Shape descrimination using fourier descriptors, IEEE Trans. Syst., Man, Cyber., vol. SMC-7, pp. 170-179, 1977.

(8) C. T. Zahn & R. Z. Roskies, Fourier descriptors for plane closed curves, IEEE Trans. on Computers, vol.C-21, pp. 269-281, 1972.

(9) F. Etesami & J. J. Uicker, Automatic dimensional inspection of machine part cross-sections using fourier analysis, Computer Vision Graphics and Image

Processing, vol 29, pp. 216-247, 1985.

(10) T. P. Wallace & P. A. Wintz, An efficient three-dimensional aircraft recognition algorithm using normalized fourier descriptors, Computer Vision, Graphics and Image Processing, vol. 24, pp. 99-126, 1980.

(11) T. Pavilidas, *Structural Pattern Recognition*, Springer-Verlag, (1977).

(12) F. Alt, Digital pattern recognition by moments, Journal of the ACM, vol. 9, pp. 240-258, 1962.

(13) E. L. Hall, W. O. Crawford & F. E. Roberts, Computer classification of pneumoconiosis from radiographs of coal workers, IEEE Trans. on Bio. Med. Engr., vol. BME-22, no. 6, pp 518-527, 1975.

(14) F. W. Smith & M. H. Wright, Automatic ship photo interpretation by the method of moments, IEEE Trans. on Computers, vol. C-20, no. 9, pp. 1089-1095, 1971.

(15) S. A. Dudani, K. J. Freeding, & R. B. McGhee, Aircraft Recognition by moment invarients, IEEE Trans. on Computer, vol. C-26, pp 39-46, 1977.

(16) M. K. Hu, Visual pattern recognition by moment invariants, IRE Trans. Inform. Theory, vol. IT-8, pp. 179-187, 1962.

(17) A. Goshtasby, Template matching in rotated images, IEEE Trans. Pattern Analysis and Mach. Intell., vol. PAMI-7, pp. 338-344, 1985.

(18) A. P. Reeves, A parallel mesh moment computer, Proc. of the Sixth Inter. Conf. of Pattern Recog., pp. 465-467, 1982.

(19) J. S. Albus, *Brain, behavior and robotics*, BYTE Books, (1981).

(20) D. H. Hubel & T. N. Wiesel, Brain mechanisms of vision, Scientific American, pp. 150-162, 1979.

(21) T. A. Feldman, Connectionist models and parallelism in high level vision, Computer Vision, Graphics and Image Processing, vol. 31, pp. 178-200, 1985.

(22) D. H. Ballard, Parameter nets, Artificial Intelligence, vol. 22, pp. 235-267, 1984.

(23) J. A. Feldman & D. H. Ballard, Connectionist models and their properties, Cognitive Science, vol. 6, pp. 205-254, 1982.

(24) D. Sabbah, Computing with connections in visual recognition of origomi objects, Cognitive Science, vol. 9, pp. 25-50, 1985.

(25) R. O. Duda & P. E. Hart, Use of the Hough transformation to detect lines and curves in pictures, Comm. of the ACM, pp. 11-15, 1972.

(26) S. P. Smith & A. K. Jain, Chord distribution for shape matching, Computer Vision, Graphics and Image Processing, vol. 20, pp. 259-271, 1982.

(27) L. A. Santalo, *Integral geometry and geometric probability*, Addison-Wesley, Reading, Mass. (1976).

(28) D. H. Ballard, Generalizing the Hough transform to detect arbitrary shapes, Pattern Recognition, vol. 13, pp. 111-122, 1981.

(29) M. Kushnir, K. Abe & K. Matsumoto, An application of the Hough transform to the recognition of printed Hebrew characters, Pattern Recognition, vol. 6, pp. 183-191, 1983.

(30) S. D. Shapiro, Use of the Hough transform for image data compression, Pattern Recognition, vol. 12, pp. 333-337, 1980.

(31) K. A. Lantz, C. M. Brown & D. H. Ballard, Model-driven vision using procedure decription : Motivation and application to photointerpretation and medical diagnosis, Proc 22nd International Symp., Society of Photo-optical Instrumentation Engineers, 1978.

(32) H. Wechsler & J. Sklansky, Automatic detection of ribs in chest radiographs, Pattern Recognition, vol. 9, pp. 21-30, 1977.

(33) G. C. Stockman A. K. Agrawala, Equivalence of Hough curve detection to template matching, Comm. of ACM, vol. 20, pp. 820-822, 1977.

(34) T. M. Van Veen & F. C. A. Groen, Discretization error in the Hough transform, Pattern Recognition, vol. 14, pp. 137-145, 1981.

(35) J. Sklansky, On the Hough technique for curve detection, IEEE Trans. on Computers, vol. c-27, pp. 923-926, 1978.

(36) D. H. Ballard & C. M. Brown, *Computer Vision*, Prentice-Hall (1982).

(37) L. S. Davis, Hierarchical generalized Hough transforms and line-segment based generalized Hough transforms, Pattern Recognition, vol. 15, pp. 277-285, 1982.

(38) J. L. Turney, T. N. Mudge & R. A. Volz, Recognizing partially occluded parts, IEEE Trans. Pattern Analysis and Mach. Intell., vol. PAMI-7, no. 4, 1985.

(39) W. Chang & D. Casasent, Chord distribution in pattern recognition: Distortion invariance and parameter estimation, SPIE, vol. 521, pp. 2-6, 1984.

(40) D. J. H. Moore & D. J. Parker, Analysis of global pattern features, Pattern Recognition, vol. 6, pp. 149-164, 1974.

(41) Z. You & A. K. Jain, Performance evaluation of shape matching via chord length distribution, Computer Vision, Graphics and Image processing, vol. 28, pp. 185-198, 1984.

(42) J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press (1982).

APPENDIX A.

PSEUDOCODE FOR MODULE, COMPARE

INPUT :

1) A library shape description, S.

2) An Unknown view, V.

OUTPUT :

1) Parameters specifying the location and orientation of the instance of S most supported by V.

2) Magnitude of the maximum element in the reference point voting array.

COMMENTS :

Module, Compare, takes as input a library shape description and the unknown view and carries out a voting procedure for the location and orientation of the library object in the unknown view. This voting procedure counts the number of chords in the unknown view that support each location and orientation of the library object. Note that each chord, s, in S has a voting vector attached to it that points to a reference point in the shape description, S.

Algorithm Compare.

Initialize the reference point voting array, RP, and the orientation voting array, OV, to be zero.

For each chord, v, in V do

Let S' be the subset of S, such that for each element, s', $|meanGreylevel(s') - meanGreylevel(v)| < T_1$

For each s' in S' do

if $|length(s') - length(v)| < T_2$ then

if $match(s',v)$ then
construct a voting vector, l, relative to v, using the reference vector, l', of s'.

Increment $RP(x,y)$ by one, where (x,y) is the location pointed to by the vector l.

endif

endif

enddo

enddo

Search RP for the maximum element, (v_x,v_y).

Let S'' be the subset of S, such that each chord, s'', voted at least once at $RP(v_x,v_y)$.

For each chord, s'', in S'' do

Let θ be the rotation s'' went through to vote for (v_x,v_y).

Increment $OV(\theta)$

enddo

Search OV for the maximum element, t.

Return v_x, v_y, t

Note :

(1) T_1, and T_2 are preset thresholds which for the experiments using Athena were in the ranges, [2,4] and [5,10], respectively.

(2) The function, *match*(a,b), computes the average of the absolute value of the difference of the I-functions of a and b. The function, *match*, compares both the I-function of a, and the reflection of the I-function of a, to the I-function of b, selects the better of the two matches, and returns true if the computed value is less than T_3. T_3 is another preset threshold which had a value in the ranges, [5,10].

APPENDIX B.

Experiment 1.

Number of chords in each library description : 400

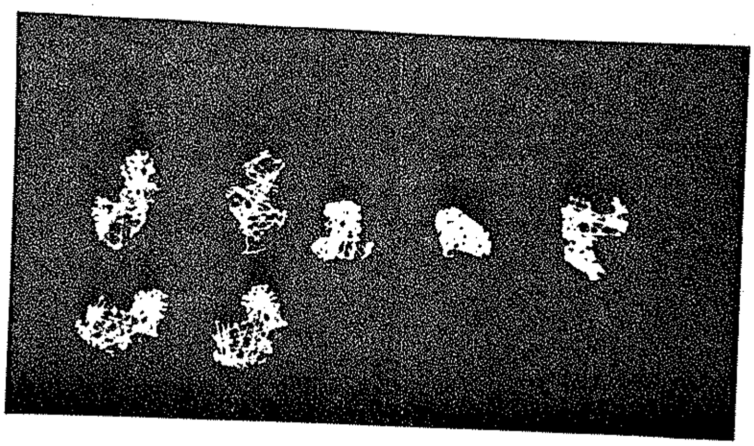
Number of chords in the unknown view : 400

Table 1.

Name	Votes	% of 400
A	56	14.00
B	22	5.05
C	36	9.00
D	28	6.00
E	7	1.65

Comments :

The unknown object is an instance of A. The angle the unknown object should be rotated by to be in the same orientation as the library is 38 degrees.



A	B	C	D	E
A'	A'.r			

Experiment 2.

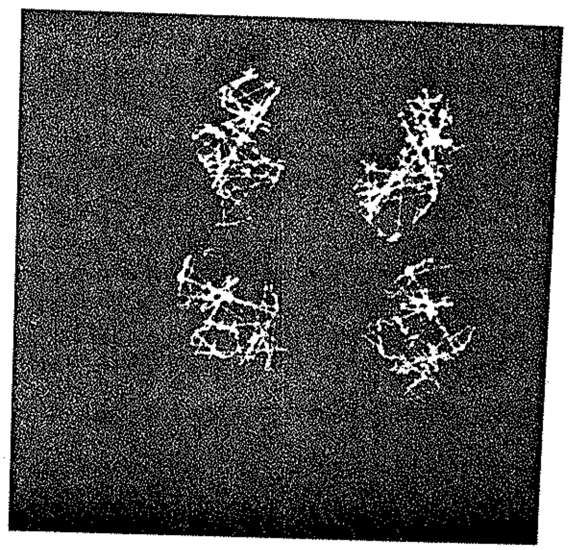
Number of chords in each object description : 400

Number of chords in the unknown view : 340.

Table 2

Name	Votes	% of 340.
A	46	13.52
B	16	4.70

The unknown view is an instance of object A. The angle the unknown object should be rotated by, to be in the same orientation as the library, is 40 degrees.



A	B
A'	A'.r

Experiment 3.

Number of chords in each library object : 300

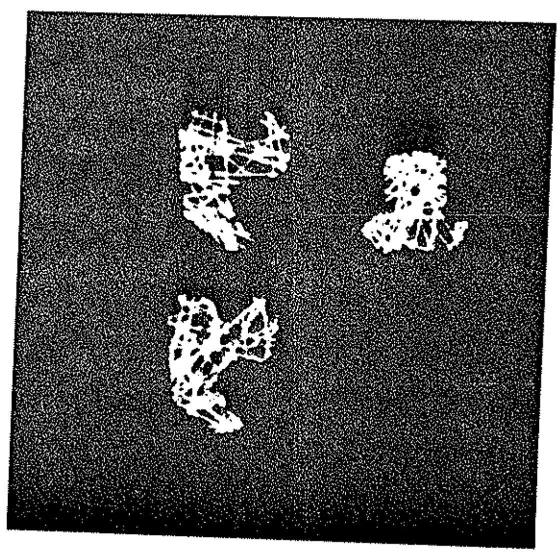
Number of chords in the unknown view : 300

Table 3.

Name	Votes	% of 300
A	24	8.00
B	4	1.03

Comments :

The unknown view is an instance of object A. The angle the unknown view should be rotated by, to be in the same orientation as the object in the library, is 1 degree.



A	B
A'	

Experiment 4.

Number of chords in each object description : 300

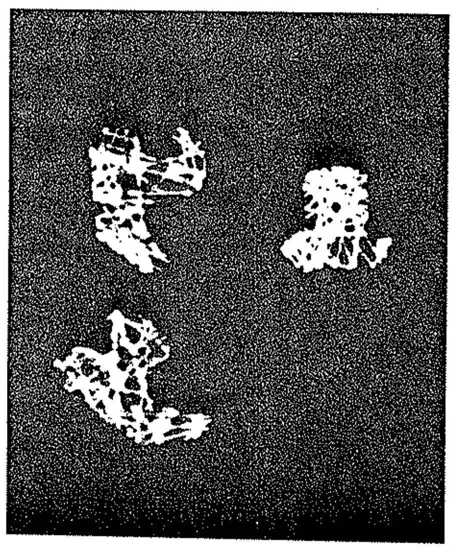
Number of chords in the unknown view : 300

Table 4.

Name	Votes	% of 300
A	42	14.00
B	6	2.00

Comments :

The unknown image is an instance of object A. The unknown image should be rotate by 44 degrees, in order to be in the same orientation as the object in the library.



A B
A'.r

Experiment 5.

Number of chords in each library description : 200.

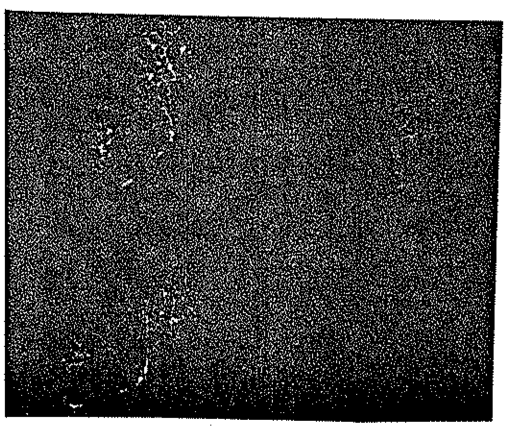
Number of chords in the unknown view : 200.

Table 5.

Name	Votes	% of 200
A	200	100.00
B	0	0.00

Comments :

The unknown view is an instance of object A in the library. The unknown view should be rotated by an angle of -25 degrees in order to be in the same orientation as the object in the library. In this example, the unknown view had exactly the same chords as object A in the library, analytically rotated by 25 degrees.



A B
A'.r

Experiment 6.

Number of chords in each library description : 200

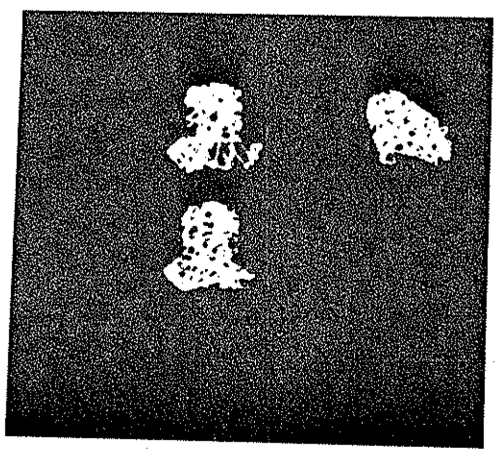
Number of chords in the unknown view : 200

Table 6.

Name	Votes	% of 200
A	33	16.05
B	8	4.00

Comments :

The unknown view is an instance of object A in the library. The angle the unknown view should rotate by in order to be in the same orientation as the object in the library is 2 degrees.



A	B
A'	

Experiment 7.

Number of chords in each library description : 200

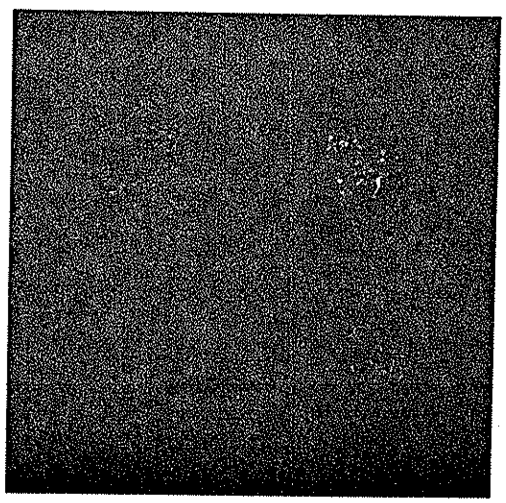
Number of chords in the unknown view : 200

Table 7.

Name	Votes	% of 200
A	11	5.05
B	92	46.00

Comments :

The unknown view is an instance of object B in the library. The angle the unknown view should rotate by in order to be in the same orientation as the object in the library is 0 degrees.



A	B
	B'

Experiment 8.

Number of chords in each object in the library : 200

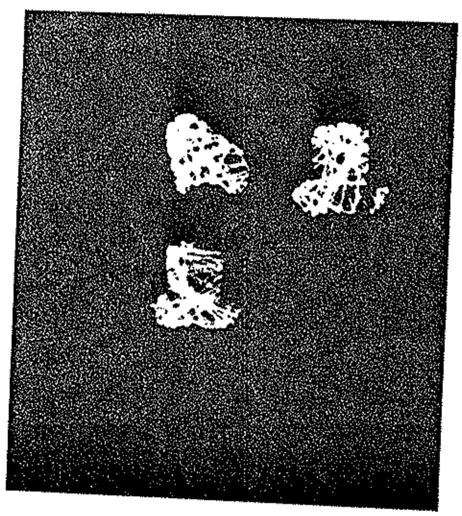
Number of chords in the unknown view : 200

Table 8.

Name	Votes	% of 200
A	15	7.50
B	0	0.00

Comments :

In this case the unknown view has the same texture as object A in the library and the same shape as object B. The unknown view was incorrectly identified as object A.



A	B
B'	

Experiment 9.

Number of chords in each library description : 400.

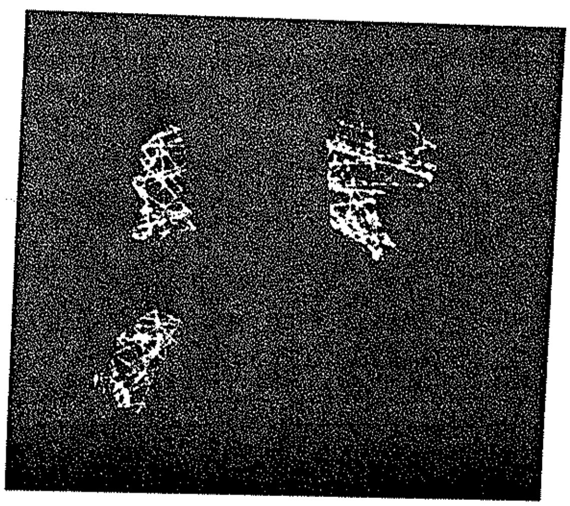
Number of chords in the unknown view : 400.

Table 9

Name	Votes	% of 400
A	68	17.00
B	10	2.02

Comments :

The unknown view is an instance of object A in the library. The unknown view should be rotated by 35 degrees to be in the same orientation as the object in the library.



A	B
A'.r	

Experiment 10.

Number of chords in each library description : 300

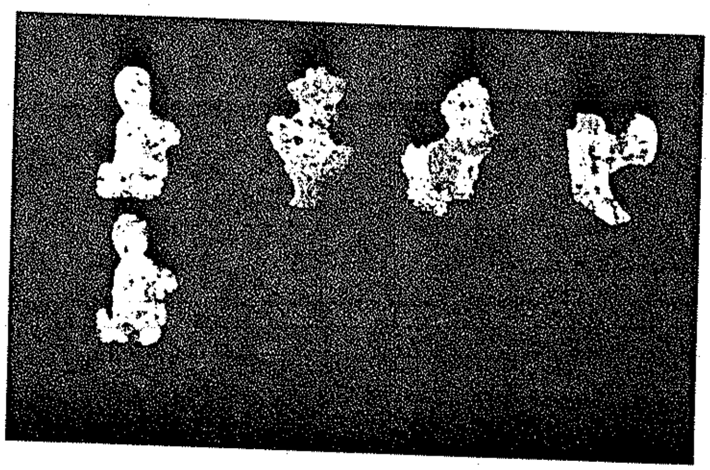
Number of chords in the unknown view : 300

Table 10.

Name	Votes	% of 300
A	32	10.06
B	13	4.03
C	13	4.03
D	4	1.03

Comments :

The unknown view is an instance of object A in the library. In order to be in the same orientation as the object in the library, the unknown view should be rotated by 0 degrees.



A	B	C	D
A'			

Experiment 11.

Number of chords in each of the library descriptions : 200

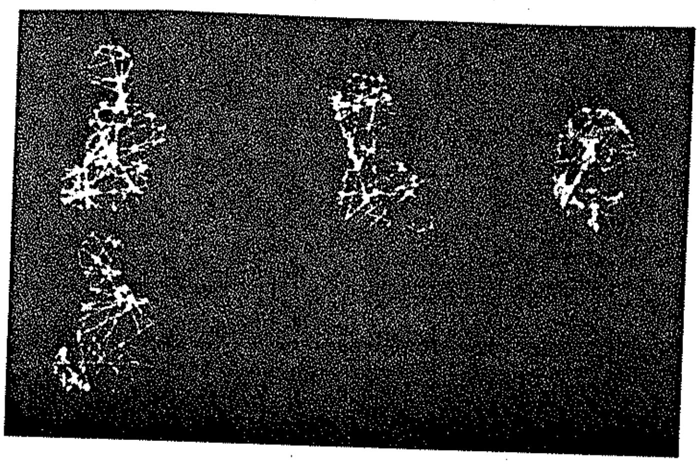
Number of chords in the unknown view : 350

Table 11

Name	Votes	% of 350
A	43	12.28
B	18	5.14
C	10	2.85

Comments :

The unknown view has been identified as an instance of object A. In order to be in the same orientation as the object in the library, the unknown view should be rotated by an angle of 1 degrees.



A	B	C
A'		