IMAGE RECOGNITION USING GENERALIZED CORRELATION

by

Lee Edwin McDonald

August 1977 UTEC-CSc-77-135

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-73-C-0363.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure

Tables

ABSTRACT

This paper investigates the use of generalized cross-correlation in pattern matching when the objects may be of one or two dimensions. Generalized correlation can be used to determine the amount of dilatation and rotation between a given template and an object, in addition to determining the relative translation. Two techniques are discussed which break this four-dimensional correlation into two two-dimensional correlations making the problem computationally feasible. The techniques were developed for a specific class of images, however they can be applied to a more general class.

CHAPTER I

INTRODUCTION

In recent years it has become practical and desirable to have the computer analyze images. An image is a representation of a three-dimensional scene which may be composed of many objects on an arbitrary background. Image analysis can be used for many purposes in a variety of fields where the capability for a machine to interpret a scene is desired. Some applications include processing of satellite photographs and automatic monitoring of production lines. The ultimate goal is to be able to determine what objects appear in a three-dimensional scene and any desired information about those objects. The problem is that each object in the image has six degrees of freedom: two translations, size (dilatation), rotation in the plane, and two rotations out of the plane. To further complicate the problem one object may partially obscure another. Although at this time a complete solution is beyond our insight and capabilities, it is hoped that analysis of simpler cases will enable us to understand the problem completely.

In this context, the work presented here is based on the restriction that all images are of a single object on a black background. However, there are no restrictions on the type of object that may appear in the image. For each object in which we are interested, we need a template which contains an instance of that object, to match with the object in the image. The object in the image may differ from the one in the template by translation, dilatation, and rotation. There are two questions which need to be answered in the context of this problem. They are, is there an instance of the template object in the image, and if so, what is the rotation, dilatation, and translation of the image object with respect to the template object. The techniques developed will then be extended to include images which meet slightly less stringent conditions.

Cross-correlation has been used in the field of pattern recognition primarily for the determination of the relative translation between the image and the template. In this paper, the term generalized correlation refers to the correlation of two functions with respect to dilatation (size), rotation, and translation. Techniques are described which enable one to use generalized correlation to determine the relative size and rotation along with the translations. When doing this, one is dealing with a four-dimensional problem which is computationally impractical. The goal of this work is to find new techniques to reduce the dimensionality of the generalized correlation computation.

Chapter II discusses cross-correlation, correlation coefficients and computing cross-correlation using Fourier

Transforms. The reader familiar with these topics may skip this chapter with no loss of continuity.

Chapter III presents generalized correlation and two methods for computing it. Generalized correlation can be broken into two two-dimensional problems. Both methods transform the image into a domain in which some degrees of freedom are eliminated. The resulting problem is easier to attack.

Chapter IV presents algorithms to compute generalized correlation for both one and two dimensional images. The images are of single objects on a black background. For each dimensionality two approaches to the computation of generalized correlation are examined.

Chapter V examines these algorithms for applicability to other types of images. Examples of the types of images considered are multiple objects on a black background and a single object on a textured background. To extend this technique to other cases, it may be necessary to preprocess the image before the correlation can be done.

Chapter VI is concerned with the problems encountered and some computational techniques used in implementing the algorithms on a digital computer. The problems are all results of the discrete and finite nature of the computer. Ways are discussed which minimize the effects of these limitations without any significant change to the algorithms.

Chapter VII presents the results of testing the system for several different types of images while Chapter VIII concludes this paper with a discussion of the possibilities for the future.

CHAPTER II

BACKGROUND

Conceptually, correlation provides a quantitative measure of the similarity of two functions. This work uses two types of correlation, cross-correlation and generalized correlation. Cross-correlation, which is discussed in this chapter, is used to determine the degree of correlation between two functions when one is translated with respect to the other. Generalized correlation will be discussed in Chapter III. This chapter may be skipped by readers familiar with cross-correlation.

Cross-Correlation

The cross-correlation function computes the correlation between two functions in terms of relative translations. In one dimension this takes into account the effects of only one variable (degree of freedom), namely the translation (shift). The cross-correlation $\phi(u)$ of two functions $f(x)$ and $g(x)$ is defined as

$$\phi(u)=\frac{1}{T_2-T_1}\int_{T_1}^{T_2} f(x)g(x+u)dx$$

where T_1 and T_2 are denote the interval of interest. Thus, for any value of u, $\phi(u)$ is the correlation between f and a version of g which has been shifted u units. In the case of two dimensional functions, the cross-correlation accounts for translations in the x and y directions and is defined by

$$\phi(u,v)=\frac{1}{T_2-T_1}\,\frac{1}{R_2-R_1}\int_{T_1}^{T_2}\int_{R_1}^{R_2} f(x,y)g(x+u,y+v)\,dy\,dx$$

where T_1, T_2, R_1, and R_2 denote the interval of interest.

One of the disadvantages of the cross-correlation function is that it gives no indication to the absolute degree of similarity. All it provides for each shift is a measure of the area of overlap between the two functions. This deficiency can be rectified by normalizing the correlation to arrive at the correlation coefficient r(u) which is defined as

$$r(u)=\frac{\phi(u)}{\frac{1}{T_2-T_1}\sqrt{\int_{T_1}^{T_2} f(x)^2dx \int_{T_1}^{T_2} g(x+u)^2dx}}$$

The correlation coefficient r(u) ranges in value from +1 to -1. It is interesting to note that $r(u)^2$ can be interpreted as the fraction of one function attributable to the other [1]. It has been shown that the correlation coefficient is an absolute measure of the closeness of two functions in a least squares sense [2]. The correlation-coefficient in two dimensions is

$$r(u,v)=\frac{\phi(u,v)}{\frac{1}{T_2-T_1}\,\frac{1}{R_2-R_1}\sqrt{\int_{T_1}^{T_2}\int_{R_1}^{R_2} f(x,y)^2dydx\int_{T_1}^{T_2}\int_{R_1}^{R_2} g(x+u,y+v)^2dydx}}$$

When using cross-correlation in two dimensional pattern matching, there are usually two objectives. The first is to determine if there is an instance of the template in the image and the second is to find where that instance occurs. Both answers can be found by examining the correlation-coefficient function r(u). This function has maximum value at the position corresponding to the most probable translation of the image relative to the template. At that point one can then make the decision whether or not the maximum is significant by examining the value of the function. This technique is often used with optical functions (also called matched filtering), but has the limitation that the template must be the same size and orientation as the image [3].

Computation of Cross-Correlation

The cross-correlation between two functions f(x) and g(x) can be computed using Fourier Transforms as follows:

$$\phi(u)=\mathcal{F}^{-1}(F(\omega)G^*(\omega))$$

where $F(\omega)$ and $G(\omega)$ are the Fourier Transforms of f(x) and g(x) respectively, and $\mathcal{F}^{-1}$ denotes the inverse Fourier Transform. The advantage of this approach to computing cross-correlation over direct integration is when working with discrete functions the Discrete Fourier Transform can be implemented in an efficient manner such that it becomes faster to cross-correlate two functions using Fourier Transforms than using the defining summation. The implementation details of cross-correlation will be presented in Chapter VI, but the underlying theme of this work is to find transform methods to enable faster computation of generalized correlation.

CHAPTER III

THE MATHEMATICS OF GENERALIZED CORRELATION

A limitation of cross-correlation is that the functions are correlated for relative translations only. The concept of generalized correlation is to correlate two functions for relative rotation and dilatation (scaling) as well as for translations. This chapter presents generalized correlation and discusses techniques for computing it. The goal of the computational techniques is to develop algorithms for efficient evaluation of the generalized correlation by using transforms to arrive in a domain in which the problem statement leads to a simple evaluation technique.

Generalized correlation will be discussed first, followed by a discussion of the computation techniques to be used. The algorithms are presented in the next chapter since each algorithm uses a different set of techniques to determine the generalized correlation. The computation of the generalized correlation is based on separating the problem into two simpler sub-problems.

Generalized Correlation

Generalized correlation extends the concept of cross-correlation to account for the ways other parameters affect the value of the correlation-coefficient. In one dimension the correlation is done for both dilatation (size) and translation. The correlation can then be defined as a function of translation, u, and size, s. This gives

$$\phi(u,s)=\frac{1}{T_2-T_1}\int_{T_1}^{T_2} f(x)g(sx+u)dx$$

with the corresponding correlation coefficient

$$r(u,s)=\frac{\phi(u,s)}{\frac{1}{T_2-T_1}\sqrt{\int_{T_1}^{T_2} f(x)^2dx\int_{T_1}^{T_2} g(sx+u)^2dx}}$$

Note: Generally, only the correlation function will be used since the correlation coefficient can be obtained at any time by dividing by the product of the norms, or rms values, of the functions over the appropriate interval.

Generalized correlation in two dimensions uses two parameters which do not appear in the cross-correlation. They are dilatation (size) and rotation (orientation). Both of these parameters can be thought of as creating new functions, but it is more revealing to think of the correlation as a function of horizontal position , u, vertical position, v, size, s, and rotation, α. This gives us

$$\phi(u,v,s,\alpha)=\frac{1}{T_2-T_1}\ \frac{1}{R_2-R_1}\ \int_{T_1}^{T_2}\int_{R_1}^{R_2} f(x,y)g(x',y')dydx$$

where

$$x'=s(x\cos\alpha+y\sin\alpha)+u$$

and

$$y'=s(y\cos\alpha - x\sin\alpha)+v$$

One of the major problems with using generalized correlation in practice is that since it is a function of four independent variables it becomes computationally impractical for all but small intervals of u,v,s, and α. If one increases the interval for each variable by the same factor k then the amount of computation increases by k^4. Furthermore, if the limits of both R and T are increased by some factor q the area of integration is increased by q^2. Consequently it is desirable to find ways to compute the four dimensional correlation other than by direct integration.

Separation of the Generalized Correlation

This section discusses two techniques of dividing the four-dimensional correlation $\phi(u,v,s,\alpha)$ or $r(u,v,s,\alpha)$ into two two-dimensional problems. Both techniques are based on the independence of the four degrees of freedom. This independence enables one to determine the values of the parameters representing the degrees of freedom separately. Note that these separations assume that each image is of a single object on a black background. The first technique for computing the generalized correlation depends on the following property of the Fourier Transform: if

$$\mathcal{F}(f(x))=F(\omega)$$

then

$$\mathcal{F}(f(ax))=F(\omega)e^{-ia\omega}$$

where $\mathcal{F}$ represents the Fourier Transform. Since translations in one domain become linear phase components in the other, the magnitude of the Fourier Transform of an image has no information concerning the location of an object. This enables one to determine the scale and rotation without having translational information that one can not interpret correctly. The portion of the Fourier Transform removed, that is the phase, contains far more information than just the translational components. Removing the phase eliminates information that prevents the Fourier Transform from being ambiguous. An example of the ambiguities that phase resolves is a reflection of the object through the origin (in one dimension, a mirror image). Determination of the scale and rotation is discussed in the next section on exponential polar coordinates. This method of separation will be referred to as the magnitude method.

The second technique for computing the generalized correlation is based on the invariance of the centroid, or first moment of an object, under rotation and scaling. The centroid of a function f(x,y) occurs at a point (a,b) given by

$$(a,b)=\frac{\int_{T_1}^{T_2}\int_{R_1}^{R_2}f(x,y)(x,y)dydx}{\int_{T_1}^{T_2}\int_{R_1}^{R_2}f(x,y)dydx}$$

Computation of the centroid immediately obtains the translation portions of the correlation. The scale and rotation factors are then obtained by first translating the object so that the centroid occurs at the origin and then correlating with respect to scale and rotation. This method will be referred to as the centroid method.

Exponential-Polar Coordinates

The previous section discussed ways of separating the generalized correlation into two lower dimensional problems. In both cases, the image is left in a domain in which translational information is not present. What is desired is to cross-correlate these two functions with respect to scale and rotation to determine what those factors are. Cross-correlation correlates functions for different shifts, hence in this case another domain is needed where scale changes are reflected as shifts in one direction and rotations as shifts in the other.

A rotation of an object, by the angle α, about the origin in a rectangular coordinate system is equivalent to shifting the object by α along the angle axis in a polar coordinate system. In making this conversion, it is necessary to insure that the correlation is not affected by the coordinate system in which the function is expressed. Examination of the Jacobian of the transformation gives

$$\iint_{Rxy} f(x,y)g(x',y')dydx=\iint_{Rr\theta} f(r,\theta)g(r',\theta')rdrd\theta$$

where

$$x'=s(x\cos\alpha+y\sin\alpha)+u$$

$$y'=s(y\cos\alpha-x\sin\alpha)+v$$

$$r'=sr$$

and $\theta'=\theta+\alpha$

for cross-correlation where Rxy and $Rr\theta$ are equivalent domains [4].

One of the disadvantages of rectangular coordinates is that scaling an object affects its description in both x and y. In polar coordinates, however, scaling affects only the radius, r. By converting to an exponential basis for r, scale factors are converted to shifts. This conversion is achieved by the change of variables $r=e^w$ where w is the new independent variable.

Figure 1 illustrates this conversion for a one dimensional signal. In one dimension the radius r is equivalent to the usual independent variable. Given f(r) and a scaled version of it, f(ar), substituting $r=e^w$ and $a=e^b$

$$
\begin{aligned}
f(r) &= f(e^w)=g(w)\\
f(ar) &= f(ae^w)\\
&= f(e^b e^w)\\
&= f(e^{w+b})=g(w+b)
\end{aligned}
$$

Figure 1 (e) and (f) show a one-dimensional example of two

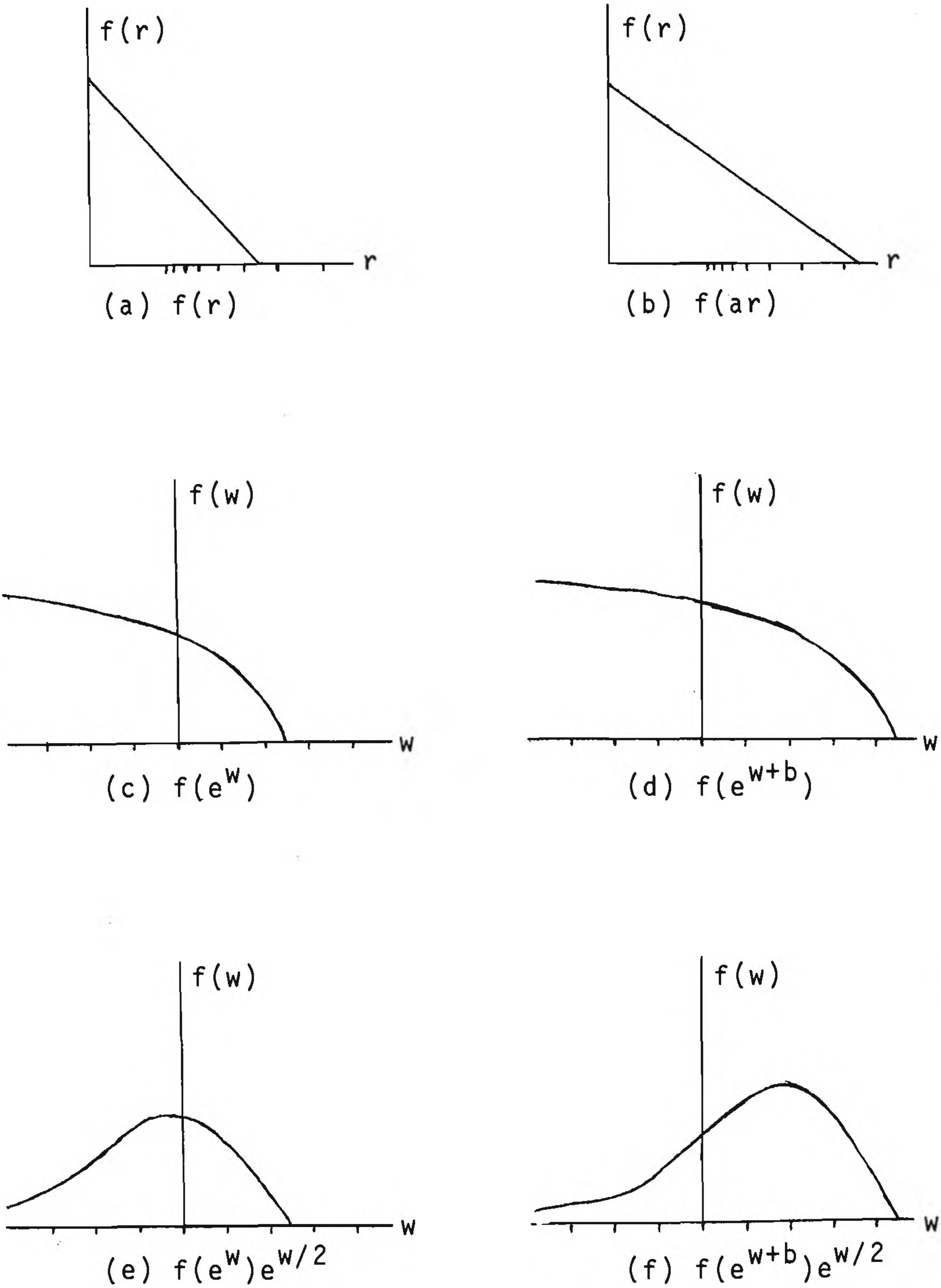


Figure 1 Exponential Coordinates

functions after the change of variables where each function has been multiplied by $e^{w/2}$ in preparation for cross-correlating. The transformed correlation is

$$\int_{Rr} f(r)g(sr)dr=\int_{Rw} f(w)g(w+v)e^{w}dw$$

where Rr and Rw denote equivalent domains.

In two dimensions the same change of variables $r=e^{w}$ is applied radially. Using this transformation, scaling in r is equivalent to shifting in w. Again, we need to insure the correlation is invariant. Here the Jacobian of the transformation $r=e^{w}$ gives

$$\iint_{Rr\theta} f(r,\theta)g(sr,\theta)rdrd\theta=\iint_{Rw\theta} f(w,\theta)g(w+v,\theta)e^{2w}dwd\theta$$

The two step change of coordinate systems shown can be thought of as a single change of variables where

$$x=e^{w}\cos\alpha$$

and

$$y=e^{w}\sin\alpha$$

This gives the same result as the two step process described above.

This one-dimensional coordinate system will be referred to as exponential coordinates, while the two-dimensional will be called exponential-polar coordinates. The advantage in converting from rectangular coordinates to exponential-polar coordinates is that scale changes map into shifts along the w axis and rotation maps into shifts along the θ

axis. A two dimensional cross-correlation can then be done to determine the scale and rotation. Note that this conversion is defined only for non-negative values of r.

Generalized correlation can determine the translational, rotational, and scale relationships between two functions. Two methods have been discussed for separating the generalized correlation into two problems which are easier to solve. Both techniques assume that the function is a single object on a black background. Exponential coordinates were developed as a domain in which scale changes are reflected as shifts along one axis and rotations are reflected as shifts along the other axis.

CHAPTER IV

ALGORITHMS FOR COMPUTING GENERALIZED CORRELATION

Generalized correlation can be useful in pattern matching. When given a template which has an instance of the object to be found, generalized correlation helps one discern if that object appears in the image. Traditionally, the template object can differ from the instance in the image only by translation, thus cross-correlation is used to find the instance. The use of generalized correlation makes it possible for the template object to differ from the instance in the image by a rotation and scale change in addition to the translations.

The previous chapter discussed the significance of generalized correlation and the mathematical techniques used in computing it. This chapter will present algorithms which use generalized correlation in pattern matching. The techniques presented compute the correlation assuming the function being correlated is of a single object on a black background. First, the magnitude and centroid methods for one-dimensional pattern matching will be described. Secondly, these methods will be discussed in relation to two-dimensional pattern matching.

One-Dimension

This work is concerned with two of the degrees of freedom of a one-dimensional function. The first degree of freedom is translation. The object of interest may occur anywhere along the independent axis. Cross-correlation is often used to determine the value of this variable. However, cross-correlation can not determine if the independent variable has been scaled. This is the second degree of freedom with which this section is concerned. As previously discussed, generalized correlation in one-dimension facilitates the computation of the most probable values for the variables which represent these derees of freedom by breaking the two-dimensional problem into two one-dimensional problems.

Magnitude Method

The algorithm for computing the one-dimensional generalized correlation by the magnitude method is outlined in Figure 2. The first step is to take the magnitude of the Fourier Transform of both the template and the image. This removes all information concerning the location of the objects.

The second step converts scale factors to shifts by converting to exponential coordinates as discussed in the exponential-polar section of Chapter III. This conversion is dependent on two factors: a) the scaling being done about

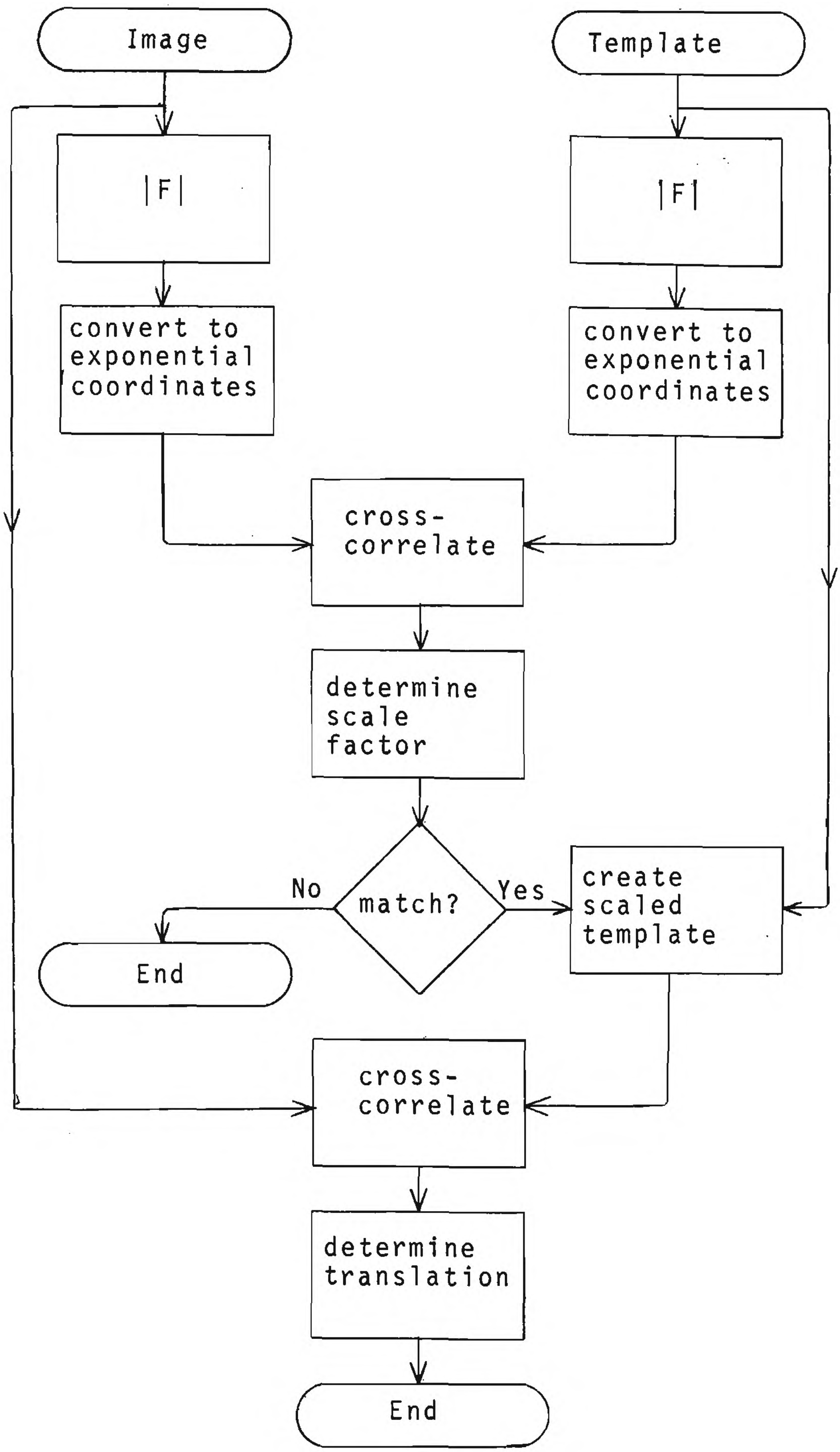


Figure 2
One-dimensional generalized correlation. Magnitude method.

the origin, and b) there being no translation between the template and the image. Using the magnitude of the Fourier Transform insures the above conditions are met, because the magnitude of the Fourier Transform of a scaled object is scaled about the origin and there is no translation in it. This makes the conversion possible. Since the magnitude of a real function is even, only the non-negative frequencies need be considered. This is a prerequisite for the use of the exponential coordinates.

The third step computes the cross-correlation of the two functions, i.e.,magnitude of template in exponential coordinates and magnitude of image in exponential coordinates. This cross-correlation is from $-\infty$ to ∞ since the coordinate change maps the frequencies between 0 and 1 into the range $-\infty$ to 0. The peak in this correlation occurs at the location b, where the true scale factor s is related to b by

$$s=e^{-b}$$

as derived in Chapter III. At this point a decision can be made whether or not the image contains an instance of the object in the template. Computing the correlation-coefficient (which ranges between +1 and -1) gives the user a basis on which to decide. The decision of which values represent a match must be determined experimentally on sample data.

If it is decided that there is a match, then there remains only to determine the amount the image has been shifted with respect to the template. First, using the scale factor already determined, a scaled version of the template is created in which the template object is the same size as the object in the image. The image is now cross-correlated with the scaled template. The peak occurs at a point u, meaning the image has been shifted by u units with respect to the template. It is not necessary that the correlation coefficient be computed at this point if the match/no match decision has been made.

Summarizing the magnitude method for one-dimensional generalized correlation, there are five steps.

1. Find the magnitude of the Fourier Transform of the image and the template
2. Convert both the image and the template to exponential coordinates
3. Cross-correlate to determine the scale factor
4. Create a scaled version of the template of the same size as the image
5. Cross-correlate to determine the translation

Centroid Method

The centroid method for computing one-dimensional generalized correlation determines where the centroid of the image is and translates the image so the centroid is at the origin. This removes the translational effects. The next step would be to convert the image into exponential coordinates. It is at this point the centroid method in one-dimension breaks down.

The conversion to exponential coordinates has two assumptions. The first is the object is scaled about the origin. The second is the values of the function f(r) for negative r do not matter. After shifting the centroid to the origin, the first condition holds, however, the second one does not. This is because there is part of the function on each side of the origin. As a result the centroid method is not used in one-dimension.

Two-Dimension

Four degrees of freedom will be considered using two-dimensional generalized correlation. Two of the degrees of freedom are the translation in x and the translation in y. An object can occur anywhere in the plane and so both translations are needed to locate it. The other two degrees of freedom are rotation and change of size (scaling). Two-dimensional cross-correlation can be used to locate an object in an image when the template object differs only by translation. However, generalized correlation can locate an object in an image when the template differs by two translations, a rotation in the plane and a change in size. Trying to correlate with respect to four independent variables is a four-dimensional problem which was discussed in Chapter III.

The four-dimensional generalized correlation can be separated into two sub-problems. The first technique of separation to be discussed is the magnitude method. This is based on using the magnitude of the Fourier Transform to remove the translational information from the analysis. The second technique is the centroid method. This method determines the translation by using the fact that the centroid of an object is invariant under rotation and scaling of the object about the centroid.

Magnitude Method

The magnitude method for computing the two-dimensional generalized correlation is very similar to the magnitude method for one-dimensional generalized correlation. The basic approach is:

1. Remove the translational information
2. Determine the scale factor and the rotation
3. Use the scale and rotation factors to help determine the translations

It is important to remember that it is assumed that the functions are each of a single object on a black background. Figure 3 outlines the flow of the algorithm.

The first step of the algorithm is take the magnitude of the two-dimensional Fourier Transform of the image and of the template. This removes the translation dependent information. The magnitude of the Fourier Transform of an image is an even function along radial lines. The importance of this is that in the process of eliminating the

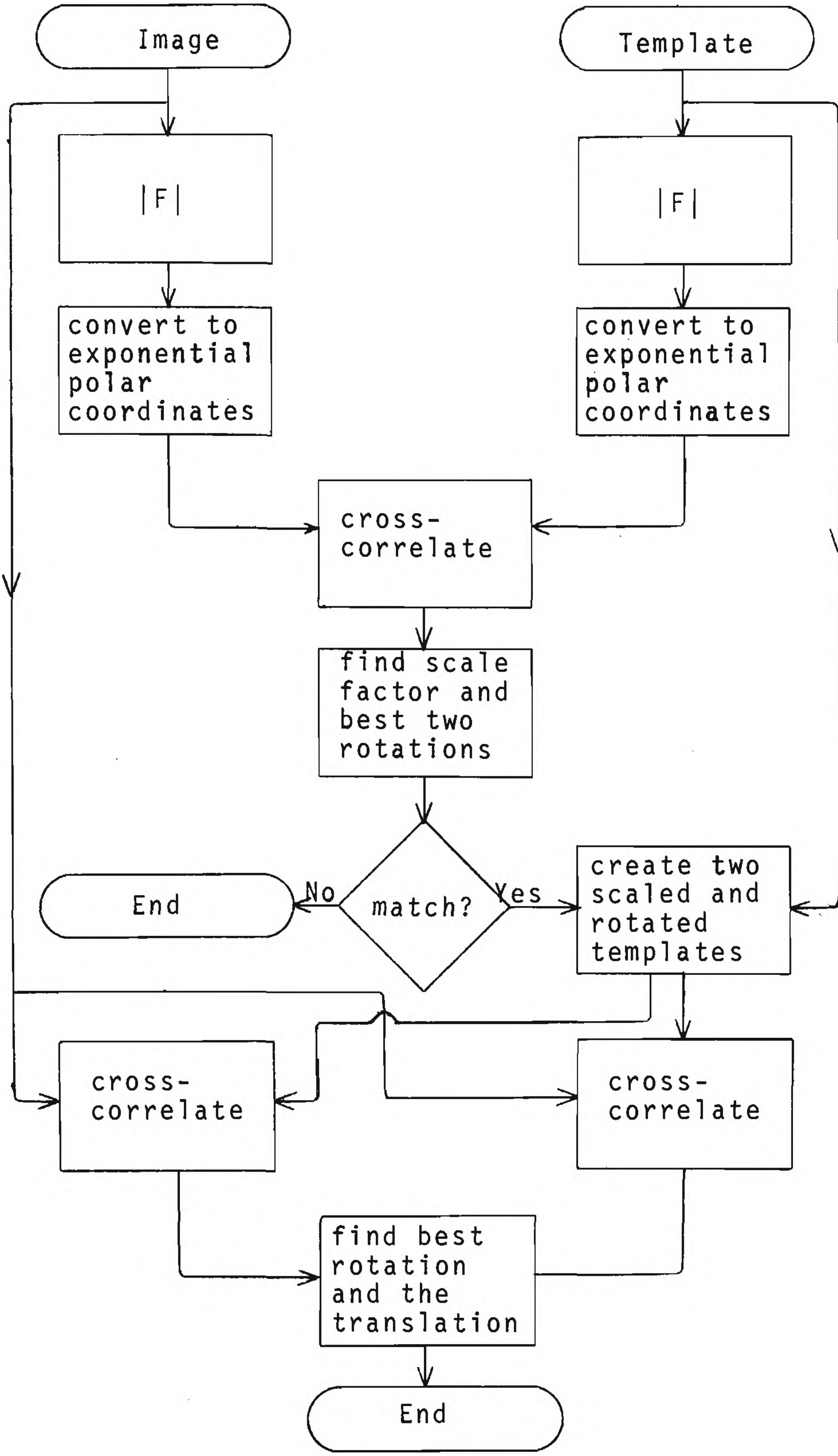


Figure 3
Two-dimensional generalized correlation. Magnitude method.

phase more information is lost than need be, including the ability to distinguish between a rotation of α and a rotation of $\alpha+\pi$. The consequences will be discussed more later.

The second step of the algorithm is to convert the magnitude from rectangular coordinates to exponential-polar coordinates. This conversion was described in Chapter II. The magnitude of the Fourier Transform is centered at the origin. Scaling and rotating an object in the image domain, scales and rotates the magnitude about the origin. This is a prerequisite for the transformation to have the desired effect.

The magnitudes of the image and the template are then cross-correlated in exponential-polar coordinates. If the image is scaled by a and rotated by α then the peak in the correlation occurs at (b,β) where

$$a = e^{-b}$$

$$\text{and either } \alpha = \beta$$

$$\text{or } \alpha = \beta+\pi .$$

The value of α can not be determined completely because the correlation is done with the magnitudes of the image and the template. At this point the decision can be made whether or not the object in the image is the same as the object in the template. The correlation coefficient can be calculated to give an indication of whether or not the functions match. The closer the value of the correlation coefficient is to 1 the more probable the two functions are the same. However, the smallest value which indicates a match has to be determined experimentally for each type of application.

The fourth step uses the information generated in the previous step to make two templates which have an instance of the object the same size as the instance in the image. In one template the object has been rotated by β and in the other the object has been rotated by $\beta + \pi$. This insures that one of the two scaled templates has the object oriented the same way as it is in the image. The last step cross-correlates the original image with the two scaled and rotated templates. The translations and the decision of which rotation is correct is made in this step. The correlation coefficient for each correlation must be computed. The correlation with the larger peak value is the one with the correct value. Furthermore, the peak occurs at (u,v) or, in other words, the image was shifted by u in one direction and v in the other direction relative to the template. The correlation coefficient computed here could be used to make the decision on whether the template and image match rather than computing it when correlating for scale and rotation.

In overview, the algorithm for computing two-dimensional generalized correlation is:

1. Compute the magnitude of the Fourier Transform of the image and of the template

2. Convert both to exponential-polar coordinates
3. Cross-correlate to find the scale factor and the two possible rotations (β and $\beta+\pi$)
4. Create two scaled and rotated templates, one for each rotation to find the correct rotation and the translation

Centroid Method

The centroid method for computing the generalized correlation in two-dimensions computes the centroid of the image and template in order to determine and remove the translation. It is necessary that the image and the template both be of a single object on a black background otherwise the location of the centroid of the image may not be at the centroid of the object. This would cause an incorrect analysis of the situation. Figure 4 outlines the flow of this algorithm.

The first step is compute the centroid of the image and of the template. Then shift the image and the template so the centroid of each is at the origin. In this step the translation has been determined and removed. What remains to be determined is the scale factor, rotation, and whether the image object is an instance of the template object.

The second step converts the image and the template from rectangular coordinates to exponential-polar coordinates. Since the centroid is invariant under scaling and rotation, the image will be scaled and rotated about the

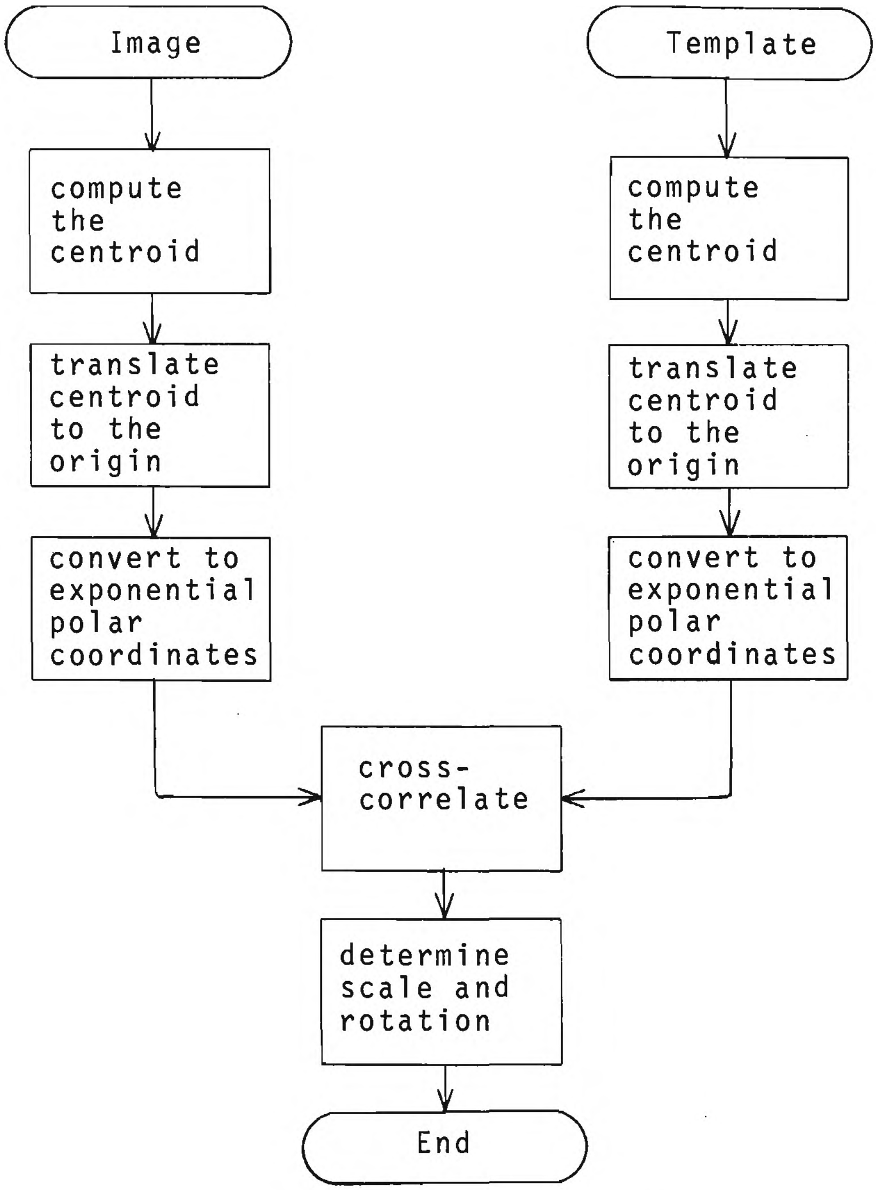


Figure 4
Two dimensional generalized correlation. Centroid method.

origin with respect to the template. This insures that the conversion to exponential-polar coordinates will have the desired results.

The third and last step is the cross-correlation to determine the scale factor and the rotation. If the image has been scaled by a and rotated by α then the peak in the cross-correlation occurs at (b,β) where

$$a = e^{-b}$$

$$\text{and } \alpha = \beta.$$

The correlation coefficient can be calculated in order to assist in the match/no match decision. Again, the closer the value of the coefficient is to 1, the more likely it is that the peak is caused by a match between the template and the image.

The centroid method algorithm is relatively short and simple. It is:

1. Compute the centroids of the template and image, shift so the centroids are at the origin
2. Convert to exponential-polar coordinates
3. Cross-correlate to determine the scale factor and the rotation

Summary

This chapter has described algorithms for using generalized correlation when the template and the image are of a single object on a black background. Algorithms for both one-dimensional and two-dimensional pattern matching were discussed and developed.

CHAPTER V

APPLICATIONS OF GENERALIZED CORRELATION TO PATTERN MATCHING

The algorithms described in Chapter IV assume both the image and the template are of a single object on a black background. Under some circumstances it is possible to use these techniques as part of a pattern matching scheme when the image is not of a single object on a black background. The three schemes presented in this chapter consider images composed of multiple objects on a black background, a single object on a textured background, and a single object with additive noise.

Multiple Objects on a Black Background

The algorithms presented for computing generalized correlation were developed for the special case of a single object on a black background. These techniques can be extended to other cases under various circumstances. One case to which these algorithms are applicable is that of multiple objects on a black background.

The case of multiple objects on a black background can not be handled directly with either the magnitude method or the centroid method. The magnitude method fails because the Fourier Transform of the image is the sum of the Fourier

Transforms of the objects in the image. The magnitude is the square of the Fourier Transform and this causes the effects of each object to be mixed in such a way that the algorithm can not sort them out. The centroid method fails because the centroid of the image is unlikely to be at the centroid of an object.

These difficulties can be avoided by appropriate preprocessing of the image. The image can be divided into pieces where each piece is of a single object on a black background. Each piece can then be used as an image in the generalized correlation procedure. The problem of dividing the image into the appropriate pieces is a special case of image segmentation. There are several techniques available for segmentation including edge detection and boundary tracing, texture classification, and various types of feature extraction [5].

The only limitations on analyzing an image with multiple objects on a black background are imposed by the limitations of current algorithms to separate the objects. As the algorithms for object separation improve, this process will become more valuable.

Single Object on an Evenly Textured Background

Images of real objects are rarely on a black background (because any surface will reflect some light). This makes it desirable to find ways of using generalized correlation when the background surface is evenly textured. This surface must be approximately the same everywhere in the image in order for this analysis to be valid.

There are two approaches to this problem. The first approach is to process the image with the texture and determine under what conditions generalized correlation will work acceptably. The second approach is remove the object from the texture and then process the object. Both the magnitude method and the centroid method will be analyzed with each approach.

Processing with Textured Background

The methods this work discusses for computing generalized correlation will not always work when the image is of an object on a textured background. Through understanding why these techniques will not always work, an understanding of when they will work can be developed. The centroid method is not appropriate for images with textured background because the texture affects the location of the centroid. Consequently, when the centroid is shifted to the origin, there is no assurance that the centroid of the object is at the origin. In fact, the presence of background texture is virtually a guarantee the centroid method will fail.

The magnitude method is not as sensitive to textured backgrounds as the centroid method. The texture does affect the magnitude of the Fourier Transform, however under some conditions this can be thought of as noise. If the amplitude or brightness of the texture is much lower than that of the object, the portions of the magnitude attributable to the object will dominate the magnitude of the image. As long as this is true the generalized correlation will be approximately correct, but the error due to the texture will be reflected by lower values for the correlation coefficient.

When the components of the magnitude due to the object no longer dominate those due to the texture, the algorithm breaks down. This can be caught by the correlation coefficient because it will decrease in value as the effect of the texture increases. The result of this is computing generalized correlation can be done with images of a single bright object on a dark background.

Separation of Object from the Texture

The second approach for processing images of a single object on an evenly textured background is remove the object from the background. This can be done by using a texture classifier to determine where the evenly textured background ends and the object begins [6]. The extracted object is then placed on black background to be used in generalized correlation.

Unfortunately the texture classifiers that are currently available in general can not do a perfect job of separating the object from the texture. What generally happens is the extracted object still has some small portions of the texture and has lost some corners or protrusions of the original object. Next, the extracted object is placed on a black background to be used in generalized correlation.

The magnitude method will work well if there is very little texture and essentially all of the object present. As the quality of the extraction goes down the ability of the magnitude method to find the correct parameters will degrade. This is because the portions of the texture that are included as part of the object cause potentially severe distortions of the magnitude.

When applying the centroid method it may suffer if the centroid of the pieces of the object not extracted and the centroid of the pieces of texture added are not very close to the centroid of the original object. If these centroids are not close together, the extracted object will not be centered properly for the conversion to exponential-polar coordinates. In the cases when the centroid of the extracted object is appropriate, this method will work. The correlation coefficient must be checked to insure that it is possible to recognize when the degradations become severe.

The ability of these algorithms to produce meaningful results when separating the object from a textured background is dependent upon the techniques available to separate the objects from the texture. Both methods become less useful as the extracted object differs more and more from the original object.

Additive Noise

Generalized correlation in the presence of additive noise is equivalent to working with a single object on a textured background where the object has been degraded. The discussion of the algorithms when working with a single object on a black background apply here. The only difference is the correlation coefficient will be lower because the object has been degraded.

Summary

This chapter has described algorithms for using generalized correlation in two-dimensional pattern matching. Two basic algorithms were used, both based on separating the generalized correlation into sub-problems. The two algorithms were the magnitude method and centroid method described in Chapter IV. These were developed for a single object on a black background in one and two dimensions. They were then examined for use with multiple objects on a black background, a single object on an evenly textured background, and an image that has been degraded by additive noise.

CHAPTER VI

IMPLEMENTATION PROBLEMS AND COMPUTATIONAL METHODS

Implementing the computation of generalized correlation by the algorithms described in Chapters IV and V on a digital computer presents several problems. This chapter presents the implementation problems, why they arise and steps to be taken to solve them. Also included is a section on computational methods.

The problems that arise are as a result of the finite nature of a digital computer. This requires images to be sampled at a finite number of points. Sampling and truncation are the fundamental issues of concern. The problems to be discussed are:

1. The initial sampling of an image
2. The infinite extent of the exponential-polar coordinate system
3. The interpolation necessary to change coordinate systems

Sampling an Image

The need for sampling the image is forced by the discrete nature of the digital computer. The sampling theorem states that in order to correctly determine the function from its samples, the sampling frequency must be at least twice the highest frequency present in the function [7]. Generally the image will have to be low pass filtered before it is sampled to meet this condition. Fortunately, many digitizers low pass filter the image as they sample. If this criterion is not met, the digitized version of the image may not be interpreted correctly.

The Infinite Nature of Exponential-Polar Coordinates

During the discussion of the conversion from rectangular to exponential coordinates, it was noted that the range 0 to 1 is mapped into the range $-\infty$ to 0. This is caused by the change in variables $r=e^w$. Thus, after the change of variables the function has infinite spread, whereas before it has a finite spread.

As part of the coordinate change, it is important to insure that the integral of the function is invariant under this transformation. In Chapter III the following transformations were derived: in one dimension

$$f(r) \longrightarrow f(e^w)e^{w/2}$$

and in two dimensions

$$f(r,\theta) \longrightarrow f(e^w,\theta)e^w .$$

The solution to the problem of the infinite extent of the transform lies in the e^w and $e^{w/2}$ terms. The e^w and $e^{w/2}$ terms become extremely small very quickly as w becomes more negative. Truncating the function for $w<b$ where $b<0$ introduces an error E which is

$$E=\int_{-\infty}^{b} f(e^{w})e^{w/2}dw$$

Approximating $f(e^{w})$ over the range $-\infty<w<0$ (equivalent to $f(r)$ for $(0<r<1)$ with a constant $k=\sup(f(0),f(1))$ gives

$$E=\int_{-\infty}^{b} ke^{w/2}dw$$
$$=ke^{w/2}<k$$

Consequently, if b is sufficiently negative, the error introduced by truncating can be kept as small as desired. Similarly, it can be shown the error caused by truncating in the exponential-polar domain is less than $2\pi\, ke^{b}$.

This analysis indicates that the error introduced by truncating the exponential and exponential-polar coordinate representations of images can be made acceptable. A simple expression which bounds the error as a function of the value of w at which the image is truncated was derived above.

Interpolation

Two factors make it necessary to interpolate the function which represents the image. The fact that the image has to be sampled coupled with the need to change coordinate systems makes it necessary to determine values of the function at points between samples. Thus it is necessary to perform some type of interpolation. The ideal interpolation scheme will be discussed first, followed by descriptions of two practical interpolation schemes.

Ideal Interpolation

Interpolation can be considered as a convolution of the function with an interpolation kernel. It is well known that if the sampled function is band limited (not aliased) the proper convolution kernel is

$$\frac{\sin(\pi x/X)}{(\pi x/X)}$$

where the samples are X units apart [8]. This kernel is referred to as sin(x)/x or sinc(x). Similarly the two-dimensional convolution kernel is

$$\frac{\sin(\pi x/X)}{(\pi x/X)} \frac{\sin(\pi y/Y)}{(\pi y/Y)}$$

This interpolation kernel will perfectly recover any function which was properly filtered before sampling. To interpolate the function at a point (x,y) the following summation is used

$$f(x,y)=\sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(kX,jY)\frac{\sin(\pi(x-kX)/X)}{\pi(x-kX)/X} \frac{\sin(\pi(y-jY)/Y)}{\pi(y-jY)/Y}$$

The difficulty with sin(x)/x interpolation is the kernel has infinite support. Consequently, it can not be used since the computations need to be finite. This leads to the next section on interpolants actually implemented and used.

Practical Interpolation

Two interpolation schemes are implemented. They are bilinear and a windowed two-dimensional $\sin(x)/x$ interpolants. The bilinear interpolant is implemented because it uses a relatively small amount of execution time. This can be very important when interpolating an image which contains a large number of points. The windowed $\sin(x)/x$ is used to approximate ideal interpolation. By comparing the results of this latter interpolant with those of the bilinear scheme, an estimate can be obtained of the error introduced by the bilinear scheme.

Bilinear Interpolation

Linear interpolation is one of the simplest and most common interpolants used. The linear interpolant is

$$f(x)=f(m)+(x-m)(f(m+1)-f(m))$$

where $m<x<m+1$ [9]. Bilinear interpolation interpolates $f(x,y)$ by performing the following linear interpolations as shown in Figure 5:

1. Linearly interpolate $f(x,n)$ where $n<y<n+1$ and $m<x<m+1$
2. Linearly interpolate $f(x,n+1)$
3. Linearly interpolate $f(x,y)$ from $f(x,n)$ and $f(x,n+1)$

Although this interpolant can be easily computed, the artifacts introduced by this scheme are not always acceptable [8]. In order to determine the effect of these artifacts, a better interpolant is used for comparison.

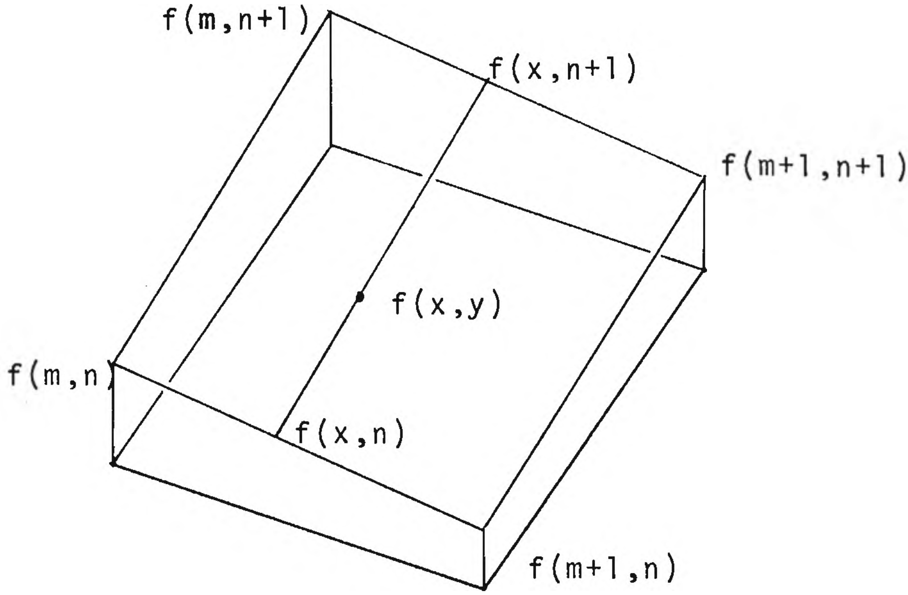


Figure 5. Bilinear interpolation.

Windowed Sin(x)/x

Ideally interpolation should be done using the sin(x)/x kernel, but the infinite extent of this kernel precludes this in practice. Truncating the sin(x)/x function produces a finite approximation that can be acceptable if the length of the truncation window is sufficient [9]. The same accuracy can be achieved with a shorter, but more sophisticated window [10]. For large windows the windowed sin(x)/x interpolant approaches the ideal sin(x)/x interpolant.

The one-dimensional interpolation formula for windowed sin(x)/x is

$$f(x)=\sum_{k=a}^{b} f(kX)\,\mathrm{sinc}(x-kX)\,W(x-kX)$$

where

$W(x)$ is the window function of length N,

a is the smallest integer $>x-N/2$,

b is the greatest integer $<x+N/2$

and X is the sampling interval.

In two-dimensions the interpolation formula is:

$$f(x,y)=\sum_{k=a}^{b}\sum_{j=c}^{d} f(kX,jY)\,\mathrm{sinc}(x-kX)\,\mathrm{sinc}(y-jY)\,W(x-kX,y-jY)$$

where

$W(x,y)$ is the window function of sized N by M,

a is the smallest integer $>x-N/2$,

b is the greatest integer $<x+N/2$,

c is the smallest integer $>y-M/2$,

d is the greatest integer $<y+M/2$,

and X and Y are the sampling intervals.

The windows used in this work were a one-dimensional Hanning window and a two-dimensional separable Hanning window. They are given by

$$W(x)=\frac{1}{2}\left[1-\cos\left(\frac{2\pi(x-N/2)}{N-1}\right)\right]$$

in one dimension, where N is the window length, and in two dimensions

$$W(x,y)=\frac{1}{4}[1-\cos(\frac{2\pi(x-N/2)}{N-1})][1-\cos(\frac{2\pi(y-M/2)}{M-1})]$$

where the window size is N by M. There are many other windows which may be used in place of the Hanning window with differing effects upon the accuracy.

Computational Methods

There are two items that need to be mentioned in regards to the computations. The first is the selection of the sampling frequency when converting to exponential-polar coordinates. The second is the technique used to compute cross-correlation.

Sampling in Exponential-Polar Coordinates

When converting from one coordinate system to another it is important that the errors introduced be minimized. One aspect of this minimization is insuring that the sampling frequency is chosen in such a way as to avoid aliasing. The way to avoid aliasing is to insure that when the samples for the exponential-polar coordinate system are placed in the rectangular grid they are never farther apart than the samples of the rectangular function. In terms of Figure 6 this means that the distance between any two adjacent radial lines,b, is never greater than the original sample spacing a. Also the distance between two adjacent samples on a radial line c, must not be greater than the original sample spacing a.

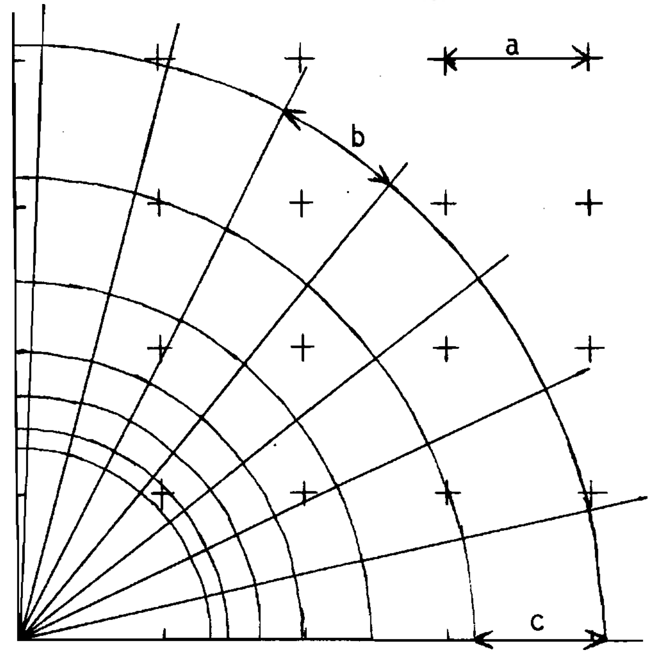


Figure 6. Exponential-polar coordinates on a rectangular grid.

The number of samples M, in the exponential domain, needed to avoid aliasing for a one-dimensional signal f(x), where f(x) is defined for 0<x<N, must be chosen such that

$$e^{M\Delta w} - e^{(M-1)\Delta w} = 1$$

and

$$e^{M\Delta w} = 1$$

where Δw is the exponential sampling interval. It can be shown that this gives $M \approx N \ln N$ [11]. For two-dimensional signals, the above is the appropriate sampling frequency radially, however the angular sampling frequency must still be determined. The angular samples are farthest apart at the maximum radius R. The number of angles at which samples must be taken, k is given by

$$k > 2\pi R.$$

There is no problem introduced in the changing of coordinate systems as long as the resampling rates are greater than or equal to the ones given above.

Computation of Cross-Correlation

Using direct summation the cross-correlation of two discrete functions of length N can be computed directly by a summation in a time proportional to N^2. Using Fourier Transforms to compute the correlation (as discussed in Chapter II) the time required becomes proportional to N ln N provided the transforms are implemented along the lines of the Cooley-Tukey algorithm [12]. Computing cross-correlations by direct summation of N by M images is proportional to N M , while using Fourier Transforms the proportionality is MN ln MN. To compute cross-correlations of sampled data this way requires using the Discrete Fourier Transform (DFT), which necessitates some precautions.

The cross-correlation c(x) of two functions f(x) and g(x) using Fourier Transforms is

$$c(x) = \mathcal{F}^{-1}(F(\omega)G(\omega))$$

where $F(\omega)$ and $G(\omega)$ are the Fourier Transforms of f(x) and g(x) respectively, and $\mathcal{F}^{-1}$ denotes the inverse Fourier Transform. When f and g are sampled functions, the discrete Fourier Transform (DFT) must be used. If interpreted correctly, this does not change the above technique for computing the cross-correlation. The DFT treats all signals as the principal period of a periodic function. The cross-correlation computed with the DFT is therefore a circular cross-correlation. Linear cross-correlation can be computed by doubling the length of f and g, by padding with zeros. This will allow the cross-correlation to avoid the periodic nature of the DFT. Implementing the DFT with a fast transform technique makes this method of computing the cross-correlation faster than direct summation.

The implementation of the algorithms for computing generalized correlation uses Fourier Transforms to compute cross-correlation. When correlating to find the translations in a two-dimensional image of size N by M, enough zeros must be added to make the DFT size 2N by 2M to insure the result is the linear cross-correlation. The cross-correlation in the exponential-polar coordinate space needs to be handled differently. In Chapter II it was explained how a rotation in rectangular coordinates map into a circular shift in exponential coordinates. Consequently, a circular correlation is needed along the angle axis. Linear correlation, however, is still needed radially. As a result, to cross-correlate two functions $f(r,\theta)$ and $g(r,\theta)$ of size N by M, zeros must be added to make the size 2N by M for the proper combination of linear and circular correlation. It is interesting to note that by using the Fourier Transform to compute correlation, the conversion from rectangular to exponential-polar coordinates and correlating using the Fourier Transform is equivalent to converting from rectangular to the usual polar coordinates and correlating using Mellin Transforms radially [11].

Summary

This chapter has dealt with a series of implementation considerations. The first section considered the issue of sampling the image correctly. The infinite extent of exponential-polar coordinate conversion forces truncation. It was shown that the error introduced can be made arbitrarily small. Since the coordinate conversion forces resampling, the issues involving interpolation were analyzed. The last section discussed sampling frequency in the exponential-polar coordinate system and computation of cross-correlation using Discrete Fourier Transform (DFT).

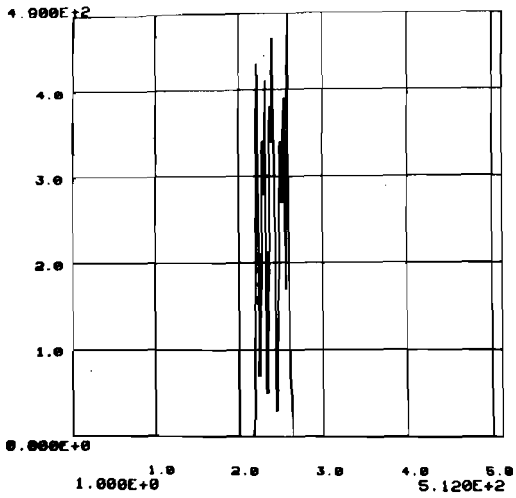
CHAPTER VII

RESULTS

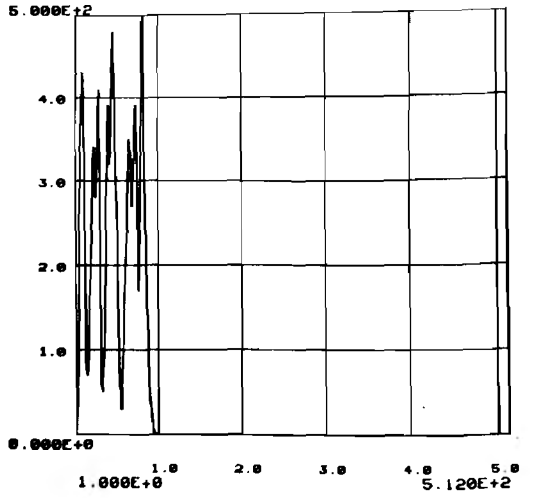
This chapter describes some results of pattern matching done by using the generalized correlation algorithms presented. Results were obtained in both one and two dimensions. The first section discusses an example of one-dimensional generalized correlation. The second section presents examples of pattern matching in two-dimensions.

One-Dimension

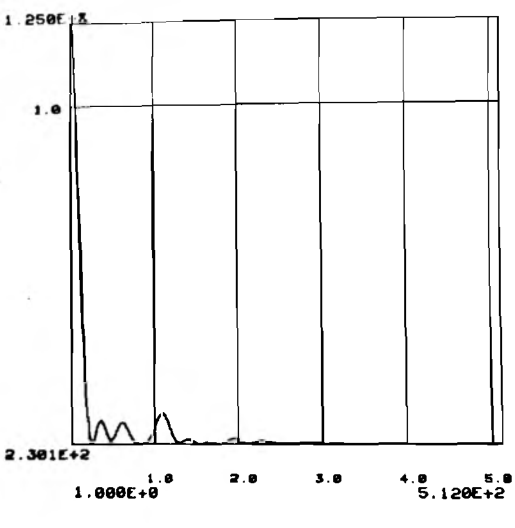
Figure 7 is a sequence showing the image function and the template function and their correlations at various steps in computation of generalized correlation. The computation was done by the magnitude method as described in Chapter IV (and outlined in Figure 2). The image in this example is one half the size of the template. Part (a) of Figure 7 is the image while part (b) is the template. The first step in computing the generalized correlation is to take the magnitude of the Fourier Transform of the image and the template. These magnitudes are shown in parts (c) and (d). The functions are then put into the exponential coordinate domain with the result shown in parts (e) and (f). Notice that the two functions do in fact appear to differ only by a translation. These two functions were then



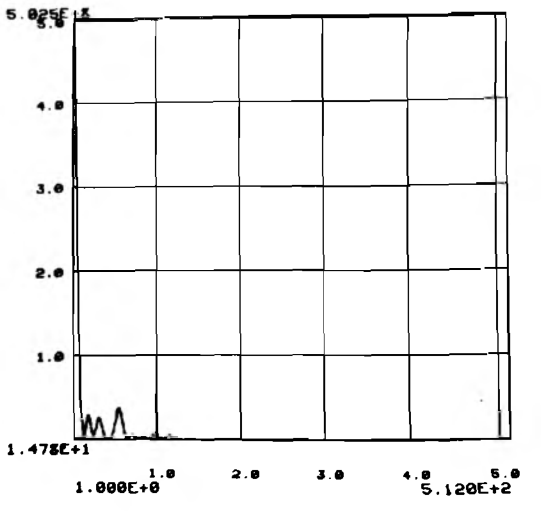
(a) Image



(b) Template

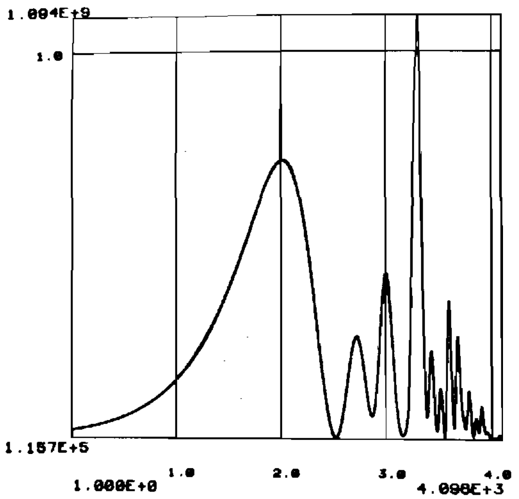


(c) Magnitude of image

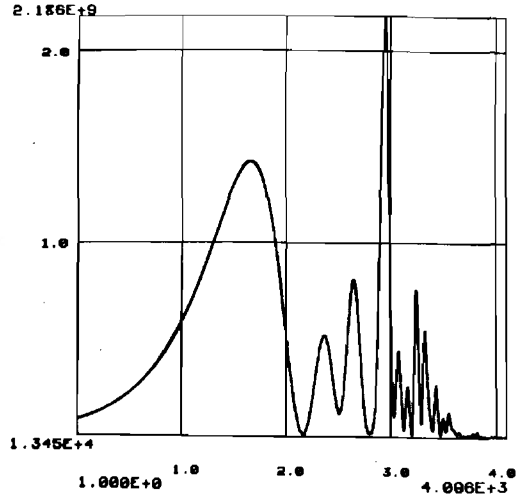


(d) Magnitude of template

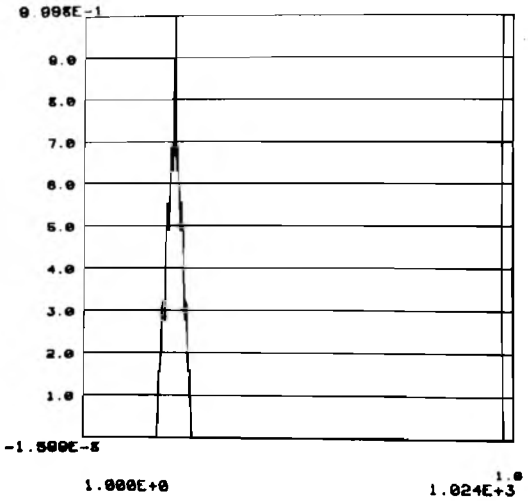
Figure 7. Generalized correlation in one-dimension. Magnitude method.



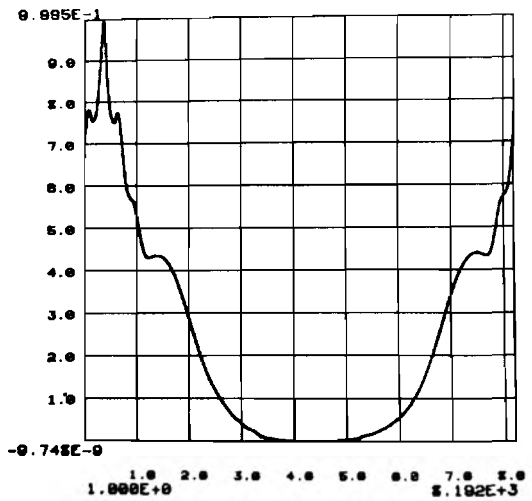
(e) Image magnitude in exponential coordinates



(f) Template magnitude in exponential coordinates



(g) Scale factor correlation



(h) Translation correlation

Figure 7. (continued)

cross-correlated giving the correlation function shown in part (g). The peak in the correlation function gives the translation in exponential coordinates which corresponds to the scale factor. A scaled template was then created and cross-correlated with the original image to get the translation, as shown in part (h). In this correlation the peak gives the relative translation between the image and the template. Table 1 summarizes the results of this experiment. The computed scale factor of .499 is as close the the actual value as can be done without careful interpolation of the correlation function. This is because the correlation is a discrete function.

In the above example, all interpolations were done using linear interpolation. The same algorithm and functions were tried using a windowed sin(x)/x interpolant. The results were identical, indicating that the error introduced by the linear interpolant was small with respect to the other sources. In conclusion, the above results strongly support the validity of the techniques developed.

TABLE 1

RESULTS OF ONE-DIMENSIONAL GENERALIZED CORRELATION

	Actual Value	Computed Value	Correlation Coefficient
Scale	.5	.499	.9995
Translation	218	218	.9998

Two-Dimensions

Two methods of computing generalized correlation in two-dimensions were discussed in Chapter III. Both are illustrated here with the same image and template in order to demonstrate the differences. The magnitude method is shown first, followed by the centroid method example.

Three examples are shown using both the magnitude and the centroid methods. The first example is referred to as "Patch" since it is composed of bicubic patches. The second example is the same as the first with a different scale factor and rotation. Since the angle of rotation is 90 degrees it will be referred to as "Ninety". In the third example the object is cross shaped hence its name is "Cross". Table 2 summarizes the relationships between each image and its corresponding template. When reading these results, it is important to remember the images and the

TABLE 2

SCALE AND ROTATIONAL RELATIONSHIPS BETWEEN IMAGES AND TEMPLATES IN TWO-DIMENSIONAL GENERALIZED CORRELATION EXAMPLES

Name of Image	Scale Factor	Rotation in Radians
Patch	.25	2.5
Ninety	.5	1.57
Cross	1.28	1.57

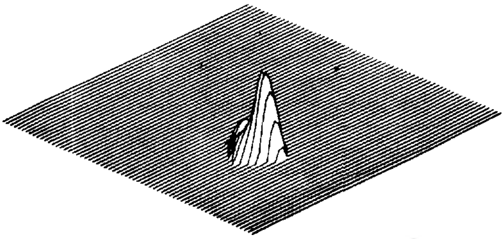
templates are described on a 64 by 64 point grid. This relatively coarse grid causes a large amount of information to be lost.

Magnitude Method

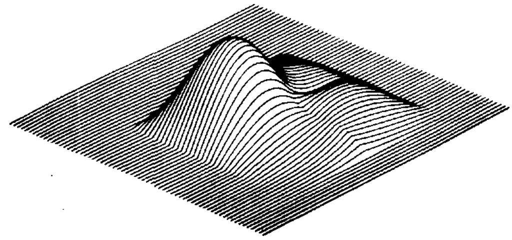
Figures 8, 9, and 10 illustrate several of the steps in computing the generalized correlation by the magnitude method for the Patch, Ninety and Cross images respectively. The image and the template are shown in parts a and b. The first step of the algorithm is take the Fourier Transform of the image and the template. The magnitudes are pictured in parts c and d. The functions which result from the conversion to exponential-polar coordinates constitute parts e and f. The cross correlation to determine the scale factor and the rotation is shown in part g.

Table 3 summarizes the results of the experiments. The translations have been omitted from the table for clarity and since the emphasis of this work is on correlation for scale and rotation.

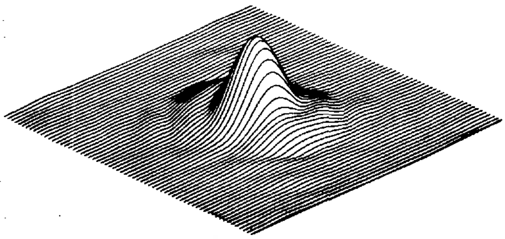
The Patch proved to be a very difficult example for two reasons. First, the scale factor of .25 on a 64 by 64 grid generates an extremely small image. The amount of information avaliable about the image is therefore quite small. The second difficulty is that the magnitude of the Fourier Transform is close to being circularly symmetric for the lower frequencies. Most of the information which indicates that the function is not circularly symmetric is



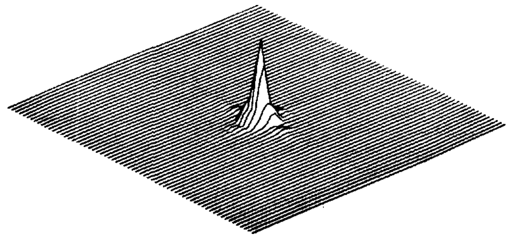
(a) Image



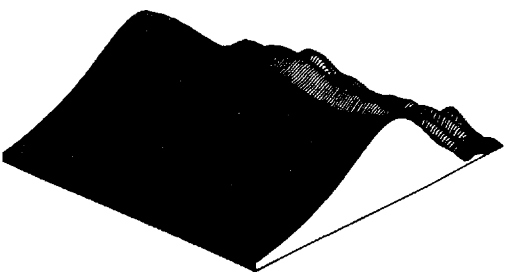
(b) Template



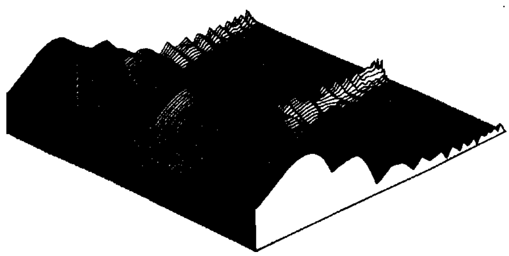
(c) Magnitude of image



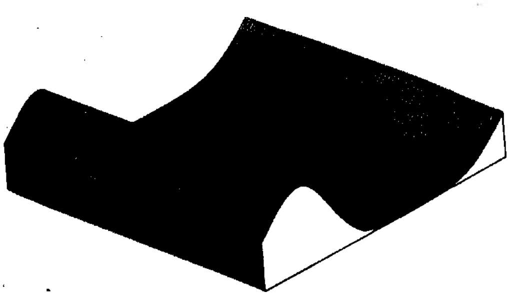
(d) Magnitude of template



(e) Image magnitude in exponential-polar coordinates

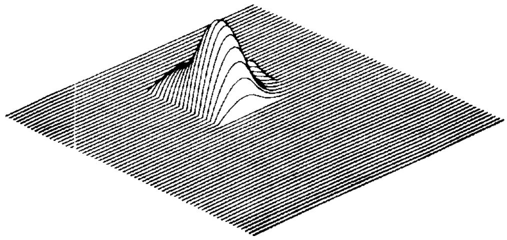


(f) Template magnitude in exponential-polar coordinates

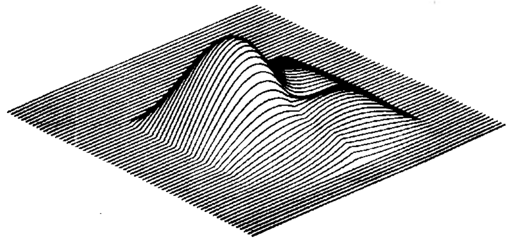


(g) Scale and rotation correlation

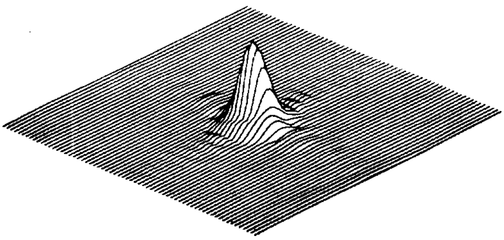
Figure 8. Generalized correlation in two-dimensions. Magnitude method for Patch image.



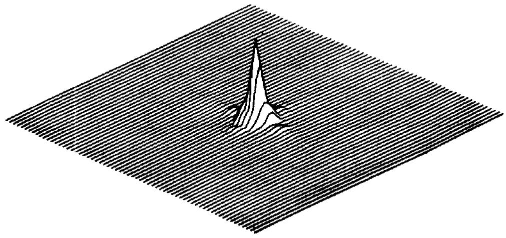
(a) Image



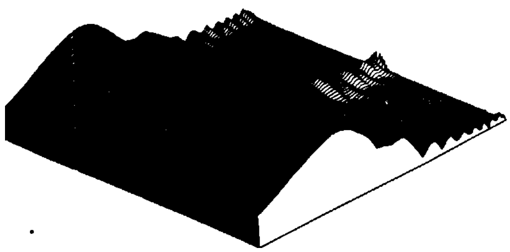
(b) Template



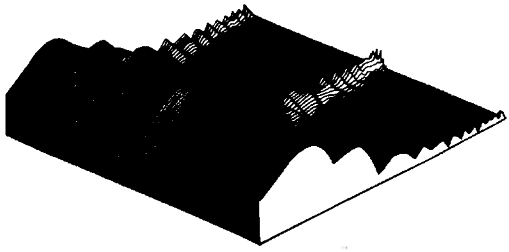
(c) Magnitude of image



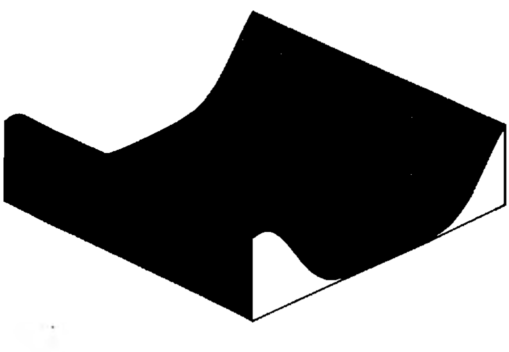
(d) Magnitude of template



(e) Image magnitude in exponential-polar coordinates

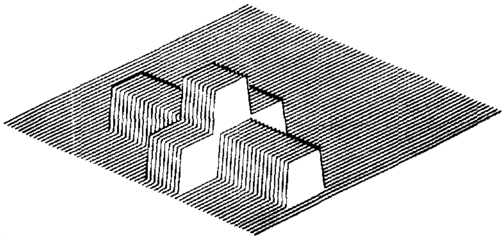


(f) Template magnitude in exponential-polar coordinates

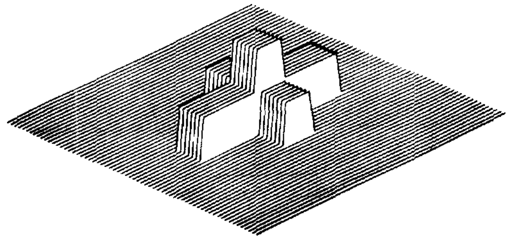


(g) Scale and rotation correlation

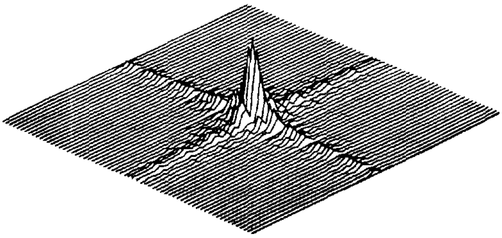
Figure 9. Generalized correlation in two-dimensions. Magnitude method for Ninety image.



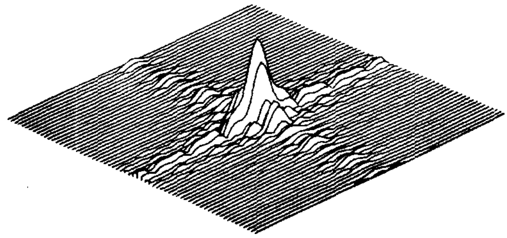
(a) Image



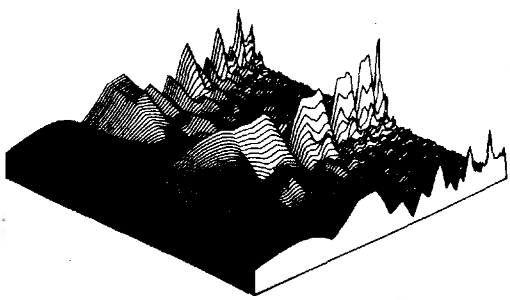
(b) Template



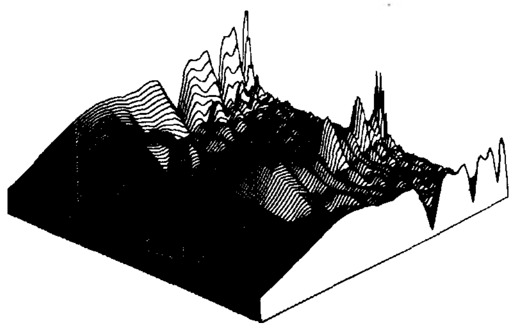
(c) Magnitude of image



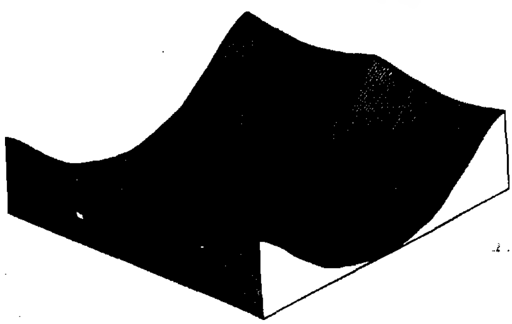
(d) Magnitude of template



(e) Image magnitude in exponential-polar coordinates



(f) Template magnitude in exponential-polar coordinates



(g) Scale and rotation correlation

Figure 10. Generalized correlation in two-dimensions. Magnitude method for Cross image.

TABLE 3

RESULTS OF TWO-DIMENSIONAL GENERALIZED CORRELATION

Name of Image	Actual		Computed by Magnitude Method			Computed by Centroid Method		
	Scale Factor	Rotation in Radians	Scale Factor	Rotation in Radians	Corre-lation Coeff.	Scale Factor	Rotation in Radians	Corre-lation Coeff.
Patch	.25	2.5	.293	1.76	.951	.275	2.38	.966
Ninety	.5	1.57	.778	3.09	.951	.500	1.57	.994
Cross	1.28	1.57	1.24	1.57	.966	1.28	1.57	.997

in the phase which is not used in this method. Since the image is so small, the lower frequencies affect the correlation more than the higher frequencies, where the effects of the rotation are more pronounced.

The Ninety example uses the same object as the Patch with a different scale factor and rotation. In this case the difficulty is again the almost circularly symmetric nature of the magnitude of the function. The Cross is an example where the magnitude method works well. In this example there is little circular symmetry in the magnitude. This is largely due to the discontinuities or sharp edges in the original function.

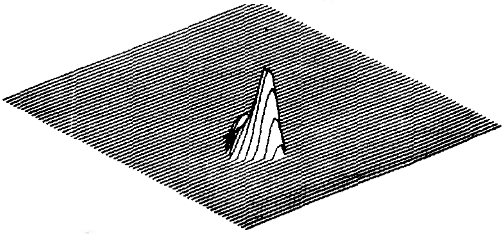
The above discussion is primarily concerned with the difficulties presented by each example. However, there are two ways major ways in which significant errors are introduced into the calculations. One source of error, which mainly affects the scale factor, is the truncation of the function in exponential-polar coordinates. This is very apparent in the Patch example and probably contributed to the poor results. The functions were truncated too close to the origin as indicated by the large value of the function where it was truncated.

The second source of error is the interpolation scheme used. A bilinear interpolant was used and its effect can be seen in in parts (e) and (f) of all three examples. The artifacts introduced by this interpolant contribute to the poor determination of the angle of rotation. In the Cross example, the scale factor was close to 1, and with the rotation being 90 degrees, the interpolant treated the image and the template almost the same.

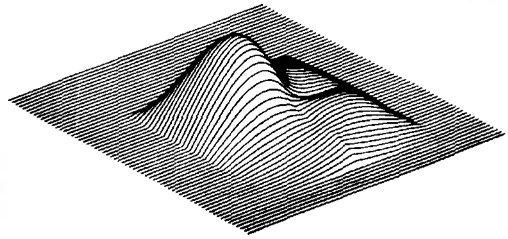
Centroid Method

The sequence of two-dimensional functions in Figures 11, 12, and 13 illustrate steps in computing generalized correlation using the centroid method for the Patch, Ninety and Cross images. The image is shown in part (a) with the template in part (b). After the location of the centroid of the image was determined, the image was shifted so the centroid was at the origin as shown in part (c). This was done with the assumption the centroid of the template was at the origin. It should be noted that this assumption need not be made since the template can also be shifted to bring its centroid to the origin. Both the image and the template were then converted to exponential-polar coordinates as shown in parts (d) and (e). Lastly, these two functions were cross-correlated to determine the scale factor and rotation, giving part (f).

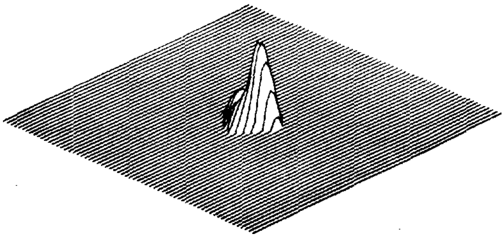
The results of these three experiments are also summarized in Table 3. In the Patch example, the centroid method, like the magnitude method, suffered from the large scale factor on a small grid. Other than this one problem, the centroid method did extremely well, finding the correct scale and rotation in both the Ninety and the Cross



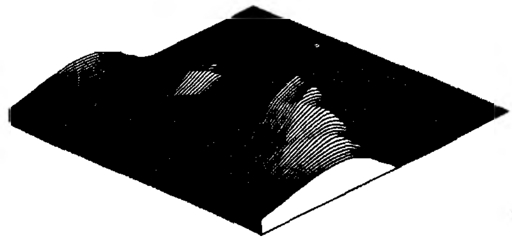
(a) Image



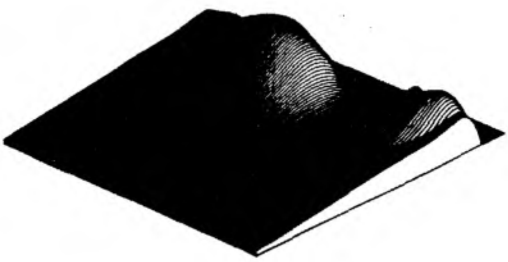
(b) Template



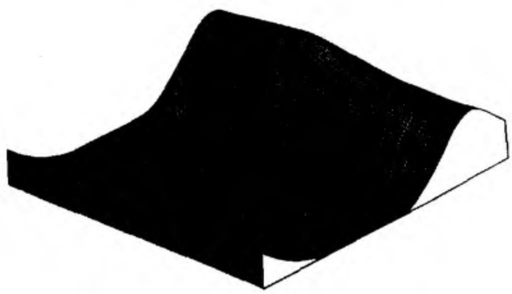
(c) Image with centroid at origin



(d) Image in exponential-polar coordinates

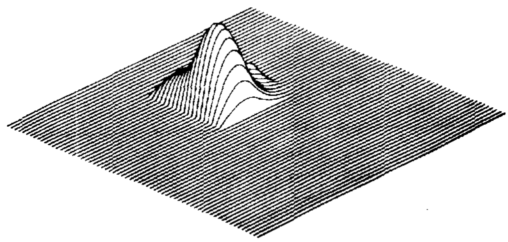


(e) Template in exponential-polar coordinates

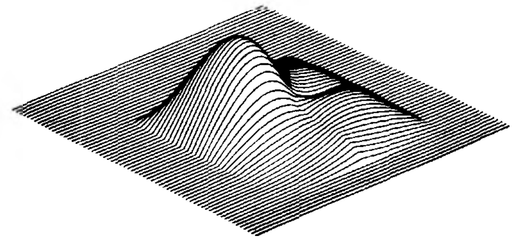


(f) Scale and rotation correlation

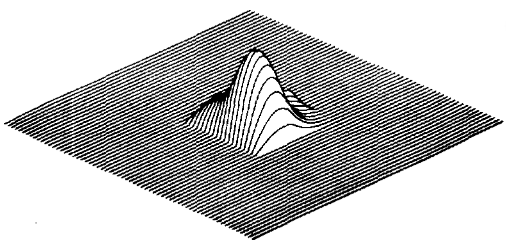
Figure 11. Generalized correlation in two-dimensions. Centroid method for Patch image.



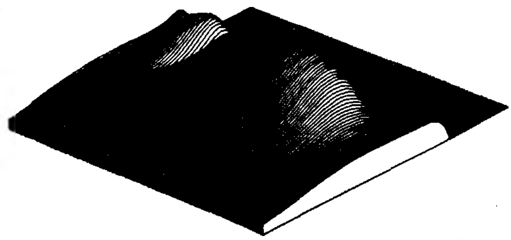
(a) Image



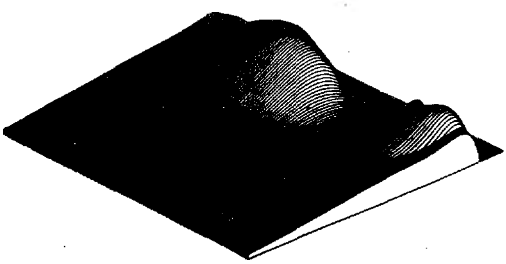
(b) Template



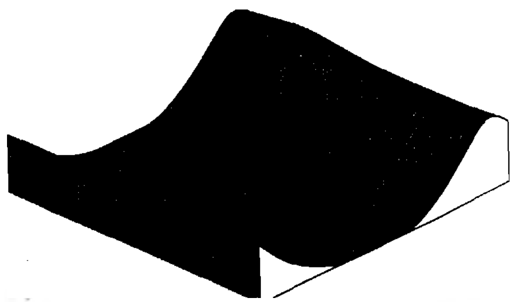
(c) Image with centroid at origin



(d) Image in exponential-polar coordinates

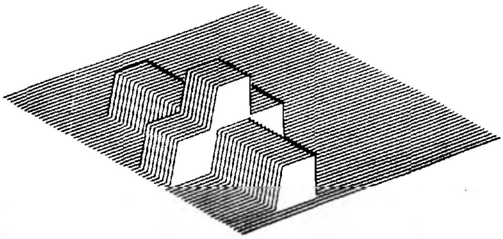


(e) Template in exponential-polar coordinates

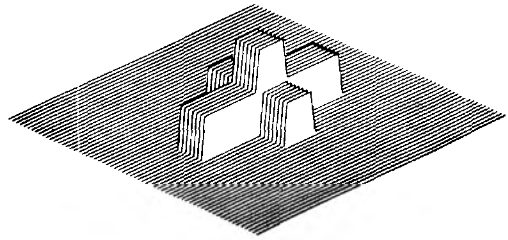


(f) Scale and rotation correlation

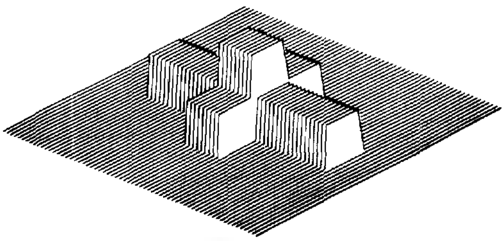
Figure 12. Generalized correlation in two-dimensions. Centroid method for Ninety image.



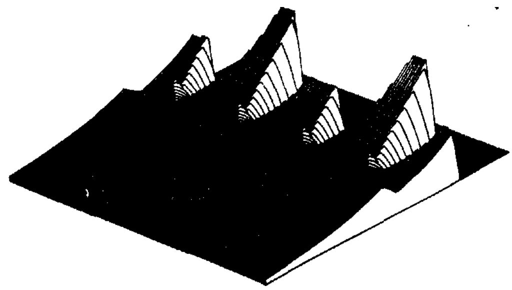
(a) Image



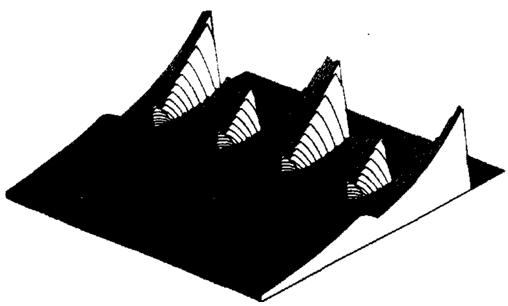
(b) Template



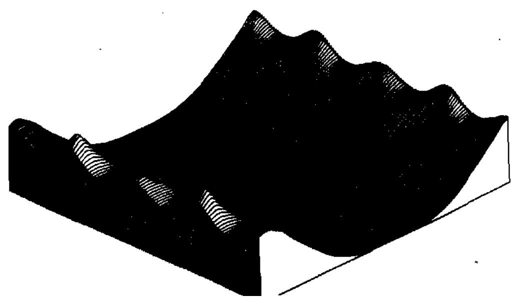
(c) Image with centroid at origin



(d) Image in exponential-polar coordinates



(e) Template in exponential-polar coordinates



(f) Scale and rotation correlation

Figure 13. Generalized correlation in two-dimensions. Centroid method for Cross image.

examples. In the Cross example, notice the four peaks in the correlation (part f), two of which are higher than the other two. The two higher peaks correspond to the two possible rotations of the symmetric object. Again, there were errors introduced from the truncation in exponential-polar coordinates and from interpolation errors. An examination of the functions in exponential-polar coordinates indicates that only the image in the Patch example has very noticeable interpolation and truncation errors.

Comparison of Magnitude and Centroid Methods

Looking at Table 3 it is obvious that the centroid method is more reliable than the magnitude method. This, combined with the greater ease of computation, makes the centroid method more attractive than the magnitude method. Unfortunately, the centroid method is probably more sensitive to noise and the presence of unwanted texture. The reliability of the magnitude method can be increased by increasing the number of samples in the image.

Summary

This chapter has presented results demonstrating the techniques developed in this work. Both the one-dimensional example and the two-dimensional centroid examples work as expected and determine correctly the relationship between the image and the template. The two-dimensional magnitude method works acceptably in only some cases because a large amount of information is lost in removing the phase.

CHAPTER VIII

CONCLUSIONS

The use of generalized correlation in the area of pattern matching was investigated. Two techniques were developed assuming the images were of a single object on a black background. These techniques were demonstrated to work well in the above case. The possibilities of extending these techniques to images that are more complex than a single object on a black background were also discussed.

Pattern matching is becoming more widely used and needed in a variety of fields. Some of these include monitoring systems, identification systems, inspection of objects (for quality control) to name but a few. Because of the prospective growth of pattern matching, it is desirable to have basic pattern matching algorithms on which to build.

Experimentation is needed to determine to what degree the methods presented for computing generalized correlation can be extended for images that are not of a single object on a black background. Experiments with higher resolution images may indicate that the techniques developed can be of greater value than indicated by the results of the limited experimentation presented here. Several other types of images were examined theoretically, namely a single object background, but these types of images need to be further examined experimentally. Other types of images, such as multiple objects on a textured background need to be considered both theoretically and experimentally.

The magnitude method suffers because all phase information is removed. If only the linear phase components could be removed then the algorithms could be simplified and made more reliable. This involves the familiar problem of phase unwrapping, hence it may not be computationally reasonable.

The slowest step in the algorithms presented is the conversion from rectangular coordinates to exponential-polar coordinates. If some way could be found to determine the same information without the extensive resampling currently required, the process could be significantly accelerated.

While the previous two suggestions would help make these algorithms more practical, the limitations imposed by the techniques used to separate the generalized correlation remain the biggest problem. New techniques for separating the generalized correlation which do not depend on the image to be of a single object on a black background need to be developed. Research in this area may provide the algorithms to make generalized correlation a powerful and useful tool in pattern matching.

LIST OF REFERENCES

1. John E. Freund, Modern Elementary Statistics, New Jersey, 1973.

2. F. H. Lange, Correlation Techniques, Princeton, 1967.

3. H. C. Andrews, Computer Techniques in Image Processing, New York, 1970.

4. W. F. Trench and B. Kolman, Multivariable Calculas with Linear Algebra and Series, New York, 1972.

5. A. Rosenfeld, Picture Processing by Computer, New York, 1969.

6. Brent Baxter, Personal Communication.

7. A. Papoulis, The Fourier Integral and its Applications, New York, 1962.

8. R. W. Schafer and L. R. Rabiner, A Digital Signal Processing Approach to Interpolation, Proc. IEEE 61 No. 6 (1973), 692-702.

9. A. F. H. Goetz et. al., Application of ERTS Images and Image Processing to Regional Geologic Problems and Geologic Mapping in Northern Arizona, JPL Tech. Report 32-1597, (1975).

10. R. W. Schafer and A. V. Oppenheim, Digital Signal Processing, New Jersey, 1975.

11. G. M. Robbins and T. S. Huang, Inverse Filtering for Linear Shift-Variant Imaging Systems, Proc. IEEE 60 No. 7 (1972), 862-872.

12. T. G. Stockham, Jr., High Speed Convolution and Correlation, AFIPS Proc. 28 (1966), 439-441.

ACKNOWLEDGEMENTS

I wish to express my appreciation to several of the people who assisted me at various stages of this work. I would like to thank Dr. Martin Newell for his assistance and guidance throughout the course of this investigation. I am also in debt to Jim Kajiya for inspiration and assistance with the mathematical formulations, to Dr. Brent Baxter for considerable help in implementing the algorithms, and to Mike Milochik for the photographic work. In addition without the support and assistance of my wife Cyndy, this work would never have been completed.