OBJECT RECOGNITION USING ANGLES IN THE PROJECTIVE PLANE

Alper Yilmaz\(^a\) and Gabor Barsai\(^b\)

\(^a\) The Ohio State University, Dept. of Geodetic Science, Photogrammetric Computer Vision Laboratory

The Ohio State University -
yilmaz.15@osu.edu

\(^b\) Gotmaps? Inc., 2981 Wicklow Rd., Columbus, OH - gabor@gotmaps.biz

ABSTRACT

In this paper, we propose a recognition technique for objects which are represented as closed contours. Our algorithm relies on the angular similarity between the contours and exploits the properties of the Fourier Transform, plane projective geometry and the wavelet transform. An important contribution of the proposed method is that it can recognize objects acquired from any viewing direction (such as oblique) and even when the objects are partially occluded, without finding any conjugate points. The experimental results on an in-house developed geographic database, and the Brown University object database show robust recognition performance.

KEYWORDS: image understanding, 2-D projective geometry, object recognition

INTRODUCTION

The ever increasing amount of digital data available from private companies, governments and the internet, suggest that the development of an automated recognition system is of utmost importance, in part due to the increased stress levels (Schabracq et al., 2001) and eye fatigue (Wang and Huang, 2004) of human operators working on such large datasets. Automated object recognition tools are commonly used in various domains including mobile mapping (Schenk, 2001), wide area surveillance (Yilmaz et al., 2004), analysis of temporal changes (Yilmaz et al., 2006) and classification and retrieval from datasets (Zhang and Lu, 2001).

Object recognition is achieved by matching individual features between various datasets, which may be composed of images, maps or object models and GIS data. Matching between the objects in these datasets can be defined as the establishment of correspondences; for example, objects extracted from a digital image are matched against digital data in a GIS. For a human operator, finding conjugate points, lines or objects is a relatively easy task within these data sets. For computer algorithms, however, recognition is a complex problem, due to the fact that partial occlusions, viewpoint changes, varying illumination, cluttered backgrounds, and intra-category appearance variations all make it necessary to develop exceedingly robust matching methods. According to Meyer (Meyer, 1993), the ultimate goal in this regard is to investigate autonomous object recognition in unconstrained environments. A recognition algorithm has to search among the possible candidate object representations to identify the best match and then determine whether the object label assigned to the entity in the image is indeed correct or incorrect (Jain and Dorai, 2000). This is a challenging task, since recognition is affected by: (i) the type of sensors used, (ii) the kinds of features extracted from an image, (iii) the class of objects that can be handled, (iv) the approaches employed to hypothesize possible matches from the image to object models, (v) the conditions for ascertaining the correctness of hypothesized matches, and (vi) the techniques to estimate the pose of the object. Among the various factors that directly affect the accuracy of a recognition algorithm, the object representation and matching procedure considerations are of high importance, since they influence the ease with which objects are recognized (Jain and Dorai, 2000).

Approaches vary in terms of how the match between the extracted and model features is achieved. Many matching strategies in use still rely on simple transformations, such as translation, rotation and scaling, without even taking into consideration occlusion and appearance variations. Recognition under more general transformations, such as affine and projective transform, or appearance variations, however, has not been fully examined, due to their complex nature. Matching methods can be grouped according to area, feature or symbolic based as illustrated in Table1 (Schenk, 2001).

Area based matching is associated with gray values. The gray value distribution of two areas (patches) is compared by correlation or least squares method. The disadvantage of this method is that it is very rigid, and does
not work well for different viewpoints, although it has been shown to work for rectified images (Schenk, 2001, chap. 10). This is similar to template matching, where a known, existing object in the image is searched for a given template that matches the query object. Feature based matching compares features extracted from the objects in various images (using edge detection, interest points or other methods) to find conjugates. The central principle underlying feature modeling is the observation that many objects can be represented by a small number of parts arranged in a characteristic configuration. The match, like strength of an edge or curvature, is measured as a cost function (Chawathe, 2007). While this method works for rectified images, but has similar problems as area based matching when using images with different viewpoints. The disadvantages of feature modeling are that it can become very computationally intensive for complex objects, or for objects with many points; along with the inaccuracies from only line correspondences (Zhang and Hans, 2002). Symbolic methods derive values from the image or digitized data: moments, wavelets, fractals or other attributes to calculate a matching score. This discussion suggests matching is not a well-posed problem. Some methods work for some cases, others for other cases; many times, due to occlusion there is no solution. A uniform matching and representation strategy that can handle all object categories and instances is unlikely to succeed for now.

Table 1. Matching method and entities

<table>
<thead>
<tr>
<th>Method</th>
<th>Measure</th>
<th>Entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>correlation, least squares</td>
<td>gray values</td>
</tr>
<tr>
<td>Feature</td>
<td>cost</td>
<td>lines, regions</td>
</tr>
<tr>
<td>Symbolic</td>
<td>cost</td>
<td>other (parameters, graphs...etc.)</td>
</tr>
</tbody>
</table>

The cost functions for symbolic matching listed in Table 1 can be stochastic or deterministic. Stochastic systems try to model variations using some form of probabilistic method, since the members of a class of objects are not exactly identical to each other. They are similar, have similar features, but are not exactly the same, due to noise and intra class variability. These methods can range from simple Bayesian classifiers to complex neural networks (Duda et al., 2000). Deterministic algorithms use different matching techniques including but not limited to: parametric methods (Lin et al., 2006), moment based (Hu, 1962) and structure based (Zhang et al., 2006) approaches. The advantage of the representation of curves/contours parametrically is that it gives a continuous function, even for discrete data, like digitized GIS data. Parametric representation involves the mathematical representation of a discrete signal, in this case the contour, and the algorithmic operations carried out on it to extract or enhance the information contained in the object of interest. Invariant moments are widely used for pattern recognition due to their discrimination power and ease of use; however, they do have disadvantages. The nth moment of a real-valued function \( f(x) \) of a real variable about a value c is

\[
\mu'_n = \int_{-\infty}^{\infty} (x - c)^n f(x) \, dx.
\]

The advantages of moments (easy to calculate) may be outweighed by their disadvantages (not intuitive) (Teague, 1980) (Zhang and Lu, 2001). In particular, it is difficult to correlate high-order moments with one of these object objects (DeValois and DeValois, 1980). A further problem is the dramatic increase in complexity with increasing order, since they are not derived from orthogonal functions and much redundant information about an object is created (Shen and Ip, 1998) (although orthogonality is not necessary for signals, and in many cases redundancy can reduce the sensitivity to noise). Many excellent papers have been written on moments up to affine transform (Hong and Zhang, 2005) (Li et al., 2006) Flusser (Flusser and Suk, 1994). Mercimek (Mercimek et al., 2005) proposes an idea for moment utilization using neural networks for 3-D object recognition, and Flusser (Flusser and Suk, 2004) and (Sivaramakrishna and Shashidharf, 1997) give examples of moment based recognition under projective transform.

Structural information involves attempting to code patterns. It implies that visual patterns can be encoded in a formal language (Leeuwenberg, 1969). Using structural information is not efficient when compared to parametric methods, and structural methods, especially those using graph-like representations, usually lead to variants of the computation-intensive graph isomorphism algorithm (Shapiro, 1979) (van Otterloo, 1991) (Zhang and Lu, 2001). Other structure based methods are from Belongie (Belongie et al., 2002), who develop a measure called object context for comparison. Shapiro (Shapiro, 1979) writes about the structure of object in and early work, about how objects can be defined, and compares different structure methods. For an extended introduction of any of these
techniques, please refer to any of the several survey papers by DeValois and DeValois (DeValois and DeValois, 1980), van Otterloo (van Otterloo, 1991), Loncaric (Loncaric, 1998) and Veltkamp (Veltkamp, 2001).

In this paper, we study the problem of matching the outlines of objects, which we will refer to as 2-D contours, using parametric method. One of the reasons for the popularity of contour-based analysis techniques is that edge detection constitutes an important aspect of object recognition by the human visual system (van Otterloo, 1991). The main motivation behind our work is that removing projectivity from objects may overcome the problem of noise sensitivity and boundary variations. This paper is a continuation of the research that appeared at the ISPRS (International Society for Photogrammetry and Remote Sensing) conference in Beijing, China, 2008. For this study, a Brown University database was used, along with several images of countries, lakes and other objects that were digitized from maps, satellite images or silhouettes to obtain the contours. We give a comparison to moment based matching using Hu’s seminal 1962 paper (Hu, 1962), combined with affine and projective moment invariance methods from Flusser’s papers (Flusser and Suk, 2004) (Flusser and Suk, 1994), with our method.

We match two objects in a deterministic formulation, by calculating the similarity score based on the comparison of the interior angles of the contours. We divide the task into two parts: we represent the contour based on Fourier descriptors, and based on these descriptors, we calculate the major and minor axes of the target and candidate objects, then change the height/width aspect ratio for the various objects to be one along the major/minor axis; next, based on wavelet software provided by (Torrence and Compo, 1998), we calculate a matching score using the wavelet power spectrum of the magnitude of the interior angles. A prominent characteristic of this method is that the comparison is scale, rotation and starting point independent. For this study, a Brown University database was used, along with several images of countries, lakes and other objects that were digitized from maps, satellite images or silhouettes to obtain the contours. The paper is organized as follows: Section 2 describes the projective geometry for a background on angular variation depending on the viewpoint in the projective plane; Section 3 outlines the methodology used: converting the x and y coordinates into periodic functions for use by the Fourier transform, leading to the location of the major/minor axes on each (candidate and target) contour, and computing the wavelet power spectrum for each contour. The major/minor axes found on each contour allows setting up distortion free (up to affine) objects. Section 4 shows the robust recognition results.

ANGLES IN THE PROJECTIVE SPACE

In Euclidean geometry the angle between two lines is computed from the dot product of their normals. Using homogeneous representation for the lines \( l = (l_1, l_2, l_3)^T \) and \( m = (m_1, m_2, m_3)^T \) with normals parallel to \( (l_1, l_2)^T \) and \( (m_1, m_2, m_3)^T \) respectively, the angle between them is

\[
\cos(\theta) = \frac{l_1 m_1 + l_2 m_2}{\sqrt{\left( l_1^2 + l_2^2 \right) \left( m_1^2 + m_2^2 \right)}}. \quad (1)
\]

The problem with this expression is that the first two components of \( l \) and \( m \) do not have well defined transformation properties under projective transformations (they are not tensors), and eq. (1) cannot be applied after a projective transformation of the plane. However, an analogous expression to eq. (1) which is invariant to projective transformations is given by:

\[
\cos(\theta) = \frac{l^T C_\infty^* m}{\sqrt{(l^T C_\infty^* l)(m^T C_\infty^* m)}}. \quad (2)
\]

where \( C_\infty^* \) is the conic dual to the circular points. In a Euclidean coordinate system eq. (2) reduces to eq. (1). It may be verified that eq. (2) is invariant to projective transformations by using the transformation rules for lines \( l' = H^{-T} l \) and dual conics \( C'^* = HC^*H^T \) under the point transformation \( x' = Hx \) (Hartley and Zisserman, 2000). From this, we can conclude that \( \cos(\theta) \) is constant if the transform parameters are known. If the transform parameters are not known, there is an apparent change in \( \cos(\theta) \) (an angle can appear to change from acute to obtuse depending on the viewpoint), but then the curvature type (convex or concave) is constant: no matter how oblique the image is, the convexity or concavity will not change. We will use this observation to compare interior angles without knowing the
transformation parameters. Another noteworthy observation is that although an affine transform does change the apparent size of the angles, this change can be "undone" without knowing the parameters, since the scale factor of the homogenous coordinates involved does not change. The scale factor stays at 1, and thus each affected angle changes in a similar manner. This is not the case in the event of a projective transform, where each angle changes in a different manner, due to the non-linear scale change of the homogenous scale.

**REPRESENTING THE OBJECT**

Once an object is detected, the next step involves the extraction of shape properties and spatial relations (Ullman, 1996). We represent objects as 2-D contours, or outlines. While establishing correspondences across images is feasible and has been a common practice among researchers, finding corresponding points between two contours is not easy and at times is impossible. Due to this observation, a common tradition has been to represent the contours in parametric form prior to any processing. Parametric contour representation can be generated by methods including but not limited to: curvature/spline based, polar coordinate (geometric) based, trigonometric based, wavelet transform based, Fourier descriptor (FD) based. In contrast to others, FD and wavelet based parametric contour representations provide continuous functions. Compared to the FD based methods, wavelet based methods, however, involve intensive computation, and it is usually not clear which basis would be a better choice to represent the contour (Zhang and Lu, 2001). Let two periodic functions \( x_n \) and \( y_n \) describe the contour, such that we treat \( x \) and \( y \) coordinates as independent dimensions. The Fourier series representation of an object contour is obtained by taking the discrete Fourier transform of the sequences \( x \) and \( y \) (Zahn and Roskies, 1972):

\[
\begin{align*}
x_n &= \sum_{k=0}^{N} G_k e^{2\pi i k n/N} = \sum_{k=0}^{N} d_k \cos(2\pi kn/N) + \sum_{k=0}^{N} d_k \cos(2\pi kn/N), \\
y_n &= \sum_{k=0}^{N} H_k e^{2\pi i k n/N} = \sum_{k=0}^{N} c_k \cos(2\pi kn/N) + \sum_{k=0}^{N} d_k \sin(2\pi kn/N),
\end{align*}
\]

where \((G_k, a_k, b_k)\) and \((H_k, c_k, d_k)\) are the coefficients for the Fourier transforms of \( x \) and \( y \) respectively, \( N \) is the number of points on the contour and \( 1 \leq n \leq N \). In order to remove bias when computing a matching score for two shapes, in equations (3) and (4), we set a standard reference to the length of the contour, and select its period as \( 2\pi \).

Formally, the arc length from an arbitrary starting point to a point \( p \) is given in terms of the contour length \( n \) and is converted to its angular representation by \( \phi = 2\pi n/ N \) (Zahn and Roskies, 1972). This conjecture is illustrated in Fig. 1 and Fig. 2, where the reconstructed \( x_n \) and \( y_n \) components for as few as six Fourier descriptors shows close to original reconstruction.

**Normalization of the Fourier Coefficients**

It is possible to normalize the Fourier coefficients so that the major and minor axes of the closed contour are oriented parallel with the coordinate axis of the given coordinate system: after normalization, the major axis is parallel with the \( x \), or horizontal, axis, the minor axis is then parallel with the \( y \), or vertical, axis. This property is displayed in Fig. (3). The procedure for normalizing the Fourier coefficients, according to Kuhl and Giardina (Kuhl and Giardina, 1982), is as follows. Let the normalized coefficients of the \( n^{th} \)
Figure 2. Reconstruction of the x coordinate of Fig. 1 using the first six Fourier descriptors out of 100. The first descriptor is the average, and is left out. (a) Second descriptor only (b) Sum of second and third descriptors (c) Sum of second, third and fourth descriptors (d) Sum of second, third, fourth and fifth descriptors (e) Sum of second, third, fourth, fifth and sixth descriptors. Notice that an increase in the number of descriptors results in a more accurate shape representation.

harmonic be \(a_n**, b_n**, c_n**\) and \(d_n**\). Then using symbols from eq. (3) and eq. (4)

\[
\begin{bmatrix}
a_n** \\
c_n**
\end{bmatrix}
= \frac{1}{E^*} \begin{bmatrix}
\cos(\Psi) & \sin(\Psi) \\
-\sin(\Psi) & \cos(\Psi)
\end{bmatrix} \begin{bmatrix}
a_n \\
c_n
\end{bmatrix}
\begin{bmatrix}
\cos(\Theta) & -\sin(\Theta) \\
\sin(\Theta) & \cos(\Theta)
\end{bmatrix}
\begin{bmatrix}
b_n** \\
d_n**
\end{bmatrix}
\]

where

\[E^* = \sqrt{a_1^* + b_1^*}, \Theta = \frac{1}{2} \arctan \frac{2(a_1 b_1 + c_1 d_1)}{a_1^2 + b_1^2 - c_1^2 - d_1^2}, \Psi = \arctan \frac{c_1}{a_1^*}\]

For further details, please see (Kuhl and Giardina, 1982). Isotropic normalization consists of translating and scaling the coordinates so that their centroid is at the origin \((0, 0)\), and the points on the contour are scaled so that their average distance from the origin is equal to \(\sqrt{2}\). If an isotropic normalization scheme is carried out before normalizing the Fourier coefficients, then \(a_1**\) and \(c_1**\) will be 0. Due to the properties of the Fourier transform, \(b_1**\) and \(d_1**\) are always 0. The end result is then a contour with \((0, 0)\) centroid and uniform scale and orientation. Using the orientation axes from the Fourier normalization, we remove the aspect (height/width) ratio along these axes for each candidate and target object, making the height and width uniform 1 value. This means that the major and minor axes lengths are equal to each other, with a uniform value of 1. This method is similar to scaling and rotating the basic ellipse (Folkers and Samet, 2002), shown in Fig. 4. With the aspect ratio removed, the distortions are removed from all objects in a similar manner. The original height/width ratio of the target object could also be used, but this requires different ratios for each target, and then for each candidate object, increasing processing time.
Figure 3. Normalizing Fourier coefficients after isotropic normalization of coordinates. (a) original orthogonal view of contour of lake Balaton (b) original affine transform of contour (c) original projective transform of contour (d) normalized orthogonal view (e) normalized affine transform (f) normalized projective transform - Note same orientation, scale for (d)-(f).

The Wavelet Power Spectrum

For series analysis where a predetermined scaling may not be appropriate because of a wide range of dominant frequencies, such as measuring angles, a method of time frequency localization that is scale independent, such as wavelet analysis, should be employed, instead of Fourier analysis. The wavelet transform can be used to analyze series that contain non-stationary power at many different frequencies. Wavelet transforms have advantages over traditional Fourier transforms for representing functions that have discontinuities and sharp peaks (Thesasiri et al., 2008). An example of the wavelet transform is the Morlet wavelet, Fig.(5) consisting of a plane wave modulated by a Gaussian:

$$\Psi(\eta) = \Pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2} \quad (6)$$

Figure 4. Normalizing the contours (a) original orthogonal view of fish contour with height/width ratio of 1 (b) projective transform of original fish contour (with correct height/width ratio) (c) the projective contour after isotropic scaling, Fourier coefficient normalization and height/width ratio of 1 - Note similar orientation, scale of (a) and (c).

where $\omega_0$ is the unitless frequency. The continuous wavelet transform of a discrete sequence $x_n$ is defined as the convolution of $x_n$ with a scaled and translated version of $\Psi(\eta)$:
\[ W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \Psi^* \left( \frac{(n'-n)\delta}{s} \right) \] 

(7)

where the (*) indicates the complex conjugate, and \( \Psi \) has been normalized. To ensure that the wavelet transforms at each scale are directly comparable to each other and to the transforms of other time series, the wavelet function at each scale \( s \) is normalized to have unit energy:

\[ \Psi(s\omega_k) = \frac{2\pi}{\delta^2} \Psi_0(s\omega_k) \] 

(8)

For further details, please see (Torrence and Compo, 1998). Since the wavelet function \( \Psi(\eta) \) is in general complex, the wavelet transform \( W_n(s) \) is also complex. The transform can then be divided into the real part, \( R W_n(s) \), and imaginary part, \( I W_n(s) \), or amplitude, \( |W_n(s)| \), and phase, \( \tan^{-1}(I W_n(s)/R W_n(s)) \). We can define the wavelet power spectrum as \( |W_n(s)|^2 \), similar to the Fourier Power Spectrum, with the difference being that the wavelet spectrum, due to the different wavelet scales, is more sensitive to variation. An additional advantage, the wavelet power spectrum allows smoothing of errors without losing the resolution of the original data (Karshenas et al., 1999).

\( \text{Figure 5.} \) (a) Morlet wavelet of arbitrary width and amplitude (b) Construction of the Morlet wavelet (blue dashed) as a sine curve (green) modulated by a Gaussian (red) (Torrence and Compo, 1998).

Matching

After reducing the projective differences between the target object and the candidate objects by applying a height/width ratio of one to the normalized target object and the candidate objects, the next step is to measure the interior angles of the normalized objects. Starting from any point on the target object and the candidate objects, surveyor’s method of angle calculation is utilized, which calculates angles from the azimuth of two conjoined lines, as shown in Fig. (6). The azimuth between two points can be calculated by

\[ \tan(\Theta) = \frac{dy}{dx}, \]

or the difference between the starting and ending y coordinate over the difference between the starting and ending x coordinate. The interior angle can then be calculated from the azimuths. As can be deduced from Fig. 6, an interior angle of 180° is representative of a straight line, and as thus, is not actually an interior angle, but represents an unbroken, straight continuation of the previous line segment. Interior angles that are close to 180° could also be considered being on a straight line: the deviation from 180° can be due to random digitizing errors and inaccuracies.

It is thus necessary to find a threshold which can discriminate whether two line segments form a straight line or not. One solution is to emulate primate vision. The performance of primate vision is superior to that of current computer vision systems, so there is potentially much to be gained by emulating it (Pasupathy...
and Connor, 1999). We selected 40° deviation from a straight line as the cut off, based on data from (Pasupathy and Connor, 1999). If an interior angle is between 140° and 220°, then it will be considered a straight line with an interior angle of 180°. This way, only the most prominent angles will be kept, the ones that provide the most stimulus to the primate brain. The position of an interior angle on an object is thus a function of the contour length from the random starting point. Due to projective distortion, noise and errors, the sum of the interior angles will rarely match, even for the exact same object. The properties of the wavelet power spectrum make it an ideal tool for comparing this data, since it can filter out errors, while not losing the resolution of the data. We take the wavelet power spectrum of this function, where the y axis represents the interior angle magnitude at each point, the x axis the contour length from the starting point. The Fourier Power spectrum can also be used, but the wavelet power spectrum is more sensitive to change, due to the varying wavelet scale, as mentioned previously. For similar objects, the power spectrum value in the software should have close values. Examples are shown in Fig. (7). We hypothesize that the closer this value is for two contours, the more similar they are. While this is not a comprehensive method, it does give a good indication of similarity. Two distinct objects can have the same wavelet power spectrum value, but when this score is combined with the analysis of the maximum and minimum interior angle values for an object, whether reconstructed candidate or target, it gives a very good indication of similarity without finding conjugate points, as shown in section 4. The maximum and minimum interior angle values for the target and reconstructed candidate objects should be closer to each other if they are similar objects, than if they are dissimilar objects. A simple comparison of the wavelet power spectrum, and then a comparison of the minimum/maximum interior angles constitute the matching score. The larger the difference, the more dissimilar the objects are.

**RESULTS**

We give a comparison of moment based matching presented in (Flusser and Suk, 2004). In our test dataset, some objects are manually extracted from maps and images, others are obtained from the silhouette database from the Brown University (Database, 2007). The contours used in this study are represented by a set of sampled points. Some contours were digitized with intentional errors, others were sampled at regular intervals from an output from an edge-detector. There is nothing unique about the detected edges, such that they are not intersection points or break points. The number of digitized points was also varied, then resampled to 100 for each contour, using linear interpolation, to ease use of the Fourier transform. A non-regular interval may also be used, but this would make the calculations more complicated. The recognition performance is evaluated by generating a confusion matrix in which the matching scores are recorded for each object against all the other objects in both datasets. The confusion matrix includes shades of gray to illustrate the strength of a match, such that white means a perfect match and black means no match. In Fig. 9, we included ground truth marked by red outlines, which provides correct clusters of matching objects. Ideally, the regions marked by red outlines, which correspond to clusters of objects, should have highest similarity and the other regions in the matrix should have no similarity. Representing a (thresholded and normalized) high match by white and no match by black color codes, the performance of the method provides shades of gray which shows robust matching performance. In order to test the robustness of the proposed technique to occlusions, we have synthetically generated occluded versions of the object, where the occlusion is varied based on the percentage of the area. The matching score, which is the vertical axis in the plots, is computed using the described matching technique.
Figure 7. Interior angles measured from invariant starting point of Fig.4 (a) (fish contour) and (c) (fish contour reprojected) and Fig.1 (a) (F15 with aspect ratio of 1) respectively. While the interior angles of the fish contours are not exactly the same (the reason that subtracting them from each other would be meaningless), they are more similar to each other than the F15 interior angles.

Figure 8. A collection of objects that are used to test the performance of the proposed approach against the invariant moment method.

The results are reported in Fig. 10 for 4 different shapes from the database and the occlusion is varied from 0% to 100%. While the results show an exponential change in performance, we should note that the matching score is still about .7 for a 80% overlap for all examples, although after that, the performance is quite poor. During our experiments, we have noted clustering objects of the same type together. An instance of this clustering is shown for the F15 and F16 shapes which are both planes. To quantify the results, the table below is used (the red lines group together similar objects in the confusion matrices):
<table>
<thead>
<tr>
<th></th>
<th>Power Spectrum</th>
<th>Invariant Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctly identify as match %</td>
<td>82</td>
<td>78</td>
</tr>
<tr>
<td>False positive match %</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>Similar object find %</td>
<td>69</td>
<td>65</td>
</tr>
<tr>
<td>Similar object false positive %</td>
<td>31</td>
<td>35</td>
</tr>
<tr>
<td>False negative match %</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 9.** The matching scores for objects in the form of a confusion matrix where each object is matched against all the other objects. The light areas represent higher match scores and the red outlines represent the ground truth of the clusters in the datasets. (a) Proposed method (b) Hu invariant moment method.

**Figure 10.** Matching score as a function of overlap for (a) Balaton (b) Mexico (c) F16 (d) Iceland shapes. Note that as the overlap increases, the matching score increases.

The wavelet power spectrum, using the aspect ratio of 1 for height/width had problems finding similar objects, perhaps it is not sensitive in current form, since it was hard to distinguish between Iceland, Pentagon and fiducials.
The results are comparable to that of the invariant moments. The above is condensed to the following table (this includes not just the highest matching scores):

<table>
<thead>
<tr>
<th></th>
<th>Power Spectrum</th>
<th>Invariant Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctly identify to projective %</td>
<td>78</td>
<td>75</td>
</tr>
<tr>
<td>Correctly identify to affine %</td>
<td>88</td>
<td>73</td>
</tr>
<tr>
<td>Correctly identify to occlusion/change%</td>
<td>70</td>
<td>71</td>
</tr>
</tbody>
</table>

No false negatives were produced, that is, two objects never had low match scores, if they were similar, using any method. Both methods have advantages and have similar scores, but the power spectrum is slightly more sensitive to shapes than the other, in fact, it may not be sensitive, since it cannot correctly distinguish similarity, although up to affine matching, it provides robust results.

**CONCLUSION**

This paper provides a novel approach to matching objects represented in the form of a contour. Compared to many other recognition methods in the literature, our method allows the extracted outlines to contain noise and occlusions. It provides a way to compare shapes quickly and effectively without calculating the transformation matrix. The proposed approach exploits the interior angles of the objects, which results in a robust and computationally simple procedure for matching shapes. An important contribution in this regard is the elimination of the point correspondences, having the same staring points on the silhouette outline and the direction of digitization of the outline. Experimental results show the robustness of the proposed method.

**REFERENCES**


