A new method for representing and matching shapes of natural objects

Thomas Bernier, Jacques-André Landry*

Department of Automated Production Engineering, École de Technologie Supérieure, 1100 Notre-Dame Ouest, Montréal, Qué., Canada H3C 1K3

Received 21 June 2002; received in revised form 19 November 2002; accepted 19 November 2002

Abstract

A new method for the representation and comparison of irregular two-dimensional shapes is presented. This method uses a polar transformation of the contour points about the geometric centre of the object. The distinctive vertices of the shape are extracted and used as comparative parameters to minimize the difference of contour distance from the centre. Experiments are performed, more than 39,000 comparisons of database shapes, provided by Sebastian et al. (ICCV (2001) 755), are made and the results are compared to those obtained therein. In addition, 450 comparisons of leaf shape are made and leaves of very similar shape are accurately distinguished. The method is shown to be invariant to translation, rotation and scaling and highly accurate in shape distinction. The method shows more tolerance to scale variation than that of Sebastian et al. (ICCV (2001) 755) and is less computationally intense.

© 2003 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Shape representation; Shape matching; Object recognition; Object classification; Polar transformation; Contour extraction

1. Introduction

There now exist a myriad of industrial applications of machine vision. Many applications fall into the broad category of inspection and quality control. In general, inspection processes require fast and robust methods to detect flaws or aberrations of product. In cases where the product is manufactured, aberrations are often easily detectable because of a clearly defined normal object pattern. In the case of less regular objects and particularly with biological objects, the acceptable model is more loosely defined which can make the detection processes considerably less robust.

Objects are very often distinguishable on the basis of their visible features, among these features, an object’s shape is frequently an important key to its recognition. The representation of shape is thoroughly discussed in Refs. [1,2] and in both, sets of criteria for the evaluation of shape are proposed. Effectively representing shape, however, still remains one of the biggest hurdles to overcome in the field of automated recognition. By effectively, it is meant: accurately enough such that the representation pertains uniquely to an object type, while still remaining loose enough to be tolerant of minor variations within various examples of said type.

Thus, a means to consistently measure shape is required to quantitatively compare them to one another. Bribiesca and Guzman [3] proposed both a numerical description of shape (shape number) and a quantitative measure of the similarity of shapes. Like most shape representation schemes, invariance to translation, rotation and scale is an important requirement of their description. This stems from the fact that by definition, repeated transformations by translation, rotation and scale do not affect the shape of an object [4]. In Flusser and Suk [5], a representation is described which is invariant to affine transformation as well as translation,
rotation and scale, thus allowing independence of view point. The discussion in this paper will be limited to representations of fixed-viewpoint closed planar curves, which are invariant to translation, rotation and scale.

In Refs. [6,7] broad reviews of varying approaches and methods of shape analysis are made. Pavlidis [6] proposes three classifications of shape analysis techniques. The first classification is based on the use of either the boundary or interior pixels for the representation. The second is whether or not the result is numeric and the third is whether or not the representation preserves the information such that reconstruction of the shape is possible from its representation. The representation method presented here is boundary based, produces a numeric result and preserves the shape information; however, the same problem is frequently approached by differently classed techniques.

With regards to similarly classed techniques, in both Refs. [8,9], closed planar curves were represented using a Fourier expansion of the function of their tangent angle and their arc length. The lower-order Fourier coefficients (descriptors) were then used to represent the shape. In a different approach to a similar problem, Lie and Chen [10] determined points of high curvature (feature points) of a closed planar curve and represented them in polar form. The method was shown to be theoretically invariant to translation and scale and comparisons were made on the basis of the variance of the ratio of two respective polar representations. Because the ratio was compared, scale normalization was considered unnecessary. Similarly, Chang et al. [11] described shapes by computing the distance of points of high curvature, using an algorithm proposed by Teh and Chin [12], to the centroid. The radial distances, however, were all normalized to the minimum to provide for scale invariance. Two contours were then assessed for shape similarity by a cyclic comparison of feature point radial distances. Ozugur et al. [13] again used the polar form but computed the distance of feature points to the centre of the smallest circle enclosing the shape. The polar representation was then broken into segments and each segment described by its Fourier coefficients and shape similarity is assessed using the Euclidean distance of the Fourier coefficients.

Shape recognition is also frequently performed by methods based on the interior points of objects. Moment-based analysis of shape was performed early on by Hu [14] and more recently by Zhu et al. [15] and provides a shape-preserving, numerical description, invariant to translation, rotation and scale. The medial axis transform, introduced by Blum [16], based on interior points, however, provides a graphical representation of the shape. The shock graph is a variant of the medial axis transformation and is effectively used to match shape by Sebastian et al. [17]; however it is not scale invariant.

In this paper, a new representation of shape is proposed for planar closed curves (two-dimensional object contours) which is invariant to translation, rotation and scale. In an effort to reduce the abstraction of the representation, the shape is represented directly by its contour in terms of the angle from the centroid and normalized radial length. This is in contrast to Zahn and Roskies [8], Bennet and McDonald [9] and Ozugur et al. [13] where curvature information was transformed to the frequency domain. This method also differs from that of Lie and Chen [10] and Chang et al. [11] first in that the points of high curvature (vertices) are found by performing a first-order derivative of their polar representation, and next, in that vertices are used only as a rotational reference and the whole contour is compared for similarity, thus providing sensitivity to contour segments lying between vertices without inordinate sensitivity to noise. In addition, while Lie and Chen [10] and Chang et al. [11] require somewhat rigid feature point similarity for comparison, this method can iteratively adjust until an equal number of vertices are found, thus significantly broadening its comparative ability. Finally, while the above methods resulted in either similar or not similar (binary) assessment, this method results in a normalized scalar assessment of shape similarity. Furthermore, by retaining all of the contour information, this method allows for analysis of dissimilarities both in terms of location and proportion, thus enabling future work on tolerance to occlusion.

As mentioned, shape is only one among many visible attributes used in the recognition of objects. The described strategy of representation and matching of shape will ultimately function as one among many components in a larger multifaceted approach to an adaptable object recognition system. In this paper a new method of boundary, based shape representation and matching is described and results for shape matching are presented and compared to those obtained by the method of Sebastian et al. [17].

2. General description of the approach

In this paper a method is described which consists of two integrally linked aspects. The first is representation and the second is matching, through which two representations can be compared. While segmentation is required prior to representation, it is not the topic of the paper and is discussed only briefly in Section 2.1. Section 2.2 describes the method of representation in detail and Section 2.3 describes the comparison or matching of representations.

2.1. Segmentation and contour extraction

The shape representation method described here assumes that the object has been fully segmented from the image such that all pixels representing the object’s shape have been identified as distinct from those pertaining to the rest of the image. In this paper, a local diffusive segmentation method [18] was used. There exist a wide variety of ways to achieve segmentation; however, it is not the subject of this paper. All contiguous pixels, which share a given point-based characteristic of the object or are surrounded by those that do, are
considered as object pixels and those outside the included region are considered as background. The result is a group of contiguous pixels, which collectively represent the object.

The boundary pixels of the object are then extracted from the segmented object pixels by a simple iterative trace, around the outside of the object, that continues until the starting point is reached. This results in a second group of pixels collectively representing the object’s exterior contour [18]. It should be noted here that, because only the exterior contour points are considered, the possibility of having holes in the object is precluded. In order to study the inside contour of the object, the contour of the hole would be extracted and handled separately as an independent shape.

2.2. Shape representation

In order to permit representational invariance to position, rotation and scale, the geometric centre of the shape is selected as a reference point. The coordinate location of the centroid is calculated using the following equation:

\[
(x_c, y_c)_{\text{centre}} = \left( \frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n} \right),
\]

where \(n\) is the number of points in object.

From the geometric centre point, the distance and angle of each contour point are calculated using Eqs. (2) and (3), respectively, as shown in Fig. 1:

\[
\text{Distance}(d) = \sqrt{(x - x_c)^2 + (y - y_c)^2},
\]

where \(x_c, y_c\) are the coordinates of centroid, and \(x, y\) the coordinates of boundary point.

\[
\text{Angle}(\theta) = \begin{cases} 
\pi/2 & \text{if } \Delta x = 0 \text{ and } \Delta y \geq 0, \\
3\pi/2 & \text{if } \Delta x = 0 \text{ and } \Delta y < 0, \\
\arctan(\frac{\Delta y}{\Delta x}) + \pi & \text{if } \Delta x < 0, \\
\arctan(\frac{\Delta y}{\Delta x}) + 2\pi & \text{if } \Delta x > 0 \text{ and } \Delta y < 0, \\
\arctan(\frac{\Delta y}{\Delta x}) & \text{if } \Delta x > 0 \text{ and } \Delta y \geq 0,
\end{cases}
\]

Fig. 1. Distance and angle of the contour points relative to centre.

where \(\Delta y = y - y_c, \Delta x = x - x_c, x_c, y_c\) are the coordinates of centroid, and \(x, y\) the coordinates of boundary point.

The result of the calculations is a representation of the contour as described by polar coordinates. This coordinate substitution is performed only to simplify the handling of scale and rotational variations, which can each be accomplished through the scaling of a single variable. In order to provide for scale invariance, the maximum distance is computed and all distances are normalized to it. Thus all values fall between 0 and 1 regardless of the scale of the image or object. Similarly, in order to provide for rotational invariance, the phase of the polar representation is shifted by the angle associated with the maximum distance, such that the maximum is associated with an angle of zero (Fig. 2).

In Bernier and Landry [19], as with many shape recognition processes, the final shape-based distinction of objects was made through the application of a template. This template was both size and orientation specific and thus had to be repeatedly applied at every possible orientation to check for a match. It was also necessary that a template,
specific to the objects being recognized and their scale, be created manually prior to the recognition process. The repeated applications of the two-dimensional template were cumbersome and time consuming. In addition, with spores of similar shape but of different size, the template would fail to find a match. By representing the contour in this manner, effectively a similar template is created but is described through dynamic and shape-specific correlation. Thus, objects of varying scale and orientation will ideally represent similarly although by a different number of contour points (Fig. 3).

2.3. Shape matching

In an ideal case the similarity of two shapes of varying scale and orientation as shown in Fig. 4 is inversely proportional to the area lying between their respective polar representations. The relation of similarity to the area lying between two polar representations is

\[
\text{Similarity of shape} \sim \left( \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \sqrt{(R_0 - R_0')^2 + \theta} \right)^{-1},
\]

where \(\theta\) is the angle of deflection from maximum radius \((0 - 2\pi)\) and \(R_0\) the fraction of max radius at angle \(\theta\).

However, the ideal case assumes that the maximum radius will always occur at the same point while nature holds many exceptions to this assumption. The two leaves shown in Fig. 5 are an example of such exceptions.

In order to deal with the likely event of objects having different maximum radii, a new method was developed to determine the orientation at which the similarity is optimized.
Optimization of similarity is simply finding the relative orientation in which the area lying between the respective polar representations of the shapes is minimized. Intuitively, one would determine that orientation by simply calculating the value at all possible orientations and finding the minimum. However, the outline of an object can be composed of several thousand points, furthermore scale invariance dictates that two outlines may be composed of a very different number of points, thus determination of optimal orientation cannot be achieved via simple minimization.

A set of scale independent, object-specific reference points is required. By performing a smoothed first-order derivative (Eq. (5)) on the polar representation of the data, the local minima and maxima can be determined by finding the zero crossings of the resulting derivation (Fig. 6):

$$\frac{d}{d\theta} R(\theta_p) = \frac{\sum_{i=1}^{n} R(\theta_{p-i}) - \sum_{i=1}^{n} R(\theta_{p+i})}{\sum_{i=1}^{n}(\theta_{p-i}) - \sum_{i=1}^{n}(\theta_{p+i})},$$  \hspace{1cm} (5)$$

where \( \theta_{p+i} \) is the \( i \)th value of \( \theta \) after \( \theta_p(0 - 2\pi) \), \( R(\theta_p) \) the fraction of max radius at angle \( \theta_p \), and \( n \) the scope of smoothing.

The local minima and maxima obviously correspond to the morphological vertices and depressions, which are often distinctive characteristics of shapes. The scope of the smoothing \( n \) determines the \textit{localness} of the maxima and minima. In general, a scope of \( n \approx \pi/10 \) yields only the major vertices but this can be adjusted until the appropriate vertices are found. The benefit of finding the vertices of the object is in reducing the very high and varying number of boundary points to a very low and distinctive number of vertices.

The number of vertices used to represent the model object determines the number required and the scope of smoothing is iteratively adjusted until an equal number of vertices are found for the unknown object. The requirement of aligning the orientation of the two objects with a similar number of vertices is then reduced to finding the angle of rotation that leads to a minimization of the Euclidean distance between corresponding vertices:

$$\text{Optimal angle } (\phi) \sim \min_{\phi} \left( \sum_{p=0}^{P-1} \sqrt{(R_p \cos \theta_p - R'_{p+s} \cos(\theta'_p + \phi))^2 + (R_p \sin \theta_p - R'_{p+s} \sin(\theta'_p + \phi))^2} \right),$$ \hspace{1cm} (6)$$

where \( \phi \) is the change in angle of orientation of object, \( \theta_{p+s} \) the angle of deflection of vertex \( (p+s) \) modulo \( P \), \( R_p \) the normalized radius of vertex \( p \), \( P \) the number of vertices, and \( s \) the offset of optimal correspondence of vertices.
However, this equation assumes that the optimal correspondence of vertices is known. The offset value \( s \) must first be determined using a similar method shown in the following equation:

\[
\text{Optimal offset} \ (s) = \min_{s=0}^{P-1} \left[ \sum_{p=0}^{P-1} \sqrt{(R_p \cos \theta_p - R'_{p+s} \cos \theta'_{p+s})^2 + (R_p \sin \theta_p - R'_{p+s} \sin \theta'_{p+s})^2} \right], \tag{7}
\]

where \( \theta_p \) is the angle of deflection of vertex \( p \), \( \theta_{p+s} \) the angle of deflection of vertex \( p + s \) (modulo \( P \)), \( R_p \) the normalized radius of vertex \( p \), \( P \) the number of vertices, \( s \) the offset of corresponding vertices.

Clearly, the matching of the peaks alone does not give a very accurate assessment of shape similarity since most of what defines a shape lies between the peaks. Thus, once the optimal alignment of vertices is achieved, the similarity of the whole boundary is assessed and the orientation is again adjusted to minimize the difference. As mentioned, due to scale variations, the line segment between any pair of corresponding vertices can be composed of a significantly different number of points from one object to another. Consequently, the area between the two line segments is assessed at equal intervals along the arc lengths rather than at every point and calculated using the equation

\[
\text{Similarity of shape} = 1 - \left( \frac{1}{SP} \sum_{p=0}^{P-1} \sum_{i=0}^{I-1} \sqrt{(R_i \cos \theta_i - R'_{i+s} \cos \theta'_{i+s})^2 + (R_i \sin \theta_i - R'_{i+s} \sin \theta'_{i+s})^2} \right), \tag{8}
\]

where \( \theta_i \) is the angle of deflection at the \( i \)th interval after vertex \( p \), \( R_i \) the fraction of max radius at angle \( \theta_i \), \( \theta'_{i+s} \) the shifted angle of deflection at the \( i \)th interval after corresponding vertex \( p \), \( R'_{i+s} \) the fraction of max radius at angle \( \theta'_{i+s} \), \( I \) the number of intervals between vertices, and \( P \) the number of vertices.

The above equation assumes that the maximum possible distance between two corresponding edge points is equal to the maximum radius (normalized to 1). Thus a similarity of zero is effectively impossible and the result will never be negative. Any arbitrary two-dimensional shape can then be compared to any other by the outlined method, with the exception of the special case of the circle.

A circle by definition has no vertices and thus will tend to give unpredictable values. Therefore, prior to performing the derivation on the polar data, shapes are checked for circularity. Circularly can be conveniently assessed directly from the polar data. A perfect circle has constant radial value, thus any deviation in radial length would indicate non-circularity. A perfect circle cannot be represented by a digital matrix, thus it is harder to detect. However, if the average deviation from the maximum radius is less than 5% or the number of vertices is less than two, the user is notified and given the option to assume that the object is circular.

When a circular object is being compared, the derivation is simply not performed and no vertices are found. A circle’s can still be compared to other objects by the above method but the orientation is assumed irrelevant. Thus for any two objects—circular or otherwise and including those with multiple boundary points for a given angle (i.e. spirals)—the similarity in shape can be assessed and quantified by a scalar value (Fig. 7).
3. Experimental results

In an effort to demonstrate the performance of the method presented here, two sets of experiments were performed. First, in order to provide a point of reference, the method was performed on the test shapes used in Sebastian et al. [17] and the results were compared to those obtained by the authors. Secondly, the method was performed on natural images of leaves to demonstrate the accuracy and ability to distinguish between closely similar shapes. The work of Sebastian et al. [17] was chosen as a reference because both their goals and the resulting quantization of shape similarity were very similar to those of the method presented here and because their method provides a highly accurate distinction between shapes.

3.1. Experiment 1—shape database

Fig. 7 shows the set of 99 shapes graciously made available by Sebastian et al. [17]. The set includes nine categories with 11 shapes in each category. Each of the shapes was used as a model to which all the other shapes were compared, thus 9801 shape comparisons were made for each repetition of the experiment.

Ideal results would be that the 11 closest matches (including the model itself) all be of the same category as the model as well as a sharp decrease in similarity for the objects outside of the category. Some of the results obtained by Sebastian et al. [17] are shown in Fig. 9. Their % recognition rates for the shapes in Fig. 8 were (100, 100, 100, 99, 99, 97, 96, 95, 87), respectively, describing the rate at which the nth nearest match was in the same category as the model.

In Fig. 10, the 16 nearest matches, obtained by the method presented here, and their match values are depicted for some of the shapes. Note that the match value depicted in Fig. 10 is simply the 1’s complement of their similarity to the model (as calculated by Eq. (8)), multiplied by 1000 to facilitate comparison to the values obtained by Sebastian et al. [17].

It is clear that the categorization is not as thorough as that of Sebastian et al. [17]; however, other issues are of significance. Foremost, the method presented here was never intended nor designed to be tolerant to occlusion. This fact can be seen in the trials with the rabbits and the human forms, only complete shapes are near matches, i.e. the human form missing a leg was not a near match, in fact it was the 44th match for the human model shown in Fig. 10. The same is true for the rabbit shapes without ears and those with occluding masks. Furthermore, the nearest matches for the wrenches were always of a similar type wrench (i.e. double ended or single ended). Lastly, in most cases, the largest difference in consecutive values lies between the last shape of the category and the first shape outside the category, thus allowing a broad margin for thresholding. Table 1 shows the number of times for each category that the nth nearest matches are in the appropriate category.

While the results are lower than those of Sebastian et al. [17], in the categories of human figures, rabbits and hands, shapes with missing or occluded parts were simply not considered as similar. This clearly both alters the results considerably and demonstrates a high potential for fine discrimination, which would have many applications in inspection processes. If the cases involving a high degree of occlusion are excluded (i.e. cases in which a major portion of the shape is missing or a mask is added to the shape), the results are considerably higher and are shown in Table 2.

More significantly, in order to demonstrate scale and rotational invariance, the experiment was repeated with the image of Fig. 8 increased in scale to 400% (Table 3), scaled down to 50% (Table 4) and finally with the image both
scaled down to 50% and rotated by 90° (Table 5). The original models (of original scale) were used to search for similar objects in an identical manner.

With very few exceptions, the results of the matching are identical regardless of their scale or rotation. The few exceptions can confidently be attributed to the pixelation error.
Table 1
Resulting nearest matches by category

<table>
<thead>
<tr>
<th>Category</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Rabbits</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Planes</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Greebles</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Wrenches</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Hands</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Humans</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Quadrupeds</td>
<td>11</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>%</td>
<td>100.0</td>
<td>98.0</td>
<td>93.9</td>
<td>92.9</td>
<td>85.9</td>
<td>76.8</td>
<td>70.7</td>
<td>65.7</td>
<td>53.5</td>
<td>37.4</td>
<td>25.3</td>
</tr>
</tbody>
</table>

Table 2
Nearest matches by category with occluded shapes omitted

<table>
<thead>
<tr>
<th>Category</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>91</td>
<td>91</td>
<td>55</td>
<td>36</td>
<td>18</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Rabbits</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>84</td>
<td>75</td>
<td>38</td>
<td>86</td>
<td>90</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Planes</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>91</td>
<td>91</td>
<td>55</td>
<td>36</td>
<td>18</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Greebles</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>84</td>
<td>75</td>
<td>38</td>
<td>86</td>
<td>90</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Wrenches</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>91</td>
<td>91</td>
<td>55</td>
<td>36</td>
<td>18</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Hands</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>91</td>
<td>91</td>
<td>55</td>
<td>36</td>
<td>18</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Humans</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>91</td>
<td>91</td>
<td>55</td>
<td>36</td>
<td>18</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Quadrupeds</td>
<td>100</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>64</td>
<td>45</td>
<td>9</td>
<td>27</td>
<td>45</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>Rays</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>%</td>
<td>100.0</td>
<td>97.0</td>
<td>96.0</td>
<td>90.9</td>
<td>81.8</td>
<td>77.8</td>
<td>72.3</td>
<td>57.3</td>
<td>38.8</td>
<td>26.4</td>
<td></td>
</tr>
</tbody>
</table>
Table 3
Resulting nearest matches by category for 400% scale image

<table>
<thead>
<tr>
<th>Category</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Rabbits</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Planes</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Greebles</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Wrenches</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Hands</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Humans</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>%</td>
<td>100.0</td>
<td>98.0</td>
<td>94.9</td>
<td>92.9</td>
<td>85.9</td>
<td>74.7</td>
<td>73.7</td>
<td>65.7</td>
<td>54.5</td>
<td>43.4</td>
<td>36.4</td>
</tr>
</tbody>
</table>

Table 4
Resulting nearest matches by category for 50% scale image

<table>
<thead>
<tr>
<th>Category</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Rabbits</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Planes</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Greebles</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Wrenches</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Hands</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Humans</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Quadrupeds</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>%</td>
<td>100.0</td>
<td>96.0</td>
<td>93.9</td>
<td>88.9</td>
<td>76.8</td>
<td>73.7</td>
<td>62.6</td>
<td>63.6</td>
<td>51.5</td>
<td>31.3</td>
<td>23.2</td>
</tr>
</tbody>
</table>

Table 5
Resulting nearest matches by category for 50% scale and 90° rotated image

<table>
<thead>
<tr>
<th>Category</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Rabbits</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Planes</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Greebles</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Wrenches</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Hands</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Humans</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Quadrupeds</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>%</td>
<td>100.0</td>
<td>96.0</td>
<td>93.9</td>
<td>88.9</td>
<td>76.8</td>
<td>73.7</td>
<td>62.6</td>
<td>64.6</td>
<td>51.5</td>
<td>32.3</td>
<td>22.2</td>
</tr>
</tbody>
</table>

introduced by the scaling of the image (reducing to 50% altered the proportion of the finer parts of the shapes). This is in contrast to the results obtained by Sebastian et al. [17] which were tolerant to only modest amounts (< 3 times) of scaling. The 16 nearest matches for two selected objects and their comparative match values for each repetition of
the experiment are shown in Fig. 11. Note that in all cases, the shape corresponding to the model was selected as the first match and for all trials there was little or no difference in the matchings of the 50% scale image and the 50% scale and rotated image.

In the original image, each shape was represented by about $125 \times 125$ pixels. The comparison of one model to each of the 99 shapes required about 13.6 0.125 s per comparison. These results were obtained on a 1.4 GHz AMD machine and were run under Windows 2000. Sebastian et al. [17] reported their algorithm, to compare one shape to another, taking about 3–5 min on an SGI Indigo II (195 MHz). The speed, however, has since been increased by 50–100 times through the implementation of a coarse-scale measure. The speed of performance for the method shown here will also require improvement for real-time applications but the code is far from being optimized and the run time can still be significantly reduced.

3.2. Experiment 2—natural shapes

Fig. 12(a) shows a set of 30 leaves of various species. Amongst the leaves 11 are sumac and 4 mint, which are quite similar in shape and identified in Fig. 12(b). Each of these leaves is represented by approximately $50 \times 100$ pixels. Experiment 1 was repeated with the leaves such that each of the sumac and mint leaves was selected as models and compared to all of the other leaves in the image.

The 16 nearest matches for selected leaves are shown in Fig. 13. In all, the % rate of correct categorization for all of the 11 sumac leaves is $(100, 100, 100, 100, 100, 100, 100, 100, 100, 100)$ in order of the $n$th nearest match and $(100, 100, 75)$ for the 4 sumac leaves.

The results again demonstrate both the method’s high sensitivity to minute differences in shape and invariance to position, scale and orientation of the object. The comparison of one leaf to all of the others required 3.75 or 0.125 s per comparison on a 1.4 GHz AMD machine.

4. Conclusion

The main objective of this paper was to introduce and demonstrate a novel and robust method of representing and comparing shapes of irregular objects. More than 39 000 comparisons were made among the shapes provided by Sebastian et al. [17]. Another 450 comparisons of leaf shapes were made and the presented method was shown to be accurate and invariant to translation, scale and rotation. For the shapes provided by Sebastian et al. [17], the results showed high accuracy ($> 85\%$) for the first four matches regardless of their scale or orientation. By omitting those that were
occluded, the accuracy was increased to over 90% for the first four matches. For the leaf shapes, the categorization was perfect (100%) in all cases except one misclassification of a sumac leaf. The speed of the process is proportional to the number of pixels representing the shape’s contour; however, for the shapes used here the time required was on average 0.137 and 0.125 s per comparison for the shapes provided by Sebastian et al. [17] and the leaf shapes, respectively. Given that the process was not designed for a specific application, it is difficult to assess whether or not it is performed at an adequate speed; however, the speed could be significantly increased through code optimization. As an example, by running a preliminary filter using moment analysis, thus reducing the number of invocations of the main algorithm, the required time was reduced to an average of 0.048 s per comparison (∼20 comparisons per second) with identical accuracy for the shapes provided by Sebastian et al. [17].

In general, the experiments demonstrate very encouraging results for the recognition of and distinction between various types of objects. In comparison to the method of Sebastian et al. [17], the method showed more invariance to scale and is less computationally demanding; however, less tolerant to occlusion and marginally less accurate at this stage of development. This is due both to differing objectives and to the new method never having been designed to tolerate occlusion. However, due to the nature of the method, partial

---

Fig. 12. Leaf image (a) and identification key (b).

Fig. 13. Selected results of leaf matching.
matches of contour segments are easily determined and occlusion could be handled with only minor refinements. Thus, while less accurate in terms of occlusion, the method presented here compared favourably in terms of speed, spatial invariance and fine distinction of non-occluded shapes.

5. Summary

In this paper, a new method for the representation and comparison of the shape of natural objects is presented. This method is one among many currently under development by the authors. Ultimately, these methods will act together in an adaptable automated recognition system for the detection and distinction of natural objects.

While many approaches have been made towards the description of shape, few have been successful in the recognition of non-specific natural objects. The presented process describes the shape of an object using a normalized, phase-shifted list of the boundary polar coordinates. The description is invariant to translation, rotation and scale.

The paper briefly outlines the method of segmentation of the object and contour extraction with which this process begins. The methods of representation and comparison are then described in detail. Experiments are performed and more than 39,000 comparisons of database shapes (provided by Sebastian et al. [1]), are made and the results are compared to those obtained therein. In addition, 450 comparisons of leaf shapes are made and leaves of very similar shape are accurately distinguished. The results of the experiments demonstrate that the proposed method is very effective at distinguishing between very similar shapes.

References


About the Author—THOMAS BERNIER holds a doctorate degree in Biosystems Engineering from McGill University and works primarily in the development of an automated system of adaptable recognition of biological objects. He is currently employed as a Research Associate at LIVIA (Laboratory for Imagery, Vision and Artificial Intelligence) of L’École de Technologie Supérieure in Montréal, Quebec and is continuing his research in adaptive computer vision.

About the Author—JACQUES-ANDRÉ LANDRY is a University Scholar from McGill University and holds a Doctorate degree in Biosystems Engineering, with a specialization in process control applied to agro-biosystems. He currently holds an academic position in the Department of Automated Production at l’École de technologie supérieure in Montréal, following 14 years of research and teaching at McGill University. His main areas of research are process control for agro-biosystems, machine vision for irregular/biological object recognition, and remote sensing for precision agriculture.