



PERGAMON

Pattern Recognition 35 (2002) 1083–1097

**PATTERN
RECOGNITION**

THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Invariant characterisation of the Hough transform for pose estimation of arbitrary shapes

Alberto S. Aguado^{a,*}, Eugenia Montiel^b, Mark S. Nixon^c

^a*Department of Electronic and Electrical Engineering, University of Surrey, Guildford, Surrey GU2 7XH, UK*

^b*iMAGIS/GRAVIR-IMAG, INRIA Rhône-Alpes ZIRST - 655 avenue de l'Europe, France*

^c*Electronics and Computer Science, University of Southampton, UK*

Received 13 December 2000; received in revised form 1 May 2001; accepted 1 May 2001

Abstract

In this paper, we develop a new formulation and methodology for including invariance in a general form of the Hough transform. Essentially, the transformations that control a shape's appearance are extracted using invariance, for arbitrary shapes with a continuous description. We first develop a formal definition of the Hough transform mapping for arbitrary shapes and general transformations. We then include an invariant characterisation of shapes and develop and apply our new technique to extract shapes under similarity and affine transformations. Our formulation and implementation is based directly on parametric curves and so avoids the use of indexed look-up tables. This confers the attributes of a continuous shape description avoiding discretisation problems inherent in earlier formulations. To obtain an invariant characterisation, each point in the model is related to a collection of other points defining a geometric arrangement. This characterisation does not require the computation of properties for lines or other primitives that compose the model, but is based solely on the local geometry of the points on shapes. The transformation is obtained by solving for the parameters of the curve according to an arrangement of points defined for a point in the image and a corresponding arrangement of points for a point in the model with the same invariant properties. The location parameters can be gathered in a 2D accumulator space independent of the transformation and of a shape's complexity. Experimental results show that the new technique is capable of extracting arbitrary shapes under occlusion and when the image contains significant noise. © 2002 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Shape extraction; Hough transform; Object recognition; Invariance; Similarity transformations; Affine transformations

1. Introduction

The characterisation and extraction of objects based on shape information has been an extensive field of research in image analysis. Shapes offer an important feature for object characterisation since they contain valuable information independent of photometric phenomena such as

changes in illumination, contrast and colour. However, the appearance of shapes changes with viewpoint. Accordingly, an important topic of research in object extraction has focused on characterising and extracting shapes under different image transformations. These techniques generally use a model shape obtained from a reference image. Thus, shape extraction is achieved by determining the transformation which maps the model into the primitive in the image. This paradigm is commonly referred to as model-based shape extraction [1–3].

* Contact author. Tel.: +44-1483-876044; fax: +44-1483-534139.

E-mail address: a.aguado@eim.surrey.ac.uk (A.S. Aguado).

Many methods of shape extraction have used the concept of invariance to overcome the problem of characterising shapes under different transformations. Invariance is a fundamental concept in geometry [4,5]. Currently, the study of invariance is an active field of research in computer vision. The general approach to shape extraction via invariance features has been based on three main steps [6,7]. Firstly, some significant matching structures of the image and the model are identified. Secondly, the image transformation which maps the model structures into the image is computed. Finally, the consistency of the transformation is evaluated to determine if the structure in the image corresponds to the structures in the model. In the earliest methods, these steps were based on a verification mechanism [8] wherein new transformations were obtained by trying different structures in the image. Transformations were evaluated until a strong consistency between the image and model features was determined. These original ideas have motivated the development of other techniques cited under different names such as cluster methods, pose clustering, evidence gathering, geometric hashing and hypothesis accumulation [6,7,9–12]. These techniques develop a different way of evaluating the consistency of the transformation. Instead of performing a verification process for each transformation of a template, the set of possible transformations is considered as a “cluster space” according to a suitable parameterisation. Each point in the image space then provides evidence that can be gathered in the cluster space. When all image points have been considered, each element in the cluster space contains a measure of the consistency of the transformation. In this way the local search given by the verification procedure is replaced by a global clustering approach. Although some techniques have considered efficient search techniques to perform the verification process [11,13], it is generally recognised that evidence gathering is computationally more attractive, in spite of requiring significantly more memory for implementation.

By definition, cluster methods and verification methods of matching are robust. In these methods, we can consider that all the values of potential transformations between a model and data in an image define a probability distribution, thus the best transformation is estimated by the mode, that is, by the maximum value. It is well known that the mode is a robust estimator [14] since it is not biased due to outliers. In shape extraction, outliers correspond to transformations obtained by mapping model structures to noise or to other data that do not define the primitive in the image. In consequence, clustering techniques should be able to handle incorrect evidence. However, an error analysis of clustering methods concludes that they can fail to provide the correct solution [15]. According to error analysis, the problem of clustering approaches is that they are based on primitives that are generally unreliable. For example, if we use

lines to estimate the scale of a shape, then the length of the lines must be precisely known. Errors provide a factor of uncertainty which generates considerable incorrect evidence [15]. Thus, the computation of properties useful for matching requires efficient and accurate methods for detecting primitive structures. Unfortunately, noise, occlusion, and changes in position and illumination make it difficult to obtain an accurate estimate of measured properties. Additionally, some primitives are adequate for characterising only one type of object such as polyhedra and might fail in a general context. In general, the inclusion of non-geometric parameters in the characterisation of a shape can reduce excessive accumulation of votes [12]. However, properties such as colour or brightness can change due to illumination conditions and provide a less generic characterisation of an object than the exclusive use of the borders of a shape. Here, we are particularly interested in the geometric invariance characterisation of shapes.

In this paper, we consider the problem of shape extraction by matching only points in a shape. That is, instead of searching for lines or other image primitives, we study how invariant evidence can be gathered when points in a shape are matched to points in the model. The advantage is that uncertainty (by occlusion or imperfect detection) is lower for points than for higher level primitives. We use a gathering evidence mechanism defined by the Hough Transform (HT) [16]. In the HT, a straight line in an image has a dual representation as a point in the parameter space. A discrete version of this space is commonly referred to as the Hough space or accumulator space and it is used to gather evidence by exploiting the dual representation of a point in the image. In the HT, this point is obtained by incrementing the elements of the accumulator space that define the trace of the dual line of each edge point in the image. A concurrent point in the parameter space, defined by collinear points in the image, is incremented several times forming a local maximum in the accumulator space. These ideas have been extended to extract quadratic forms and arbitrary shapes [17–19].

In this paper, we avoid the representation of models by indexed look-up tables and we directly use parametric curves to represent shapes under a transformation. In our approach, the problem of shape extraction is characterised, not by the properties of curves, but by the transformation. For example, the problem of shape extraction of an ellipse is studied as the problem of extracting a circular shape under an affine transformation. Thus, instead of performing shape extraction by considering particular geometric properties of the ellipse, we use a more general approach based on invariant properties that are generic for any shape under an affine mapping. By considering arbitrary shapes and arbitrary transformations, it is possible to develop a general formulation and methodology for including invariance in a general form of the HT. In

this way, we formulate a formal definition of invariant characterisation within the HT, and we then develop and implement particular techniques for similarity and affine transformations.

In the HT, a point in an image and a point in a curve define a point spread function (psf) that represents the trace of a curve in an accumulation space whose size is exponentially proportional to the number of free parameters of the shape. By using an invariant characterisation, a psf can be reduced to a collection of points in a 2D space whose arguments represent shape's location. The size of this space is independent of the transformation or complexity of the shape, thus, invariance provides a general approach for reducing the computational requirements of the gathering process in the HT. In order to characterise invariance, each point in a model is related to a collection of other points defining a geometric arrangement whose properties remain invariant under a particular transformation. Thus, each point in the model has an invariant characterisation which does not depend on the location of primitives but on the local geometry. The transformation is obtained by solving for the parameters according to the arrangement of points defined for a point in the image and a corresponding arrangement of points for the model with the same invariant properties. Since for a given point a great number of potential arrangements can be defined, then to avoid combinatorial complexity only selected points in the curve are considered as candidates of the arrangement. In this approach occlusion and points that do not belong to the primitive generate wrong evidence, however, there is no uncertainty associated with the data. Experimental results show that the techniques are capable of extracting arbitrary shapes under occlusion and when the image contains noise.

2. Definitions and notation

In this paper, we consider that a curve is composed of a collection of points in \mathfrak{R}^2 . The domain of the curve is given by a closed interval in \mathfrak{R}^1 . A point in the domain is mapped into a point in the curve, thus a model shape can be defined as a parametric curve $\mathbf{v}(s)$. A curve is the trace of a continuous function which defines a rule of correspondence between the parameter s and a set of points in the Euclidean plane. Here, a curve is represented in the orthogonal form $\mathbf{v}(s) = v_x(s)U_x + v_y(s)U_y$ where $U_x = [1, 0]$ and $U_y = [0, 1]$ are the two orthonormal vectors that span \mathfrak{R}^2 . The functions $v_x(s)$ and $v_y(s)$ define a correspondence between the parameters s and \mathfrak{R}^1 . Here, the notation $\mathbf{v}(s)$ can be ambiguous since it can represent either the rule of correspondence or a particular point. To distinguish between these two interpretations, we use a sub-index in the parameter s to indicate a point. That

is, $\mathbf{v}(s_0)$ is a point in $\mathbf{v}(s)$ for $s_0 \in D$ and D represents the domain of the curve. Thus, $\mathbf{v}(s_1)$ and $\mathbf{v}(s_2)$ denote two different points in a curve. The slope of the tangent to a point is denoted as $G(\mathbf{v}(s_0))$.

A transformation f is a function that maps each point $\mathbf{v}(s)$ into a new curve $f(\mathbf{v}(s))$. A parametric transformation f_a defines a family of transformations. The transformation of a curve by a parametric transformation is denoted as $f(\mathbf{a}, \mathbf{v}(s))$ where \mathbf{a} is a vector of parameters of the transformation. We name the curve obtained by applying a parametric transformation a parametric model, and it is denoted as $\mathbf{w}(s, \mathbf{a})$. The function $\mathbf{z}(s, \mathbf{a})$ denotes a parametric model with translation. In accordance with the notation of $\mathbf{v}(s)$, $\mathbf{z}(s_0, \mathbf{a})$ denotes a point in $\mathbf{z}(s, \mathbf{a})$, and $\mathbf{w}(s_0, \mathbf{a})$ denotes a point in $\mathbf{w}(s, \mathbf{a})$ for a particular value of the parameters.

In this paper, an image is a collection of points in \mathfrak{R}^2 , where a point is denoted as λ_i . Thus, λ_0 and λ_1 denote two different points in an image. The coordinates of a point are denoted as $(\lambda_{x_i}, \lambda_{y_i})$. The relationship $\lambda_0 = \mathbf{z}(s_0, \mathbf{a})$ means that the coordinates of the point λ_0 and the coordinates of the point $\mathbf{z}(s_0, \mathbf{a})$ are the same. In such case, we say that λ_0 and $\mathbf{z}(s_0, \mathbf{a})$ are corresponding points. Collections of points in an image in a parametric curve and in a parametric model are denoted by upper case letters. If a collection of points is defined based on the value of some point or points, then the dependent points are indicated in parentheses. Thus, $\mathbf{W}(\lambda_0)$ defines a collection of points which satisfy a given condition with respect to the point λ_0 .

3. Robust estimation and evidence gathering of arbitrary shapes

Model shape extraction is fundamentally a problem of analysis of regression where parametric models are fitted to observable data [20]. A shape is extracted when unknown location parameters of the parametric model are estimated. In order to formalise this aspect, we consider a parametric model defined by a shape model $\mathbf{v}(s)$ and a parametric transformation f_a . The shape model represents the border of an object and the transformation defines a family of shapes which characterises the potential manner in which a shape can appear in an image. Thus, a parametric model is defined as $\mathbf{w}(s, \mathbf{a}) = f(\mathbf{a}, \mathbf{v}(s))$, where \mathbf{a} is a vector that contains the transformation parameters. We consider that f_a does not contain any translation term, thus the model shape $\mathbf{w}(s, \mathbf{a})$ is relative to the origin. In this way a primitive in an image is represented by a translation of the parametric model. That is,

$$\mathbf{z}(s, \mathbf{a}) = \mathbf{w}(s, \mathbf{a}) + \mathbf{b}, \quad (1)$$

where the point $\mathbf{b} = (a_0, b_0)$ defines the position of the primitive in an image. Thus, shape extraction is solved by determining the parameters in \mathbf{a} and the position \mathbf{b} such

that they define a curve $z(s, \mathbf{a})$ that best fits image data. The estimation of the parameters is affected by missing data due to occlusion or noise due to background objects in the scene. Points that do not form part of the fitting curve $z(s, \mathbf{a})$ are called outliers. Since, in general, these points are not characterised by a Gaussian distribution, then techniques based on the minimisation of quadratic or absolute errors are biased. Additional problems due to change in the position of the objects have been discussed in Refs. [21,22].

The basic principle of robust methods is to perform a fitting without outliers. For images, the best fit is the curve that passes through a maximum number of points, thus any point that does not lie on the curve is considered as an outlier. In this case, we can observe that the parameters form a distribution, wherein the probability that a set of parameters provides the best fit is directly proportional to the number of points that fit the parametric model. Then the best fit is given by the mode of the distribution. The nature of outliers in images means that robust methods of shape extraction can deal with more outliers than methods of general robust regression. For example, in some cases curve fitting based on the HT can handle data comprising of 90% outliers whilst for methods of robust statistics the best that can be expected is 50% [23].

To determine the best model we can perform a verification process of all the potential parameterised models in the image or alternatively, in a more computationally efficient approach, it is possible to gather evidence of the parameters by the HT [24,25]. Evidence can be gathered by considering all the values of the parameters defined when a point in an image is matched to a point in the model. For Eq. (1), if a point in an image λ_0 is matched to a point in the model $z(s_0, \mathbf{a})$, then we can solve for the location parameter as a function of λ_0 and \mathbf{a} . That is, $\mathbf{b}(\lambda_0, \mathbf{a}) = \mathbf{b}(a_0, b_0)$ for

$$\mathbf{b}(\lambda_0, \mathbf{a}) = \lambda_0 - \mathbf{w}(s_0, \mathbf{a}). \quad (2)$$

This relationship establishes that the location can be determined given a point in the image (i.e., λ_0), and the parameters of the transformation (i.e., \mathbf{a}). Since the parameters of the transformation are unknown, then each combination of the values in \mathbf{a} defines a potential value of \mathbf{b} . That is, the local matching of a point in the model and a point in the image define a hyper-surface of potential location and parameters of the transformation. This hyper-surface defines the psf of the point λ_0 . Only one point in the psf defines the values of \mathbf{a} and \mathbf{b} which define the primitive in the image. These values are given by the intersection of all psfs formed by considering all the points in the image and their corresponding points in the model. In the HT, the intersection is computed by increasing the element of an accumulator space that

forms the trace of each psf and then searching for a maximum. Thus, the elements that are incremented in the accumulator space are given by

$$\{(\mathbf{b}, \mathbf{a}) | \mathbf{b} = \lambda_0 - \mathbf{w}(s_0, \mathbf{a}), \mathbf{w}(s_0, \mathbf{a}) = \mathbf{f}(\mathbf{a}, v(s_0))\} \quad \forall \lambda_0 \in I, \\ \forall v(s_0). \quad (3)$$

This equation defines a general HT mapping for arbitrary shapes and transformations: a set of parameters in the accumulator space is determined for each point in the image and each point in the model. However, this equation is not suited to direct implementation due to the computational complexity required to store and process the accumulator space defined by the location parameters \mathbf{b} and the transformation parameters \mathbf{a} . Additionally, Eq. (3) presupposes that we have a finite number of points in the model (i.e., discrete set). The rest of this paper will focus on reducing the computational requirements in Eq. (3) and on providing an analytic way of computing Eq. (3) for models represented by curves under similarity and affine transformations.

4. Generalised invariant HT

Although the HT can be formulated for gathering evidence of arbitrary shapes and general transformations, the computation requires excessive resources. This problem is inherent in the HT's definition and its solution has been studied by several authors [18,19]. Techniques of parameter space reduction exploit selected geometrical properties of shapes such as straight lines, circles or ellipses, to reduce the dimensionality of the accumulator space. In our case this approach cannot be followed since we cannot presuppose any geometric constraint in a general shape model. However, we can consider geometric properties of the transformation. These properties can be characterised by invariant features. Thus, the HT can be developed as a pose estimation mechanism that gathers shape evidence according to the geometric invariant properties of the points in an image.

We define a function Q (where $Q: \mathfrak{R}^2 \rightarrow \mathfrak{R}$) which computes a feature on a single point in a curve. The function Q is invariant with respect to f_a if the same feature is obtained from a point in the curve $v(s)$, and from the corresponding point in a curve obtained by applying the transformation f_a . That is, if $Q(\mathbf{f}(\mathbf{a}, v(s_0))) = Q(v(s_0))$. If Q is invariant under translation, then according to the definitions in the previous section, it is possible to establish the relationship,

$$Q(z(s_0, \mathbf{a})) = Q(\mathbf{f}(\mathbf{a}, v(s_0))) = Q(v(s_0)). \quad (4)$$

Thereby, for a point λ_0 there exists a model point $v(s_0)$ such that,

$$Q(\lambda_0) = Q(v(s_0)). \quad (5)$$

To gather evidence, we can constrain the elements of the accumulator space in Eq. (3) by considering only the elements for which Eq. (5) holds. It is important to notice that generally several points in a shape can be characterised by the same invariant value. That is, if we compute the value of an invariant feature at an image point and we determine all the points in the model which have that value, then we obtain a collection of points. The uniqueness or variety of values for the points in a shape, as well as the complexity in the computation and accuracy, are important considerations in the definition and evaluation of an invariant function.

To constrain Eq. (3) we determine, by Eq. (5), the potential points in the curve for a given image point. These points can be represented as

$$\mathbf{W}(\lambda_0) = \{\mathbf{v}(s_j) | Q(\lambda_0) - Q(\mathbf{v}(s_j)) = 0\}. \quad (6)$$

Thus, instead of gathering evidence by considering all the points of $\mathbf{v}(s)$ for each point λ_0 in I (i.e., $\forall \lambda_0 \in I, \forall \mathbf{v}(s_0)$), evidence can be gathered by considering only the points in $\mathbf{W}(\lambda_0)$ for each point λ_0 in I (i.e., $\forall \lambda_0 \in I, \forall \mathbf{v}(s_0) \in \mathbf{W}(\lambda_0)$). A further and more significant simplification can be achieved if we consider that the matching process in Eq. (6) is an invariant mapping, so we can find the transformation by solving for \mathbf{a} in $Q(\lambda_0) = Q(\mathbf{f}(\mathbf{a}, \mathbf{v}(s_0)))$. The solution is multi-valued. That is, for each point $\lambda_0 \in I$, and each point $\mathbf{v}(s_0) \in \mathbf{W}(\lambda_0)$ we can determine a different transformation that maps the model point into the image point. The set of solutions is denoted as $f^A(\lambda_0, \mathbf{v}(s_0))$. That is,

$$f^A(\lambda_0, \mathbf{v}(s_0)) = \{\mathbf{a} | Q(\lambda_0) = Q(\mathbf{f}(\mathbf{a}, \mathbf{v}(s_0)))\}. \quad (7)$$

Thus, the position of the image point λ_0 and the transformation parameters can be gathered independently. Accordingly the HT mapping in Eq. (3) can be redefined as

$$\{\mathbf{b} | \mathbf{b} = \lambda_0 - \mathbf{f}(\mathbf{a}, \mathbf{v}(s_0)), \mathbf{a} \in f^A(\lambda_0, \mathbf{v}(s_0))\} \quad \forall \lambda_0 \in I, \\ \forall \mathbf{v}(s_0) \in \mathbf{W}(\lambda_0). \quad (8)$$

Consequently, if we establish an invariant function Q for a family of transformations f_a , we can characterise equivalent objects, and thereby solve the extraction problem by searching for the location of a shape in a 2D accumulator space. The size of this accumulator is independent of the complexity of the object or of the generality of the transformation. The transformation parameters can be determined by gathering evidence according to the mapping,

$$\{\mathbf{a} | \mathbf{a} = f^A(\lambda_0, \mathbf{v}(s_0))\} \quad \forall \lambda_0 \in I, \quad \forall \mathbf{v}(s_0) \in \mathbf{W}(\lambda_0). \quad (9)$$

Alternatively, we can exploit the values of the location parameters to solve for some of the parameters in \mathbf{a} , reducing the computational burden. This approach will be considered in further detail in Section 5. According to Eqs. (8) and (9) the extraction process described by three

general steps in Section 1 can be defined for the extraction of a shape via a general invariant form of the HT as follows. First, for each point in the image identify a potential set of points in the model by matching invariant features according to Eq. (6). Second, the image transformation that maps the point in the image and the point in the model is determined by Eq. (7). Finally, evidence of the location position and transformation parameters is gathered by the HT mappings in Eqs. (8) and (9). The remainder of this paper will be concerned with characterising the function Q and solving the analytic expression in (6) and (7) when f_a is defined by similarity and affine transformations.

5. Similarity transformations

5.1. Parametric model

The transformation f_a defines a family of shapes that can be located in an image. The generality of this family is related to the generality of the transformation. Similarity transformations are defined by a change in rotation and scale and are of practical significance since they can be used to model changes in the distance between an object and a camera, as well as variation in the orientation within the plane of view.

A parametric model for similarity transformations is defined by multiplying a model shape $\mathbf{v}(s)$ by a scalar value and by a rotation matrix. Thus, the mapping in Eq. (2) is given by $\mathbf{a} = (l, \rho)$, where l and ρ are the scale and rotation parameters, respectively. Here, the parametric model is represented by an orthogonal decomposition of the form

$$\mathbf{w}(s, \mathbf{a}) = w_x(s, \mathbf{a})\mathbf{U}_x + w_y(s, \mathbf{a})\mathbf{U}_y \quad (10)$$

for $w_x(s, \mathbf{a}) = lR_x(s, \rho)$, $w_y(s, \mathbf{a}) = lR_y(s, \rho)$ and $[R_x(s, \rho) \ R_y(s, \rho)]$ is the result of multiplying the model vector $[\mathbf{v}_x(s) \ \mathbf{v}_y(s)]$ by a rotation matrix.

5.2. Geometric invariance

The definition of invariance is based on a characterisation of a shape by a function Q . For similarity transformations this function is derived from the concept of angle. The measure of an angle is the fundamental invariant for similarity transformations since its value does not change with changes in scale or rotation. Clearly, a single point does not provide enough information for an invariant geometric characterisation, so it is necessary to consider the relationship of the point with other points in the curve. In this way, the invariant of a point can be defined by considering the angle formed by the lines that join the point with other two points in the curve. In order to distinguish between the point that is characterised by the invariant, we use the sub-index s_0 , while the points

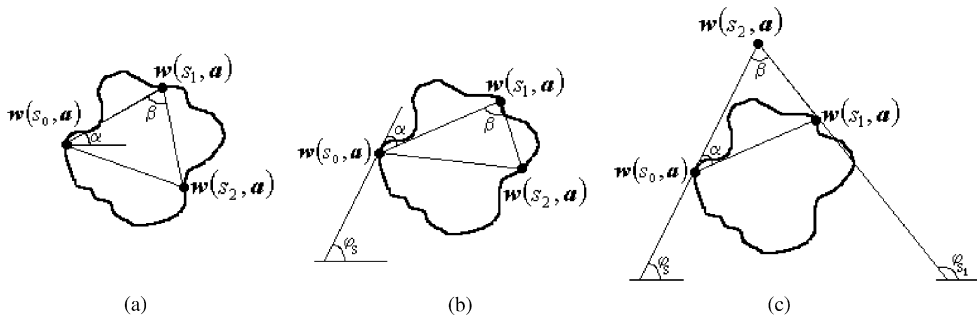


Fig. 1. Arrangements of points. (a) General triangle. (b) Triangle defined relative to gradient direction. (c) Natural triangle defined by a pole–polar relationship.

used to define the angle are denoted by the sub-indices s_1 and s_2 . An invariant characterisation of the point $w(s_0, \mathbf{a})$ in the parametric model is given by the tangent of one angle in the triangle formed by the three points. That is,

$$Q_{sim}(w(s_0, \mathbf{a}), w(s_1, \mathbf{a}), w(s_2, \mathbf{a})) = \frac{V_{x_2} V_{y_1} - V_{x_1} V_{y_2}}{V_{x_2} V_{x_1} - V_{y_1} V_{y_2}}, \tag{11}$$

for

$$V_{x_i} = w_x(s_i, \mathbf{a}) - w_x(s_0, \mathbf{a}), \quad V_{y_i} = w_y(s_i, \mathbf{a}) - w_y(s_0, \mathbf{a}).$$

For completeness, the proof is included in Appendix A. It is important to notice that although the number of parameters in Q has been increased with respect to the definition presented in Section 4, the function maintains the same meaning as in Eq. (8). That is, it provides a characterisation of a single point in the model (i.e., for $w(s_0, \mathbf{a})$).

Although Eq. (11) defines a unique invariant for the points in the model, there exist different ways in which the points can be chosen. Three alternative ways are shown in Fig. 1. In a straightforward approach, the three points in Eq. (11) can represent points of the shape which satisfy some given geometric condition. Since angles remain invariant under similarity transformations, then we can consider the arrangement of points which forms a triangle with fixed angles as shown in Fig. 1(a) [26]. Thus, the points $w(s_1, \mathbf{a})$ and $w(s_2, \mathbf{a})$ can be chosen by searching for points in the lines that pass from $w(s_0, \mathbf{a})$ and that have a fixed orientation defined by the fixed angles α and β . The problem with this arrangement is that it is dependent on rotation, since the angles are not measured relative to the orientation of the shape. Thus, when the shape is rotated, the position of $w(s_0, \mathbf{a})$ and $w(s_1, \mathbf{a})$ cannot be determined without involving the rotation parameter. That is, given a point $w(s_0, \mathbf{a})$ and a fixed value of α , the point $w(s_1, \mathbf{a})$ must satisfy the relationship $(w_y(s_1, \mathbf{a}) - w_y(s_0, \mathbf{a})) / (w_x(s_1, \mathbf{a}) - w_x(s_0, \mathbf{a})) = \alpha + \rho$. Since ρ is unknown, any point $w(s_1, \mathbf{a})$ can be an element

of the arrangement. Thus, it is necessary to consider all points combinations for all values of the rotation.

A better definition of the points that characterise the invariance relationship can be obtained by defining their position with reference to the point $w(s_0, \mathbf{a})$. Two of these definitions are shown in Fig. 1(b) and (c). In Fig. 1(b) the point $w(s_1, \mathbf{a})$ is defined relative to the gradient direction of the point $w(s_0, \mathbf{a})$. Thus, the point $w(s_1, \mathbf{a})$ can be located independent of rotation by searching on a line passing through $w(s_0, \mathbf{a})$ whose slope is $(w_y(s_1, \mathbf{a}) - w_y(s_0, \mathbf{a})) / (w_x(s_1, \mathbf{a}) - w_x(s_0, \mathbf{a})) = G(w(s_1, \mathbf{a})) + \rho$. Since β is relative to the line which joins the points $w(s_0, \mathbf{a})$ and $w(s_1, \mathbf{a})$, then the point $w(s_2, \mathbf{a})$ is also determined, again independent of the rotation.

The collection of points shown in Fig. 1(c) uses the value of the gradient direction at two points in the curve to define an invariant angle. In this case, the third point is defined by the intersection of the tangents to two points in the curve. The advantage of this definition is that it characterises a point by only one other point in the image and this point does not require a particular geometric relationship to the other two. The definition of the third point is based on the pole–polar form of the shape and it has been previously used for the extraction of circles and ellipses [27,28]. In [29–32] this relationship is exploited to define indexed tables suited to invariant extraction of shapes by the HT. These tables store the position of the centre of a shape as a function of the invariant properties in the pole–polar form. The geometry used in Ref. [31] considers pairs of edge points with the same gradient direction and by fixing β to 90° . The disadvantage of this definition is that since a point can only be paired to other point with the same slope, the technique can suffer when there are missing data. In Ref. [29] a more general relationship is defined by considering complementary angles of α and β . Since the table includes all the possible values of α and β , the complexity of the technique is similar to the one obtained by the arrangement of points shown in Fig. 1(a). In Ref. [32] different complementary angles of α and β are used to define a 2D indexed table. In this

case, invariant properties are not exploited and the technique is associated with a four-dimensional parameter space. As in the case of Ref. [29] geometric relationships are used to solve for the parameters of scale and rotation necessary to obtain the position of the shape and pairs of points are considered in a computationally intensive process. In the next section, we consider the arrangement of points in Fig. 1(c) to formulate the analytic definition of the collection of points in $W(\lambda_0)$.

5.3. Local matching

According to Eq. (6), invariant shape extraction requires establishment of a collection of points $W(\lambda_0)$ which are composed of all the points $\mathbf{v}(s_0)$ for which the relationship $Q_{sim}(\lambda_0, \lambda_1, \lambda_2) = Q_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1), \mathbf{v}(s_2))$ holds. In the case of the arrangement of points in Fig. 1(c), the invariant feature only depends on two points. Thus, we can consider the simplified relationship $Q_{sim}(\lambda_0, \lambda_1) = Q_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1))$. The form of each invariance in this relationship can be obtained by simplifying Eq. (11) by using the pole–polar relationships. Thus, invariant features for the points in the model and for the points in the image can be expressed as a function of the position of the points and their gradient direction. There exist two different ways in which the invariance defined in Eq. (11) can be used to define invariance features in the arrangement shown in Fig. 1(c): (1) as a measure of the angle α , between the line joining s_0 and s_1 , and the tangent at s_1 ; and (2) as a measure of the angle β defined by the intersection of the lines. For the first case we have that

$$Q_{sim}(\lambda_0, \lambda_1) = \frac{G(\lambda_0) - \gamma(s_0, s_1)}{1 + G(\lambda_0)\gamma(s_0, s_1)},$$

$$Q_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1)) = \frac{G(\mathbf{v}(s_0)) - \phi(s_0, s_1)}{1 + G(\mathbf{v}(s_0))\phi(s_0, s_1)} \quad (12)$$

for $\gamma(s_0, s_1) = (\lambda_{y1} - \lambda_{y0}) / (\lambda_{x1} - \lambda_{x0})$ and $\phi(s_0, s_1) = (v_y(s_1) - v_y(s_0)) / (v_x(s_1) - v_x(s_0))$.

As a measure of the angle β the invariant arrangement of points is given by

$$Q'_{sim}(\lambda_0, \lambda_1) = \frac{G(\lambda_0) - G(\lambda_1)}{1 + G(\lambda_0)G(\lambda_1)},$$

$$Q'_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1)) = \frac{G(\mathbf{v}(s_0)) - G(\mathbf{v}(s_1))}{1 + G(\mathbf{v}(s_0))G(\mathbf{v}(s_1))}. \quad (13)$$

Thus, from Eq. (6), we observe that the problem of obtaining a definition of $W(\lambda_0)$ can be formulated as the problem of obtaining the points $\mathbf{v}(s_0)$ such that the invariant features $Q_{sim}(\lambda_0, \lambda_1)$ and $Q'_{sim}(\lambda_0, \lambda_1)$ computed from an image are equal to the invariant features $Q_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1))$ and $Q'_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1))$ in the model. That is, for two points in an image, we have a pair

of equations that defines a pair of points in the model with the same invariant characterisation. More formally,

$$W_{sim}(\lambda_0, \lambda_1) = \{\mathbf{v}(s_0), \mathbf{v}(s_1) | Q_{sim}(\lambda_0, \lambda_1) = Q_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1)), Q'_{sim}(\lambda_0, \lambda_1) = Q'_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1))\}. \quad (14)$$

The pair of simultaneous equations can be written as

$$Q_{sim}(\lambda_0, \lambda_1) = (G(\mathbf{v}(s_0)) - \phi(s_0, s_1)) / (1 + G(\mathbf{v}(s_0))\phi(s_0, s_1)),$$

$$Q'_{sim}(\lambda_0, \lambda_1) = (G(\mathbf{v}(s_0)) - G(\mathbf{v}(s_1))) / (1 + G(\mathbf{v}(s_0))G(\mathbf{v}(s_1))). \quad (15)$$

By considering a function D that obtains the derivative of the component at a point in the curve (i.e., $G(\mathbf{v}(s_0)) = D(v_y(s_0)) / D(v_x(s_0))$) and the definition of $\phi(s_0, s_1)$, the equations in Eq. (15) can be rewritten as

$$(v_x(s_1) - v_x(s_0))(Q_{sim}(\lambda_0, \lambda_1)D(v_x(s_0)) - D(v_y(s_0)))$$

$$+ (v_y(s_1) - v_y(s_0))(Q_{sim}(\lambda_0, \lambda_1)D(v_y(s_0))$$

$$+ D(v_x(s_0))) = 0.$$

$$D(v_x(s_1))D(v_y(s_0)) - D(v_x(s_0))D(v_y(s_1))$$

$$- Q'_{sim}(\lambda_0, \lambda_1)(D(v_x(s_0))D(v_x(s_1))$$

$$+ D(v_y(s_0))D(v_y(s_1))) = 0. \quad (16)$$

5.4. Parameters of the transformation

A shape's rotation and scale can be obtained by matching two points of the curve with two points in the model. That is, for each pair of points $\lambda_0, \lambda_1 \in I$, and each pair of points $\mathbf{v}(s_0)$ and $\mathbf{v}(s_1)$ such that $Q_{sim}(\lambda_0, \lambda_1) = Q_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1))$ and $Q'_{sim}(\lambda_0, \lambda_1) = Q'_{sim}(\mathbf{v}(s_0), \mathbf{v}(s_1))$, we can define a function of the form of Eq. (7) as

$$f^A(\lambda_0, \lambda_1, \mathbf{v}(s_0), \mathbf{v}(s_1)) = \{l, \rho | Q_{sim}(\lambda_0, \lambda_1) = Q_{sim}(f((l, \rho), \mathbf{v}(s_0)), f((l, \rho), \mathbf{v}(s_1)))\}. \quad (17)$$

By using the geometric properties of a pair of points, it is possible to obtain an explicit form of Eq. (17) as

$$f^A(\lambda_0, \lambda_1, \mathbf{v}(s_0), \mathbf{v}(s_1)): \rho = \tan^{-1} \left(\frac{\lambda_{y1} - \lambda_{y0}}{\lambda_{x1} - \lambda_{x0}} \right)$$

$$- \tan^{-1} \left(\frac{v_y(s_1) - v_y(s_0)}{v_x(s_1) - v_x(s_0)} \right), l = \frac{|\lambda_1 - \lambda_0|}{|\mathbf{v}(s_1) - \mathbf{v}(s_0)|}. \quad (18)$$

5.5. Implementation and examples

In order to gather evidence for a model shape under a similarity transformation, we can use the definitions

in Eqs. (14) and (17) in Eqs. (8) and (9). Thus, evidence can be gathered by using a pair of mappings from the image space into two independent 2D accumulator spaces. In implementation, the evidence gathering process can be divided into four steps: (1) for each point in the image λ_0 chose another point λ_1 and compute $Q_{sim}(\lambda_0, \lambda_1)$ and $Q'_{sim}(\lambda_0, \lambda_1)$ according to Eqs. (12) and (13); (2) use these values in the system of Eq. (16) to find all the points $v(s_0)$ and $v(s_1)$ that satisfy Eq. (14); (3) use the points λ_0 , λ_1 , $v(s_0)$ and $v(s_1)$ in Eq. (18) to find the parameters of the transformation and increment the associated element in the accumulator space; finally, (4) compute the location parameter according to Eq. (8). The location accumulator has the same resolution than the image. In our examples the scale and rotation accumulators have half the size of the image in the scale axis and one degree resolution in the rotation axis. Since there are numerical errors in the computation of the gradient direction (in our implementation errors are about 0.2 radians) and thus in the invariance, it is convenient to gather evidence in a collection of cells determined using a Gaussian distribution.

Even though each point can be characterised by taking only one other point, it is necessary to ensure that both points belong to the same primitive. Accordingly, the point λ_1 is only selected if it is within a specified distance of λ_0 . In our implementation, λ_1 must be closer than one quarter of the image length. We develop the curve $v(s)$ as an orthogonal Fourier expansion as presented in Ref. [33]. In this representation, derivatives are easily computed. An explanation of the use of elliptic Fourier coefficients for model extraction can be found in Ref. [34]. The system of equations in Eq. (16) can be solved by using a successive approximation method. In our implementation, we use an alternative approach that finds a set of potential values of s_0 and s_1 and then verifies that the relationships in Eq. (16) are true. Potential values of s_0 and s_1 are obtained by considering gradient direction information. Based on gradient direction information it is possible to obtain the pair of independent equations

$$G(\lambda_0)D(g(s_0, \rho)) - D(h(s_0, \rho)) = 0,$$

$$G(\lambda_1)D(g(s_1, \rho)) - D(h(s_1, \rho)) = 0. \quad (19)$$

These equations depend on the value of ρ , thus each possible value of rotation provides a potential value of the parameters s_0 and s_1 . These potential values are considered in Eq. (16). Fig. 2 shows an example of the accumulation process for similarity transformations. This figure contains a shape and a collection of random dots. We use similar images to estimate the minimum signal-to-noise ratio for which the maximum in the accumulator still yields the correct location of a shape. Accordingly, the number of points that forms the shape

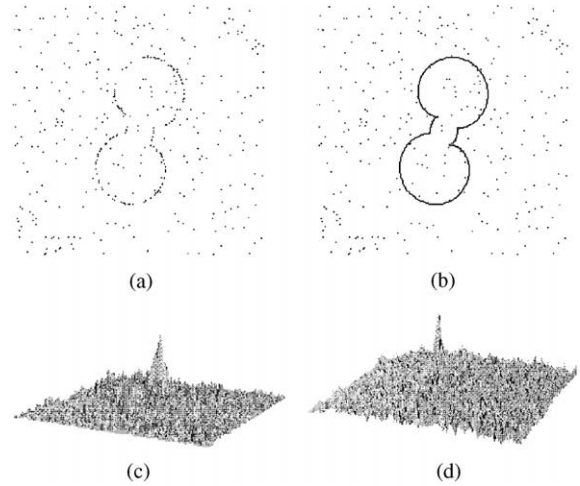


Fig. 2. (a) Example of an image where only 35% of the data belong to the primitive. (b) Result of the extraction process. (c) Location accumulator. (d) Scale and rotation accumulator.

is reduced in the same proportion to the increase in noise points. In the example in Fig. 2(a) 65% of the data correspond to noise. Fig. 2(b) shows the result of the extraction process superimposed on the original image. Fig. 2(c) shows that the location accumulator contains a well-defined peak. (In our experiments the technique can actually provide an accurate extraction when almost 90% of image points are noise. With this percentage it is difficult even to locate a shape by visual inspection.) Fig. 2(d) shows the accumulation obtained by gathering evidence of the scale and rotation parameters. The peaks in the accumulators define an accurate value of the parameters of the shape. In this example, gradient direction was not computed from the image, but from the analytic definition of the curve. This allows us to quantify only the error due to clutter and occlusion.

Fig. 3 shows an example of the accumulation process on a real image. The model shape was obtained from a binary image of 128×128 pixels. Fig. 3(a) shows the image model and the parameterised model. This parameterised model was used to accumulate evidence for the edge information computed from the image in Fig. 3(b) and shown in Fig. 3(c). In this image the number of points that do not belong to the primitive is significantly more than the model points and some edge data of the shape have not been detected. The result of the extraction process is presented in Fig. 3(d) and superimposed as a thick black border. Fig. 3(e) shows the final accumulator for the centre parameters, and Fig. 3(f) shows the final accumulator for the rotation and scale parameters. The parameters given by the peaks of these accumulators were used to define the shape

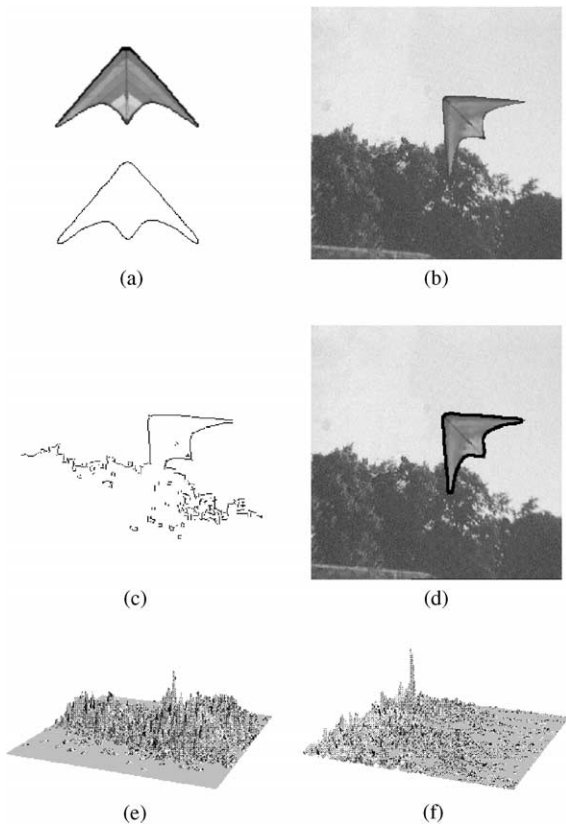


Fig. 3. Shape extraction. (a) Model shape. (b) Raw image. (c) Image edges. (d) Result. (e) Location accumulator. (f) Scale and rotation accumulator.

superimposed in Fig. 3(d) which matches with the image data.

6. Affine transformations

6.1. Parametric model

Although similarity transformations constitute a useful description for shape extraction, more general transformations are required to model the changes in position and orientation of a shape in three-dimensional space. Affine transformations are defined by a linear mapping and can provide an accurate approximation of the three-dimensional orientation of a planar shape when the depth of an object is insignificant compared to the viewing distance. Since this type of transformation is more general than similarity transformations, it contains more parameters and the geometry of the definition of the invariant features is more complex.

A parametric model for an affine transformation can be obtained by multiplying a model shape $v(s)$ by a

linear transformation. Thus, Eq. (2) can be defined by the four transformation parameters $\mathbf{a} = (A, B, C, D)$ for the orthogonal components of the parametric model in Eq. (10) defined as,

$$\begin{aligned} w_x(s, \mathbf{a}) &= Av_x(s) + Bv_y(s), \\ w_y(s, \mathbf{a}) &= Cv_x(s) + Dv_y(s). \end{aligned} \quad (20)$$

6.2. Geometric invariance

The family of shapes defined by affine transformations is a linear combination of point coordinates. Thus, geometric relationships based on properties computed on a pair of parallel straight lines, such as slope and distance ratio, remain invariant. Although affine invariant features are defined by properties of straight lines, this does not mean that the extraction process implies the use of primitives of polygonal shapes. That is, in a manner analogous to the inclusion of invariance under similarity transformations, invariance under affine transformations can be defined by considering a collection of points that forms a particular geometric arrangement. The distance ratio between two parallel lines can be defined by four points, thus an invariant characterisation of a point $w(s_0, \mathbf{a})$ in the parametric model can be obtained by considering three additional points. If the two points $w(s_0, \mathbf{a})$ and $w(s_1, \mathbf{a})$ define a parallel line to the line formed by the points $w(s_2, \mathbf{a})$ and $w(s_3, \mathbf{a})$, then the distance ratio invariant for a point $w(s_0, \mathbf{a})$ can be defined as

$$\begin{aligned} Q_{aff}(w(s_0, \mathbf{a}), w(s_1, \mathbf{a}), w(s_2, \mathbf{a}), w(s_3, \mathbf{a})) \\ = \frac{w_x(s_1, \mathbf{a}) - w_x(s_0, \mathbf{a})}{w_x(s_3, \mathbf{a}) - w_x(s_2, \mathbf{a})} = \frac{w_y(s_1, \mathbf{a}) - w_y(s_0, \mathbf{a})}{w_y(s_3, \mathbf{a}) - w_y(s_2, \mathbf{a})}. \end{aligned} \quad (21)$$

For completeness, the proof of this invariance is given in Appendix B. Although Eq. (11) defines a unique invariance, it is possible to consider alternative information in the image to define the four points in the geometric arrangement. Fig. 4(a) shows a geometric arrangement wherein the four points that define the parallel lines are points of the shape. The arrangement in Fig. 4(b) reduces the number of points required to define the invariant feature in a manner analogous to the reduction presented in Fig. 1(c). In this case, only three points are on the shape, whilst the fourth one is determined indirectly by using gradient direction information. In Fig. 4(b), the point $w(s_1, \mathbf{a})$ is defined by the intersection of the tangent lines to the points $w(s_0, \mathbf{a})$ and $w(s_3, \mathbf{a})$.

The invariance properties of the lines in Fig. 4(b) can be proved by extending the results in Appendix B. By denoting m_{s_0} and m'_{s_0} as the gradient direction in the point $w(s_0, \mathbf{a})$ before and after the transformation, and

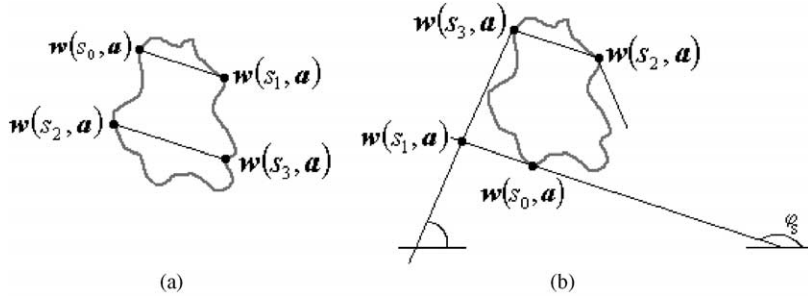


Fig. 4. Arrangement of points to define two parallel lines. (a) Four point definition. (b) Three point definition.

by considering the definition in Eq. (20) we have that

$$m'_{s_0} = \frac{w'_y(s_0)}{w'_x(s_0)} = \frac{Cv'_x(s_0) + Dv'_y(s_0)}{Av'_x(s_0) + Bv'_y(s_0)} = \frac{C + Dm_{s_0}}{A + Bm_{s_0}}. \quad (22)$$

According to the results in Appendix B the slope of the line formed by the points $w(s_2, a)$ and $w(s_3, a)$, after the transformation is

$$m'_{s_2s_3} = \frac{C + Dm_{s_2s_3}}{A + Bm_{s_2s_3}}. \quad (23)$$

Since the gradient direction and the slope of the line are equal before the transformation, then $m_{s_2s_3} = m_{s_0}$. Thus, according to Eqs. (22) and (23) $m'_{s_2s_3} = m'_{s_0}$ which means that the lines are parallel after the transformation.

6.3. Local matching

By considering the arrangement of points in Fig. 4(b), and according to Eq. (6), the collection of points $W(\lambda_0)$ is composed of all the points $v(s_0)$ for which the relationship $Q_{aff}(\lambda_0, \lambda_2, \lambda_3) = Q_{aff}(v(s_0), v(s_2), v(s_3))$ holds. The invariant function $Q_{aff}(\cdot)$ in this relationship can be developed as a particular case of Eq. (21) defined according to the geometric relationships in Fig. 4(b). That is,

$$Q_{aff}(\lambda_0, \lambda_2, \lambda_3) = \frac{\lambda_{x_1} - \lambda_{x_0}}{\lambda_{x_3} - \lambda_{x_2}},$$

$$Q_{aff}(v(s_0), v(s_2), v(s_3)) = \frac{v_x(s_1) - v_x(s_0)}{v_x(s_3) - v_x(s_2)} \quad (24)$$

for

$$\begin{aligned} \lambda_{x_1} &= (\lambda_{y_0} - \lambda_{y_3} + G(\lambda_3)\lambda_{x_3} + G(\lambda_0)\lambda_{x_0}) / (G(\lambda_3) \\ &\quad + G(\lambda_0)), \lambda_{y_1} \\ &= (G(\lambda_3)G(\lambda_0)(\lambda_{x_3} - \lambda_{x_0}) + G(\lambda_3)\lambda_{y_0} \\ &\quad - G(\lambda_0)\lambda_{y_3}) / (G(\lambda_3) + G(\lambda_0)) \end{aligned} \quad (25)$$

and $v(s_1)$ is defined in a similar fashion. By considering the second part of Eq. (21) we obtain the invariance definition,

$$Q'_{aff}(\lambda_0, \lambda_2, \lambda_3) = \frac{\lambda_{y_1} - \lambda_{y_0}}{\lambda_{y_3} - \lambda_{y_2}},$$

$$Q'_{aff}(v(s_0), v(s_2), v(s_3)) = \frac{v_y(s_1) - v_y(s_0)}{v_y(s_3) - v_y(s_2)}. \quad (26)$$

Thus, $W(\lambda_0)$ is defined by all the points in the model such that the invariant features $Q_{aff}(\lambda_0, \lambda_2, \lambda_3)$ and $Q'_{aff}(\lambda_0, \lambda_2, \lambda_3)$ computed from an image are equal to the invariant features $Q_{aff}(v(s_0), v(s_2), v(s_3))$ and $Q'_{aff}(v(s_0), v(s_2), v(s_3))$ on the model shape. That is,

$$\begin{aligned} W_{aff}(\lambda_0, \lambda_2, \lambda_3) &= \{v(s_0), v(s_2), v(s_3) | Q_{aff}(\lambda_0, \lambda_2, \lambda_3) \\ &= Q_{aff}(v(s_0), v(s_2), v(s_3)), \\ Q'_{aff}(\lambda_0, \lambda_2, \lambda_3) &= Q'_{aff}(v(s_0), v(s_2), v(s_3))\}. \end{aligned} \quad (27)$$

This equation and the geometry of the distance ratio define a system of three simultaneous equations. That is

$$\begin{aligned} (Q_{aff}(\lambda_0, \lambda_2, \lambda_3))(v_x(s_3) - v_x(s_2)) - v_x(s_1) + v_x(s_0) &= 0, \\ (Q_{aff}(\lambda_0, \lambda_2, \lambda_3))(v_y(s_3) - v_y(s_2)) - v_y(s_1) + v_y(s_0) &= 0, \\ G(v(s_0))(v_x(s_3) - v_x(s_2)) - v_y(s_3) + v_y(s_2) &= 0. \end{aligned} \quad (28)$$

The first two equations constrain the two orthonormal coordinates of the curve. The third equation ensures that the arrangement of points forms two parallel lines. That is, the gradient direction at the point λ_0 equals the slope of the line which joins the points λ_2 and λ_3 .

6.4. Parameters of the transformation

The matching of three points in the curve to three points in the model is sufficient to obtain the parameters A, B, C, D of the transformation. The solution of the parameters of the transformation is defined by the function in Eq. (7), and can be developed in a manner analogous

to Eq. (17). In this case the parameters are obtained by the system

$$f^A(\lambda_0, \lambda_2, \lambda_3, \mathbf{v}(s_0), \mathbf{v}(s_2), \mathbf{v}(s_3)):$$

$$\begin{aligned} \lambda_{x_0} - \lambda_{x_2} &= A(v_x(s_0) - v_x(s_2)) + B(v_y(s_0) - v_y(s_2)), \\ \lambda_{x_0} - \lambda_{x_3} &= A(v_x(s_0) - v_x(s_3)) + B(v_y(s_0) - v_y(s_3)), \\ \lambda_{y_0} - \lambda_{y_2} &= C(v_x(s_0) - v_x(s_2)) + D(v_y(s_0) - v_y(s_2)), \\ \lambda_{y_0} - \lambda_{y_3} &= C(v_x(s_0) - v_x(s_3)) + D(v_y(s_0) - v_y(s_3)). \end{aligned} \quad (29)$$

This system defines a gathering process in a four-dimensional parameter space. This space can be reduced by using the information of the position of a shape. Thus, after the location parameters have been obtained, we can gather evidence of the transformation parameters by using the known shape location. This two-stage approach simplifies Eq. (29) into two independent systems of two equations. That is, the gathering process can be performed in two 2D accumulators. The simplification of Eq. (29) by including the location parameters is given by

$$f^A(\lambda_0, \lambda_2, \mathbf{v}(s_0), \mathbf{v}(s_2)):$$

$$\begin{aligned} \lambda_{x_0} - a_0 &= Av_x(s_0) + Bv_y(s_0), \\ \lambda_{y_0} - b_0 &= Cv_x(s_0) + Dv_y(s_0), \\ \lambda_{x_2} - a_0 &= Av_x(s_2) + Bv_y(s_2), \\ \lambda_{y_2} - b_0 &= Cv_x(s_2) + Dv_y(s_2). \end{aligned} \quad (30)$$

6.5. Implementation and examples

The definitions presented in Eqs. (27) and (30) can be used in Eqs. (8) and (9) to develop an evidence gathering process for affine transformations. Evidence can be obtained by considering the invariant properties of three image points. The definition of Eqs. (27) and (30) in Eqs. (8) and (9) define three mappings from the image space into three 2D accumulator spaces. According to these definitions, the process of evidence gathering can be divided into five main steps as follows: (1) for each point λ_0 in the image choose a pair of points λ_2 and λ_3 such that the line which passes through the points has the same slope of the gradient direction of the point λ_0 . From these, compute the value of $Q_{aff}(\lambda_0, \lambda_2, \lambda_3)$ and $Q'_{aff}(\lambda_0, \lambda_2, \lambda_3)$ according to Eqs. (24) and (26), respectively; (2) use these values in the system of equations in Eq. (28) to find all the points $\mathbf{v}(s_0)$, $\mathbf{v}(s_2)$ and $\mathbf{v}(s_3)$ that satisfy Eq. (27); (3) use the points λ_0 , λ_2 , λ_3 , $\mathbf{v}(s_0)$, $\mathbf{v}(s_2)$ and $\mathbf{v}(s_3)$ in Eq. (29) to solve for the parameters A, B, C, D ; (4) compute the location parameter according to Eq. (8) and increment the associated element in the corresponding position in the accumulator space; finally

Eq. (5) after all evidence has been gathered and the location parameter is known, repeat step Eq. (1) and use the points in the system of equations in Eq. (30) to solve for the parameters of the transformation. Based on these parameters increase the corresponding element in the accumulator that represents the parameters A and C , and in the accumulator that represents the parameters B and D . The values of these parameters are unbounded. In our implementation, we used the *atan* function to transform the values of A, B, C and D to values between 0° and 360° . Thus, the accumulator has a resolution of 1° . A more efficient and accurate implementation can be achieved by using multiresolution techniques [35].

In this implementation, the solution for Eq. (28) imposes significant computational load since it contains three unknown values. In our implementation, we reduce the complexity by considering that the point λ_2 can only be chosen from a selected collection of landmark points. Landmark points are identified by their high curvature. Thus, for each point λ_0 we consider a point λ_2 in the image with a high curvature. In order to select the point λ_3 from the image, we search for it on the trace of the straight line which passes through the point λ_2 and whose slope is equal to the slope of the point λ_0 . In order to solve the system in Eq. (28), we consider each point $\mathbf{v}(s_2)$ with high curvature in the model. We find the point $\mathbf{v}(s_0)$ by considering that $G(\mathbf{v}(s_0)) = G(\lambda_0) + \rho$ for fixed values of ρ . Thus, for each combination of $\mathbf{v}(s_2)$ (i.e. for all the curvature points) and $\mathbf{v}(s_0)$ (i.e. all points whose gradient is $G(\lambda_0) + \rho$) we determine the point $\mathbf{v}(s_3)$ that satisfies the last relationship in Eq. (28). This triplet of points is used to gather evidence only if the two first conditions in Eq. (28) are satisfied.

Fig. 5 shows an example of the evidence gathered for affine transformations. Fig. 5(a) shows the model shape used in this example. Points of high curvature are defined as the zeroes of the derivative of the tangent angle of the Fourier expansion and they are marked with small circles. To quantify which percentage of points is necessary to obtain evidence that can locate the shape we generated a collection of synthetic images containing a transformed version of the model and random noise. The number of points that form the model shape was reduced in proportion to the amount of noise points added to the image. In our experiments we maintained three points of high curvature in correct positions (those which lie on the shape), whilst the other five were positioned at random. In the example shown in Fig. 5, the model shape is composed of 35% of the points (i.e., 65% outliers). Fig. 5(c) shows the result obtained by the gathering process superimposed on the edge image. The accumulator presented in Fig. 5(d) shows a well-defined peak that defines the location of the centre of the shape. The accumulators presented in Fig. 5(e) and (f) correspond to the parameters of the transformation. In our experiments, if

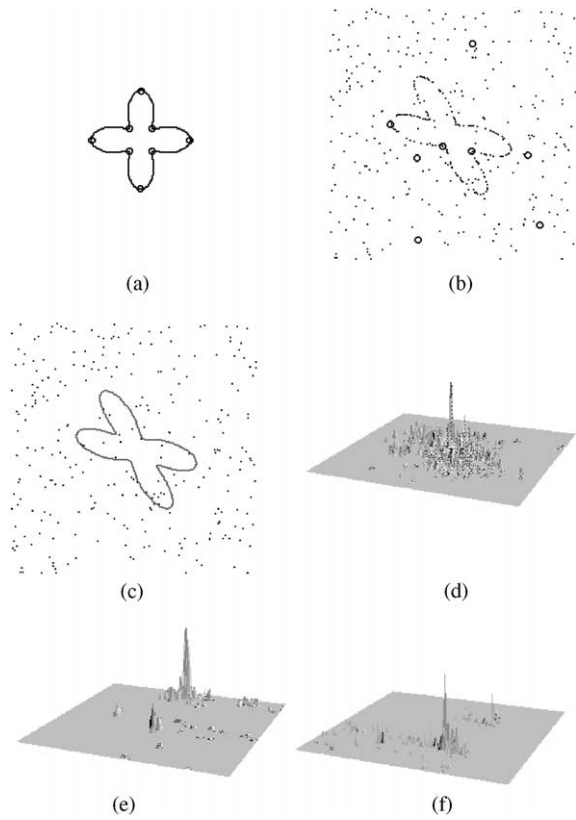


Fig. 5. (a) Model shape. (b) Example of an image where 35% of the data belong to the primitive. (c) Result of the extraction process. (d) Location accumulator. (e) Result of the accumulation of the parameters A and C of the transformation. (f) Result of the accumulation of the parameters B and D of the transformation.

we can locate at least three points of high curvature in their correct position and if the point in the shape, then an accurate value of the parameters of the primitive in the image can be obtained when almost only 15% of data actually belong to the primitive. Similar to the example in Fig. 2, gradient direction was computed from the analytic definition of the curve.

An example of the extraction process applied to a real image is presented in Fig. 6. The model used for the extraction process is the same one used in Fig. 3. In this model, we identified four points of high curvature. We apply the extraction process to the image shown in Fig. 6(a). In this image, the object suffers a transformation which changes its relative orientation. The edges of this image are shown in Fig. 6(b). Small circles indicate land-mark points, those with high curvature. Only some points in the model can be identified and one is actually in error (the one in the centre of the shape). The result of the extraction process is shown in Fig. 6(c) superimposed as a thick black border on the original image. The

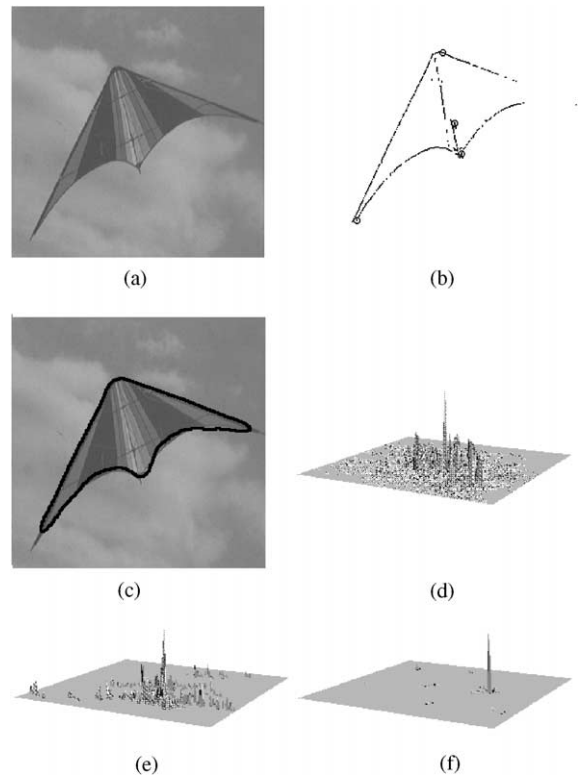


Fig. 6. Shape extraction. (a) Raw image. (b) Image edges and high curvature points. (c) Result. (d) Location accumulator. (e) Accumulator for the parameters A and C . (f) Accumulator for the parameters B and D .

parameters of the shape are determined by the maxima of the accumulators in Fig. 6(d)–(f). These accumulators show that the formulation can provide an effective evidence gathering process, since there is a good match between the result and the primitive in the image, and the accumulator peak is well defined.

7. Conclusions

In this paper, we have developed a formulation and methodology for including invariance in a general form of the Hough transform. This formulation extends the definition of the HT in two ways. First, it includes a parameterised model shape composed of a model shape and a transformation. Second, it includes invariant properties of the transformation to reduce the dimensionality of the HT mapping. The basis of our methodology is a continuous shape description which affords immunity to the discretisation effects usually experienced with tabular shape representations. The invariant characterisation studied in this paper is based on the local geometry of the points on the shape. The advantage of this characterisation is that

it significantly reduces the uncertainty associated with the use of higher level primitives such as lines or curves.

We have considered parametric models defined by similarity and affine transformations. For these transformations there exists an established theoretical development for the generation of invariant features. We capitalise on these concepts to develop several alternative arrangements of points that can be used to characterise a shape for invariant evidence gathering. Invariance reduces the complexity of the HT, and it is possible to compute the location of a shape using a 2D accumulator space independent of the complexity of the shape and the generality of the transformation. However, the complexity of determining corresponding arrangements of points in the model and in the shape is directly related to the generality of the transformation.

For affine transformations, we have included characteristic points defined as points of high curvature to reduce the potential solutions in the local matching process that identifies image points and model points. An effective strategy that reduces the number of correspondences between the model and the image is indispensable since the generality of the transformation increases the geometric complexity of the features identified in the shape. The inclusion of features that reduce the number of potential correspondences between local data on the image and the model, can include other local curve descriptors or features such a colour or texture. We have shown that the technique can obtain adequate results when some features are generated by background objects or are missed. Thus, the presented approach can provide an adequate way of combining the geometry of the model with other features in the model.

Appendix A. Invariance under similarity transformations

By considering the definitions in Eq. (10), Eq. (11) can be developed as,

$$Q_{sim}(\mathbf{w}(s_0, \mathbf{a}), \mathbf{w}(s_1, \mathbf{a}), \mathbf{w}(s_2, \mathbf{a})) = \frac{l^2((R_x(s_2, \rho) - R_x(s_0, \rho))(R_y(s_1, \rho) - R_y(s_0, \rho)) - (R_x(s_1, \rho) - R_x(s_0, \rho))(R_y(s_2, \rho) - R_y(s_0, \rho)))}{l^2((R_x(s_2, \rho) - R_x(s_0, \rho))(R_x(s_1, \rho) - R_x(s_0, \rho)) + (R_y(s_1, \rho) - R_y(s_0, \rho))(R_y(s_2, \rho) - R_y(s_0, \rho)))}. \quad (\text{A.1})$$

for $R_x(s, \rho) = v_x \cos(\rho) - v_y \sin(\rho)$ and $R_y(s, \rho) = v_x \sin(\rho) + v_y \cos(\rho)$.

This means that the function $Q_{sim}(\mathbf{w}(s_0, \mathbf{a}), \mathbf{w}(s_1, \mathbf{a}), \mathbf{w}(s_2, \mathbf{a}))$ is invariant to scale. We can show that Eq. (A.1) is invariant to rotation by considering the terms in the division separately. The numerator can be developed as

$$(v_{x_2} \cos(\rho) - v_{y_2} \sin(\rho))(v_{x_1} \sin(\rho) + v_{y_1} \cos(\rho)) - (v_{x_1} \cos(\rho) - v_{y_1} \sin(\rho))(v_{x_2} \sin(\rho) + v_{y_2} \cos(\rho)) \quad (\text{A.2})$$

where

$$v_{x_1} = (v_x(s_1) - v_x(s_0)), \quad v_{y_1} = (v_y(s_1) - v_y(s_0)), \\ v_{x_2} = (v_x(s_2) - v_x(s_0)), \quad v_{y_2} = (v_y(s_2) - v_y(s_0)).$$

By algebraic manipulation, Eq. (A.2) can be simplified to,

$$v_{x_2} v_{y_1} - v_{y_2} v_{x_1}, \quad (\text{A.3})$$

which is independent of rotation.

In an analogous way the denominator can be written as

$$(v_{x_2} \cos(\rho) - v_{y_2} \sin(\rho))(v_{x_1} \cos(\rho) - v_{y_1} \sin(\rho)) + (v_{x_1} \sin(\rho) + v_{y_1} \cos(\rho))(v_{x_2} \sin(\rho) + v_{y_2} \cos(\rho)), \quad (\text{A.4})$$

which can be reduced to,

$$v_{x_2} v_{x_1} + v_{y_2} v_{y_1}. \quad (\text{A.5})$$

Consequently,

$$Q_{sim}(\mathbf{w}(s_0, \mathbf{a}), \mathbf{w}(s_1, \mathbf{a}), \mathbf{w}(s_2, \mathbf{a})) = \frac{v_{x_2} v_{y_1} - v_{y_2} v_{x_1}}{v_{x_2} v_{x_1} + v_{y_2} v_{y_1}}, \quad (\text{A.6})$$

is independent of rotation and scale.

Appendix B. Invariance under affine transformations

Invariance under affine transformations implies that the slope of the lines through the points $\mathbf{v}(s_0)$ and $\mathbf{v}(s_1)$ and a parallel line through the points $\mathbf{v}(s_2)$ and $\mathbf{v}(s_3)$ is the same after transformation. Since these lines are parallel, then

$$m_{s_0s_1} = \frac{v_y(s_1) - v_y(s_0)}{v_x(s_1) - v_x(s_0)} = m_{s_2s_3} = \frac{v_y(s_3) - v_y(s_2)}{v_x(s_3) - v_x(s_2)}. \quad (\text{B.1})$$

After an affine transformation the lines are defined by the points $\mathbf{w}(s_0, \mathbf{a})$ and $\mathbf{w}(s_1, \mathbf{a})$, and by the points $\mathbf{w}(s_2, \mathbf{a})$ and $\mathbf{w}(s_3, \mathbf{a})$, respectively.

The slope of these lines is given by

$$m'_{s_0s_1} = \frac{C(v_y(s_1) - v_y(s_0)) + D(v_y(s_1) - v_y(s_0))}{A(v_x(s_1) - v_x(s_0)) + B(v_x(s_1) - v_x(s_0))}, \\ m'_{s_2s_3} = \frac{C(v_y(s_3) - v_y(s_2)) + D(v_y(s_3) - v_y(s_2))}{A(v_x(s_3) - v_x(s_2)) + B(v_x(s_3) - v_x(s_2))}, \quad (\text{B.2})$$

or

$$m'_{ss_1} = \frac{C + Dm_{ss_1}}{A + Bm_{ss_1}}, \quad m'_{s_2s_3} = \frac{C + Dm_{s_2s_3}}{A + Bm_{s_2s_3}}, \quad (\text{B.3})$$

and since $m_{ss_1} = m_{s_2s_3}$, then the lines after the transformation are parallel. That is, $m'_{ss_1} = m'_{s_2s_3}$.

Invariance of the distance ratio under affine transformations is defined by $d_0/d_1 = d_2/d_3$, where d_0 is the distance between the points $v(s_0)$ and $v(s_1)$, d_1 is the distance between the points $v(s_2)$ and $v(s_3)$, d_2 is the distance between the points $w(s_0, \mathbf{a})$ and $w(s_1, \mathbf{a})$, and d_3 is the distance between the points $w(s_2, \mathbf{a})$ and $w(s_3, \mathbf{a})$. We can express the first distance ratio as

$$\frac{d_0}{d_1} = \sqrt{\frac{(v_x(s_1) - v_x(s_0))^2 + (v_y(s_1) - v_y(s_0))^2}{(v_x(s_3) - v_x(s_2))^2 + (v_y(s_3) - v_y(s_2))^2}}. \quad (\text{B.4})$$

By considering the definition in Eq. (B.1),

$$\frac{d_0}{d_1} = \frac{v_x(s_1) - v_x(s_0)}{v_x(s_3) - v_x(s_2)} \quad \text{or} \quad \frac{d_0}{d_1} = \frac{v_y(s_1) - v_y(s_0)}{v_y(s_3) - v_y(s_2)}. \quad (\text{B.5})$$

Analogously,

$$\frac{d_2}{d_3} = \frac{w_x(s_1, \mathbf{a}) - w_x(s_0, \mathbf{a})}{w_x(s_3, \mathbf{a}) - w_x(s_2, \mathbf{a})} \quad \text{or} \quad \frac{d_2}{d_3} = \frac{w_y(s_1, \mathbf{a}) - w_y(s_0, \mathbf{a})}{w_y(s_3, \mathbf{a}) - w_y(s_2, \mathbf{a})}. \quad (\text{B.6})$$

That is,

$$\frac{d_2}{d_3} = \frac{A(v_x(s_1) - v_x(s_0)) + B(v_y(s_1) - v_y(s_0))}{A(v_x(s_3) - v_x(s_2)) + B(v_y(s_3) - v_y(s_2))} \quad (\text{B.7})$$

and according to the definition in Eq. (B.1),

$$\frac{d_2}{d_3} = \frac{v_x(s_1) - v_x(s_0)}{v_x(s_3) - v_x(s_2)}. \quad (\text{B.8})$$

Consequently, the distance ratio d_0/d_1 and d_2/d_3 have the same value.

References

- [1] T.O. Binford, Survey of model-based image analysis systems, *Int. J. Robot. Res.* 1 (1982) 18–64.
- [2] H. Baird, *Model-based Image Matching Using Location*, MIT Press, Cambridge MA, 1984.
- [3] R.T. Chin, C.R. Dyer, Model-based recognition in robot vision, *ACM Comput. Surveys* 18 (1986) 67–108.
- [4] O. Veblen, J.W. Young, *Projective Geometry*, Finn and Company, Boston, USA, 1910.
- [5] J.W. Young, *Projective Geometry*, The Open Court, 1930.
- [6] G. Stockman, S. Kopstein, S. Benett, Matching images to models for registration and object detection via clustering, *IEEE Trans. Pattern Anal. Mach. Intell.* 4 (3) (1982) 229–241.
- [7] G. Stockman, Object recognition and localization via pose clustering, *Comput. Vision Graphics Image Process.* 40 (1987) 361–387.
- [8] W. Perkins, A model based vision system for industrial parts, *IEEE Trans. Comput.* C-17 (1978) 126–143.
- [9] Y. Lamdan, J. Schawatz, H. Wolfon, Object recognition by affine invariant matching, *Proc. IEEE Conf. Comput. Vision Pattern Recognition* 1988, pp. 335–344.
- [10] D.H. Ballard, D. Sabbah, Viewer independent shape recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 6 (1983) 653–659.
- [11] S. Umeyama, Parameterised point pattern matching and its application to recognition of object families, *IEEE Trans. Pattern Anal. Mach. Intell.* 15 (1) (1993) 136–144.
- [12] A. Califano, R. Mohan, Multidimensional indexing for recognition of visual shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (1994) 373–392.
- [13] W.E.L. Grimson, Recognition of object families using parameterized models, in: *Proceedings of the International Conference on Computer Vision*, 1987, pp. 93–101.
- [14] P.J. Rosseeuw, A.M. Leroy, *Robust Regression and Outlier Detection*, Wiley, USA, 1987.
- [15] W.E.L. Grimson, D.P. Huttenlocher, On the sensitivity of the Hough transform for object recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (1990) 255–274.
- [16] P.V.C. Hough, Method and means for recognising complex patterns, US Patent No. 396954 (1962).
- [17] R.O. Duda, P.E. Hart, Use of the Hough transform to detect lines and curves in pictures, *Commun. ACM* 15 (1972) 11–15.
- [18] J. Illingworth, J. Kittler, A survey of the Hough transform, *Comput. Vision, Graphics Image Process.* 48 (1988) 87–116.
- [19] V.F. Leavers, Which Hough transform?, *CVGIP: Image Understanding* 58 (1993) 250–264.
- [20] G. Roth, M.D. Levine, Extracting geometric primitives, *CVGIP: Image Understanding* 58 (1993) 1–22.
- [21] F. Bookstein, Fitting conic sections to scattered data, *CVGIP: Image Understanding* 9 (1979) 56–91.
- [22] D. Fofsyth, J.L. Mundy, A. Zisserman, C.M. Brown, Projectively invariant representations using implicit algebraic curves, *Lecture Notes in Computer Science*, Vol. 427, Springer, Berlin, 1990, pp. 427–436.
- [23] G. Danuser, M. Stricker, Parametric model fitting: from Inlier characterization to outlier detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (2) (1998) 263–280.
- [24] J. Sklansky, On the Hough technique for curve detection, *IEEE Trans. Comput.* 27 (1978) 923–926.
- [25] G.C. Stockman, A.K. Agrawala, Equivalence of Hough curve detection to template matching, *Commun. ACM* 20 (1977) 820–822.
- [26] A.S. Aguado, M.E. Montiel, M.S. Nixon, Improving parameter space decomposition for the generalized Hough transform, *Proceedings of the ICIP'96 International Conference on Image Processing*, Lausanne, Switzerland, 1996, pp. 627–630.
- [27] A.S. Aguado, M.E. Montiel M.S. Nixon, On using directional information for parameter space decomposition in ellipse detection, *Pattern Recognition* 29 (3) (1996) 369–381.

- [28] H.K. Yuen, J. Illingworth, J. Kittler, Detecting partially occluded ellipses using the Hough transform, *Image Vision Comput.* 7 (1) (1989) 31–37.
- [29] J.Y. Chang, A.J. Hanson, Virtual line segment-based Hough transform, *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, 1 (1994) 57–62.
- [30] P.-K. Ser, W.-C. Siu, Object recognition with a 2-D Hough domain, *IEEE International Symposium of Circuits and Systems* 13 (1994) 9–12.
- [31] P.-K. Ser, W.-C. Siu, A new generalized Hough transform for the detection of irregular objects, *J. Visual Comm. Image Representation* 6 (3) (1995) 256–264.
- [32] T.E. Dufresne, A.P. Dhawan, Chord-tangent transformation for object recognition, *Pattern Recognition* 28 (9) (1995) 1321–1331.
- [33] F.P. Kuhl, C.R. Giardina, Elliptic Fourier features of a closed contour, *Comput. Vision, Graphics Image Process.* 18 (1982) 236–258.
- [34] A.S. Aguado, M.S. Nixon, M.E. Montiel, Parameterizing arbitrary shapes via Fourier descriptors for evidence-gathering extraction, *Comput. Vision Image Understanding* 69 (2) (1998) 202–221.
- [35] M. Atiquzzaman, Multiresolution Hough transform—An efficient method of detecting patterns in images, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (11) (1992) 1090–1095.

About the Author—ALBERTO S. AGUADO is currently a Lecturer in the Department of Electrical Engineering at the University of Surrey. Previously, he was a research fellow at INRIA Rhône-Alpes (1997–1999). He worked as a research fellow within the project Fuzzy Land Information from Environmental Remote Sensing (FLIERS) in the Image, Speech and Intelligent Systems (ISIS) Research Group in the Department of Electronics and Computer Science at the University of Southampton. In 1997, he completed his Ph. D. at the University of Southampton at the Department of Electronics and Computer Science.

About the Author—EUGENIA MONTIEL completed her Ph.D. in 1997 at the University of Southampton in the Department of Electronics and Computer Science. Her research dissertation was on recursive definition of curves. She holds a B.Eng. and a M.Sc. in Computer Science. She was a research fellow at iMAGIS at INRIA Rhône-Alpes (1998–1999). Currently, she is a research fellow at the CCSR group at the University of Surrey. She has authored several papers in shape extraction, mathematical morphology and fractals.

About the Author—MARK S. NIXON is a Reader in the Image Speech and Intelligent Systems research group, and heads the vision research team there. His team's research includes feature extraction and description, applied to medical imagery and in biometrics (particularly gait recognition). He chaired the British Machine Vision Conference BMVC98, held at the University of Southampton, and has just completed a textbook with Alberto, forthcoming with Butterworth Heinmann on feature extraction in computer vision.