
ELSEVIER

Pattern Recognition Letters 22 (2001) 959–969

Pattern Recognition Letters

www.elsevier.nl/locate/patrec

Improving the Hough Transform gathering process for affine transformations

Eugenia Montiel [a], Alberto S. Aguado [b,*], Mark S. Nixon [c]

[a] *iMAGIS, INRIA Rhône-Alpes, ZIRST – 655 avenue de l'Europe, 38330 Montbonnot Saint Martin, France*
[b] *Department of Electronic and Electrical Engineering, University of Surrey, Guildford, Surrey GU2 7XH, UK*
[c] *Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK*

Received 27 April 2000; received in revised form 23 January 2001

Abstract

In this paper, we show that significant wrong evidence can be generated when the Hough Transform (HT) is used to extract arbitrary shapes under rigid transformations. In order to reduce the amount of wrong evidence, we consider two types of constraints. First, we define constraints by considering invariant features. Secondly, we consider constraints defined via gradient direction information. Our results show that these constraints can significantly improve the gathering strategy, leading to identification of the correct parameters. The presented formulation is valid for any rigid transformations represented by affine mappings. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Shape extraction; Hough Transform; Shape discrimination; Object recognition; Geometric invariance; Geometric constraints; Affine

1. Introduction

The problem of model shape extraction can be viewed as an estimation problem in which the free parameters of an equation are determined such that they define a curve that corresponds to an image primitive (Roth and Levine, 1993). The Hough Transform (HT) is a well-established technique for shape extraction (Stockman and Agarwala, 1977; Sklansky, 1978; Illingworth and Kittler, 1993; Leavers, 1993). The technique gathers evidence for the potential parameters of the equation that defines a shape, by mapping image points into the space defined by the parameters of the curve. After gathering the evidence of all the image points, shapes are determined by local maxima in the parameter space (i.e., local peaks).

The HT was first applied to extract straight lines and was then extended to the extraction of quadratic curves such as circles and ellipses (Illingworth and Kittler, 1993; Leavers, 1993). In a broader definition, the HT can be generalised to the extraction of arbitrary models by changing the equation of the curve under detection (Sklansky, 1978). An arbitrary model can be defined by combining the equation of a shape with a parameterised transformation (Aguado et al., 1998). Thus, the parameters of the model are actually the

* Corresponding author. Tel.: +44-1483-876044; fax: +44-1483-534139.
E-mail address: A.Aguado@eim.surrey.ac.uk (A.S. Aguado).

PII: S0167-8655(01)00039-3

parameters of the transformation that represents the different appearances of a shape in an image. This generalisation has been mainly defined for similarity transformations (Ballard, 1981; Ser and Siu, 1995; Yip et al., 1995), although recently there has been an interest in affine transformations and more general transformations (Aguado et al., 1997; Dufresne and Dhawan, 1995; Lo and Tsai, 1997; Yuen and Ma, 1997).

Geometric hashing and cluster techniques (Grimson and Huttenlocher, 1990; Aguado et al., 2000; Grimson and Huttenlocher, 1991; Bhandarkar and Su, 1991; Chakravarthy and Kasturi, 1991; Lamdan and Wolfson, 1991) are closely related to the HT. Both approaches gather evidence of a shape. However, whilst geometric hashing uses primitives such as lines, or curves, the HT gathers evidence by considering the duality of edge points. This makes the two techniques fundamentally different. As consequence of their definition, geometric hashing and clustering techniques require less computational resources than the HT. However, these techniques can suffer when primitives are not accurately computed (Grimson and Huttenlocher, 1990; Aguado et al., 2000). A significant amount of work has been aimed to analyse the performance and to present alternative computation strategies in geometric hashing (e.g., Grimson and Huttenlocher, 1991; Bhandarkar and Su, 1991; Chakravarthy and Kasturi, 1991; Lamdan and Wolfson, 1991). Conversely, less interest has been given to the HT. Some works have shown that the HT can actually be generalised to extract arbitrary shapes under rigid transformations. However, there is little analysis about the process. In general, analyses on geometric hashing cannot be applied to the HT (Aguado et al., 2000).

This paper considers the application of the HT to extract shapes under affine transformations. The paper has two main contributions. First, we show that the performance of the HT can be reduced due to the generality of the transformation. Intuitively as the transformation, that defines the appearance of a model shape, becomes more general then local image information loses significance. That is, the model increases the number of possible forms to be matched. Therefore, the model can easily be matched to noise or to segments of objects that do not correspond to the description of the whole shape. The false evidence thus generated can interfere seriously with the detection process causing inaccurate, or even incorrect, results. As a second contribution, we consider whether the inclusion of geometric constraints can improve the gathering process. We consider two types of constraints. First, we define constraints based on invariant features. Secondly, we consider constraints defined via gradient direction information. Gradient constraints have been previously used for the extraction of lines, quadratic forms (Aguado et al., 1996) and for shapes under similarity transformations (Ballard, 1981). Here we show that these constraints are also very important for shapes under affine mappings. Our results show that invariant and gradient direction constraints can significantly improve the gathering strategy of the HT.

This paper is organised as follows. Section 2 introduces the notation used in this paper. For completeness, Section 3 presents the definition of the HT for arbitrary shapes. Arbitrary shapes are parameterised by a continuous curve under rigid transformations represented by affine mappings. Our formulation is valid for general geometric transformations, however, our examples and results are developed for affine transformations only. In Section 4 we discuss the source of wrong evidence in the HT. Notice that this is different to the combinatorial error discussed in (Grimson and Huttenlocher, 1990). In the HT wrong evidence is generated when the model is matched against false evidence independently of the accuracy in the data. In Section 5 we consider constraints to reduce false evidence during the gathering process. Section 7 presents implementations and examples. Section 6 includes conclusions.

2. Notation

In this paper, we consider an image and a model shape as a collection of points in the two-dimensional Euclidean space. A collection of points in the image is denoted as uppercase bold letters and a collection of points in the model is denoted as uppercase Greek characters. An image point is

denoted as a lowercase bold letter and a point in the model as a lowercase character. Several points are distinguished by sub-indices. Thus, $\boldsymbol{P} = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$ denotes a collection of three image points and $\Gamma = \{\upsilon_1, \upsilon_2, \upsilon_3\}$ denotes a collection of three model points. The co-ordinates of each point are indicated by using the sub-index x and y, thus, $\boldsymbol{p}_1 = (\boldsymbol{p}_{1x}, \boldsymbol{p}_{1y})$. A shape is defined by the curve of an arbitrary function and is denoted as $\upsilon(\ )$. A parametric model is the curve of a function with two arguments and is denoted as $\boldsymbol{w}(\boldsymbol{a}, \boldsymbol{b})$. Here, we divide the parameters into translation parameters $\boldsymbol{b}$ and what we call deformation parameters $\boldsymbol{a}$. A numerical value obtained based on the geometric properties of a single point or a collection of points is called a feature or a measure and is represented as a function. Thus, for example, $\boldsymbol{G}(\boldsymbol{p}_i)$ is the gradient direction at a point $\boldsymbol{p}_i$.

3. Arbitrary shapes under affine mappings

3.1. General HT

The HT gathers evidence of a model shape through a mapping defined between the image space and the parameter space. In this section, we are interested in obtaining a formal representation of this mapping when the model is given by an arbitrary shape under an affine mapping.

The HT can be defined for arbitrary shapes by considering two components in the definition of a model. First, we can consider a shape represented by an equation without any free parameter. Secondly, the model can be obtained by applying a parameterised transformation that defines the potential appearances of the shape. Thus, the parameters of the transformation become the parameters of the model. In order to exemplify these concepts we can consider a circle with radius unity and centred on the origin to be a shape without any parameters. If we scale the circle (in the x and y directions), rotate it and translate it, then we obtain a model defined by an ellipse. Each of the five parameters of the ellipse is related to a parameter of the transformation. Thus, if we want to find an ellipse in an image, we need to find the parameters of the transformation that map a circle into an ellipse. This idea can be generalised by replacing the circle for another curve that defines the shape of an arbitrary object. As such, we will aim at finding the parameters that map the shape of the object into the shape in an image. It is important to notice that in this approach the complexity of the extraction process is independent of the complexity of the shape. That is, if we replace the circle by a complex shape such as the profile of a mountain or the course of a river, then the model governing its appearance will still have the same number of free parameters. Therefore, the HT mapping for the circle and the complex shapes both have a five-dimensional accumulator space and the extraction process involves the same complexity.

If we consider an arbitrary shape given by the curve $\upsilon(\ )$, then a parametric model is composed of the points

$$\boldsymbol{w}(\boldsymbol{a}, \boldsymbol{b}, \upsilon) = \boldsymbol{f}(\boldsymbol{a}, \upsilon) + \boldsymbol{b} \tag{1}$$

for $\upsilon \in \upsilon(s)$ a point in the shape. Here the function $\boldsymbol{f}(\boldsymbol{a}, \upsilon)$ represents a transformation that defines a new curve by mapping the points of $\upsilon(\ )$ according to the deformation parameters $\boldsymbol{a}$. The vector $\boldsymbol{b}$ represents a translation.

If we consider an image point $\boldsymbol{p}$, then we can match this point to a point in the model. This would imply that $\boldsymbol{p} = \boldsymbol{w}(\boldsymbol{a}, \boldsymbol{b}, \upsilon)$. Thus, by solving for $\boldsymbol{b}$ in Eq. (1), we have that

$$\boldsymbol{b}(\boldsymbol{p}, \upsilon, \boldsymbol{a}) = \boldsymbol{p} - \boldsymbol{f}(\boldsymbol{a}, \upsilon). \tag{2}$$

Here, the function $\boldsymbol{b}(\boldsymbol{p}, \upsilon, \boldsymbol{a})$ represents a mapping that obtains the location parameters for each potential value of the parameters $\boldsymbol{a}$, given an image point $\boldsymbol{p}$ and a model point υ. That is, it defines the point spread function (psf). This function traces a curve in the parameter space. The HT defines the parameter space by an accumulator. Thus, evidence is gathered by increasing the elements defined by the psf. After all evidence has been gathered, then maxima in the parameter space define the best values $\boldsymbol{a}^*$ and $\boldsymbol{b}^*$ that represent the transformation that maps the model into the image.

3.2. Affine transformations

The mapping in Eq. (2) can be defined for several transformations $\boldsymbol{f}(\boldsymbol{a},\upsilon)$. For affine or linear transformations, we have that

$$\boldsymbol{f}(\boldsymbol{a},\upsilon)=\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}\begin{bmatrix}\upsilon_x\\ \upsilon_y\end{bmatrix}.$$

This transformation includes other transformations such as scale, rotation and shear and it is very useful to approximate the appearance of object under rigid motions. Let us suppose that instead of considering one single point $\boldsymbol{p}$ in the image and a single point υ in the model, we have a collection of points $\boldsymbol{P}$ in the image and a collection of points Γ in the model. Thus, we have that $\boldsymbol{P}=\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$ and $\Gamma=\{\upsilon_1,\upsilon_2,\upsilon_3\}$. By considering Eq. (2), we can obtain the simultaneous equations:

$$\begin{aligned}\boldsymbol{p}_1-\boldsymbol{p}_2&=\boldsymbol{f}(\boldsymbol{a},\upsilon_1)-\boldsymbol{f}(\boldsymbol{a},\upsilon_2),\\ \boldsymbol{p}_1-\boldsymbol{p}_3&=\boldsymbol{f}(\boldsymbol{a},\upsilon_1)-\boldsymbol{f}(\boldsymbol{a},\upsilon_3)\end{aligned}\tag{3}$$

which can be developed as two independent equations by considering the orthogonal components of each point. That is,

$$\begin{aligned}\boldsymbol{V}^{-1}\begin{bmatrix}\boldsymbol{p}_{1x}-\boldsymbol{p}_{2x}\\ \boldsymbol{p}_{1x}-\boldsymbol{p}_{3x}\end{bmatrix}&=\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix},\\ \boldsymbol{V}^{-1}\begin{bmatrix}\boldsymbol{p}_{1y}-\boldsymbol{p}_{2y}\\ \boldsymbol{p}_{1y}-\boldsymbol{p}_{3y}\end{bmatrix}&=\begin{bmatrix}\boldsymbol{C}\\ \boldsymbol{D}\end{bmatrix}\end{aligned}\tag{4}$$

for

$$\boldsymbol{V}=\begin{bmatrix}\upsilon_{1x}-\upsilon_{2x} & \upsilon_{1y}-\upsilon_{2y}\\ \upsilon_{1x}-\upsilon_{3x} & \upsilon_{1y}-\upsilon_{3y}\end{bmatrix}.$$

In general, we can redefine Eq. (2) as the pair of equations

$$\begin{aligned}\boldsymbol{b}(\boldsymbol{P},\Gamma)&=\boldsymbol{p}-\boldsymbol{f}(\boldsymbol{S}(\boldsymbol{P},\Gamma),\upsilon),\\ \boldsymbol{a}(\boldsymbol{P},\Gamma)&=\boldsymbol{S}(\boldsymbol{P},\Gamma)\end{aligned}\tag{5}$$

for $\boldsymbol{p}\in\boldsymbol{P}$ and $\upsilon\in\Gamma$. Here the function $\boldsymbol{S}(\boldsymbol{P},\Gamma)$ obtains the transformation parameters from the collections $\boldsymbol{P}$ and Γ. That is it is defined by Eq. (4).

Eq. (5) defines a general parameter decomposition of the HT. The first equation defines the location parameters independently of the parameters in $\boldsymbol{a}$. The second equation solves for the deformation parameters in $\boldsymbol{a}$ independently of the location parameters. This approach corresponds to a generalisation of the parameter space decomposition for quadratic forms. The reduction in the dimensions of the parameter space makes Eq. (5) much more amenable for practical implementation than Eq. (2).

4. Wrong evidence

In Eq. (2) we associate a point $\boldsymbol{p}$ in an image to a point υ in the model. In a straightforward implementation, we can consider for each point $\boldsymbol{p}$ all the points υ that form the model. However, this generates wrong evidence since we consider pairs $(\boldsymbol{p},\upsilon)$ that do not give the correct transformation. This wrong evidence is not related to the extension of the psf as in the case of clustering techniques (Califano and Mohan, 1994). Wrong evidence corresponds to psfs that do not define the primitive.

Eq. (2) provides the correct values of $\boldsymbol{b}$ and $\boldsymbol{a}$ only when the values of $\boldsymbol{p}$ and υ are related by Eq. (1). That is, when

$$\boldsymbol{p}=\boldsymbol{f}(\boldsymbol{a}^*,\upsilon)+\boldsymbol{b}^*,\tag{6}$$

where $\boldsymbol{b}^*$ and $\boldsymbol{a}^*$ are the parameters that map the model shape into the image primitive. Accordingly, only one point in the psf generates true evidence, the evidence for the remainder is wrong. This problem is more significant for Eq. (5) since many more pairs $(\boldsymbol{P},\Gamma)$ can be generated from the combinations associated with a collection of points. In this case the correct value of the parameters is obtained only when the points match the transformed model. That is,

$$\boldsymbol{p}_i-\boldsymbol{f}(\boldsymbol{a}^*,\upsilon_i)+\boldsymbol{b}^*=0\quad\forall\boldsymbol{p}_i\in\boldsymbol{P},\ \ \upsilon_i\in\Gamma.\tag{7}$$

If we consider $\boldsymbol{n}$ image points (i.e., $\boldsymbol{P}=\{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n\}$) in Eq. (5), then for $\boldsymbol{m}$ model points there are ${}_mC_n$ combinations of possible pairs $(\boldsymbol{P},\Gamma)$ for each set $\boldsymbol{P}$. One of these pairs gives the correct transformation whilst the others generate wrong evidence. As we shall show in the examples in Section 6, this wrong evidence can easily lead to incorrect results. The obvious solution to this problem is to control the selection of the points in $\boldsymbol{P}$ and Γ.

5. Gathering constraints

5.1. Invariance constraints

In order to reduce the wrong evidence, we can establish a mechanism aimed to select the points in $\boldsymbol{P}$ and Γ that satisfy Eq. (7). As such, we can consider a verification stage, which ensures that only the points that share a measured feature in the model and in the image are used in Eq. (5). If points share a feature, then it is probable that they correspond to the same point in the model and in the shape in the image. Accordingly, it is probable that they satisfy Eq. (7). Previous work has considered the use of intensity or chromatic attributes as a possible characterisation of image points (Grimson and Huttenlocher, 1990). Here we focus on obtaining a geometric characterisation rather than a chromatic constraint.

As a starting point we could define a constraint such that each pair $\boldsymbol{p}$ and υ have the same gradient direction. That is, if $\boldsymbol{G}(\boldsymbol{p}) = \boldsymbol{G}(\upsilon)$ (Ballard, 1981). However, this information cannot characterise points if the transformation includes rotation. Thus, an effective characterisation should be independent of the transformation that dictates a shape's appearance. Thus, we should consider a pair of points $\boldsymbol{p}$ and υ only if

$$\boldsymbol{Q}(\boldsymbol{p}) = \boldsymbol{Q}(\upsilon),$$

where $\boldsymbol{Q}$ is invariant with respect to the transformation and with respect to translation. That is,

$$\boldsymbol{Q}(\boldsymbol{f}(\boldsymbol{a}, \upsilon)) = \boldsymbol{Q}(\upsilon).$$

From Eq. (5), we have that

$$\boldsymbol{Q}(\boldsymbol{P}) = \boldsymbol{Q}(\Gamma),$$

which indicates an invariant correspondence between a collection of points.

An invariance characterisation is not unique. Thus, given a point $\boldsymbol{p}$ or a collection of points $\boldsymbol{P}$, we can identify several points υ or Γ, respectively. Accordingly, if we denote the points in the model characterised by the same invariant feature as $\boldsymbol{W}(\boldsymbol{P})$, then

$$\boldsymbol{W}(\boldsymbol{P}) = \{\Gamma \,|\, \boldsymbol{Q}(\boldsymbol{P}) - \boldsymbol{Q}(\Gamma) = 0,\ \Gamma \subset \{\upsilon\}\}, \tag{8}$$

for all the combinations of points in the model $\{\upsilon\}$. Based on this constraint, we can rewrite Eq. (5) as

$$\begin{aligned} &\boldsymbol{b}(\boldsymbol{P}, \Gamma) = \boldsymbol{p} - \boldsymbol{f}(\boldsymbol{S}(\boldsymbol{P}, \Gamma), \upsilon) \quad \forall \Gamma \in \boldsymbol{W}(\boldsymbol{P}), \\ &\boldsymbol{a}(\boldsymbol{P}, \Gamma) = \boldsymbol{S}(\boldsymbol{P}, \Gamma) \quad \forall \Gamma \in \boldsymbol{W}(\boldsymbol{P}). \end{aligned} \tag{9}$$

These equations indicate that evidence will only be gathered when the invariant feature $\boldsymbol{Q}$ in the model and in the image is the same.

5.2. Gradient direction constraints

The constraints in Eq. (9) reduce false evidence by considering a selection process, which determines whether local geometric information has the same characterisation in the model and in the image. However, this constraint relies on the supposition that the cardinality of $\boldsymbol{W}(\boldsymbol{P})$ is small and it does not consider the false evidence generated by background objects or other scene artifacts.

We can consider a verification process that determines when the transformation defined by the points $\boldsymbol{P}$ and Γ is congruent to image data, in a manner similar to backmapping. This can be formalised by considering that the solution of Eq. (9) defines the parameters of a transformation that maps the points Γ into $\boldsymbol{P}$. That is, we can consider the solution in Eq. (9) in Eq. (7). Thus, we have that

$$\boldsymbol{p}_i = \boldsymbol{f}(\boldsymbol{a}(\boldsymbol{P}, \Gamma), \upsilon_i) + \boldsymbol{b}(\boldsymbol{P}, \Gamma). \tag{10}$$

This means that if we apply the transformation defined by Eq. (9) to the model point υ_i, then we obtain the co-ordinates of the point in the image. Thus, in order to verify the validity of the transformation we can consider whether additional points in the model and not in Γ, are mapped to points in the image. In general, we can expect that for each point in the model we have a point in the image. However, the transformation will not give a perfect match since there might be noise or occlusion. Thus, we would need to consider several points to analyse whether the transformation is congruent with the image information. In order to avoid comparing a large number of points, we only compare features of the points in $\boldsymbol{P}$ with features of the points in Γ after the transformation. That is,

we compute the gradient direction $\boldsymbol{G}(\boldsymbol{p}_i)$ for $\boldsymbol{p}_i \in \boldsymbol{P}$ and compare it against the value of

$$\boldsymbol{G}(f(\boldsymbol{a}(\boldsymbol{P},\Gamma),\upsilon_i)+\boldsymbol{b}(\boldsymbol{P},\Gamma)).$$

Thus, evidence is gathered only if the value is the same for all the points in the collection. That is, we constrain the gathering process in Eq. (10) to values $(\boldsymbol{P},\Gamma)$ for which

$$\boldsymbol{G}(\boldsymbol{p}_i)=\boldsymbol{G}(\boldsymbol{f}(\boldsymbol{a}(\boldsymbol{P},\Gamma),\upsilon_i)+\boldsymbol{b}(\boldsymbol{P},\Gamma)). \tag{11}$$

Thus, the solution of Eq. (9) is considered in the gathering process if the gradient direction of the points in the image is the same that the gradient of the points in the model after transformation. Gradient direction has been previously used in several techniques to reduce the computational requirements of the HT (e.g., Tsuji and Matsumoto, 1978; Yoo and Sethi, 1993; Wu and Wang, 1993). This information cannot be used for general geometric transformations since the characterisation of points must be invariant. Nevertheless, its use in a post-verification process provides important information for shape extraction.

6. Implementation and examples

The gathering process can be performed by considering the invariant properties in Eq. (9) with the verification process in Eq. (11). In our implementation, we defined the invariant feature $\boldsymbol{Q}(\boldsymbol{P})$ as the ratio of length of two parallel line segments. That is, $\boldsymbol{Q}(\boldsymbol{P})=|\boldsymbol{p}_1-\boldsymbol{p}_2|/|\boldsymbol{p}_3-\boldsymbol{p}_4|$. Thus, the gathering process can be implemented in four main stages. First, for each point in the image, we select another three points such that they define two parallel line segments. Then, we search for a collection of points in the model which satisfy Eq. (8). That is, points for which $\boldsymbol{Q}(\boldsymbol{P})=\boldsymbol{Q}(\Gamma)$. Following this, we solve for the parameters of the transformation according to Eq. (5). Finally, we verify the constraint in Eq. (11) and we gather evidence of the parameters in three two-dimensional accumulators: one for the location parameters and the other two for the deformation parameters.

In a straightforward implementation, we need to consider all the potential combinations of four points in the image and then search for the corresponding points in the model. The number of combinations makes the selection of all the points and the search for corresponding points computationally impossible. In order to reduce the number of combinations, we constrain the points $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$ to a subset of points of high curvature. That is, given a point $\boldsymbol{p}_1$, we select the points $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$ from a list of high curvature points. Then, the point $\boldsymbol{p}_2$ is selected by looking along the line that passes through $\boldsymbol{p}_1$ and whose slope is equal to the slope of the line $\boldsymbol{p}_3\boldsymbol{p}_4$. We select points of high curvature in the image and in the model, thus reducing the combinations used to gather evidence and the search for corresponding points. If we locate m points of high curvature in an image, then we have that each point can be associated to ${}_mC_2$ lines $\boldsymbol{p}_3\boldsymbol{p}_4$. Therefore, if an image has n points, then the gathering process is repeated n_mC_2 times. For each of these four-point combinations, we must search for the corresponding four points in the model. If the model has n' points and m' points of high curvature, then we need to search for $(n_mC_2)(n'_{m'}C_2)$ pairs in the model. The number of points of high curvature can be reduced by setting a high threshold value in the selection process. When the number of high curvature points is large, the combinations can be reduced by considering just the lines within a distance from the point. That is, we can select the points $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$ as the high curvature points closest to $\boldsymbol{p}_1$. It is important to say that the reduction in the number of combinations reduces the number of votes, however the invariance constraint and the verification process still can produce a clear peak in the accumulators.

Fig. 1 shows an example of the extraction of the HT for affine shapes. Fig. 1(a) shows the model used in this example. This model is actually defined by a continuous curve with 15 Fourier coefficients (Aguado et al., 1998). The image in Fig. 1(b) contains a primitive that approximates a linear transformation of the model in Fig. 1(a). The accumulators in Fig. 1(c)–(e) were obtained by gathering evidence according to Eq. (9). These accumulators have well-defined peaks which define an accurate value of the parameters of the transformation. The pair of peaks in the accumulators of Fig. 1(d) and (e) show that the matching allows two solutions that correspond to mirror

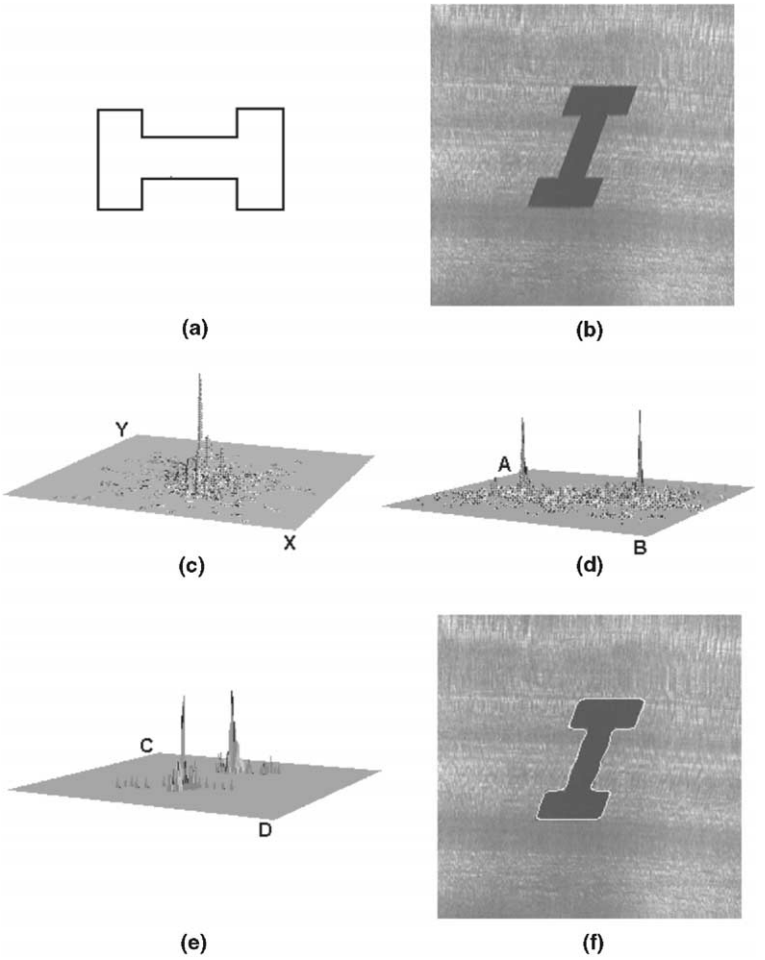

Fig. 1. Example of the extraction process under affine transformations: (a) model shape; (b) raw image; (c) accumulator for the translation parameters; (d) accumulator for the parameters A B; (e) accumulator for the parameters C D; (f) extraction result.

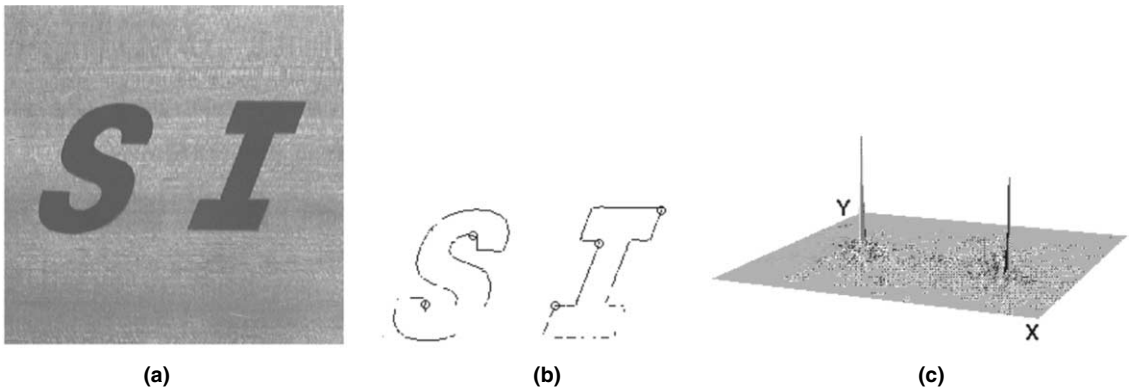

Fig. 2. Example of wrong extraction for two primitives: (a) raw image; (b) edge points; (c) translation parameters accumulator.

transformations. The shape defined by these transformations is shown in Fig. 1(f) superimposed on the original image.

The well-defined peaks in the accumulators of the example in Fig. 1 show that the mapping in Eq. (9) provides an effective approach for gathering evidence of arbitrary shapes under rigid transformations. However, the generality of the transformation can lead to incorrect results. This case is illustrated in the example as shown in Fig. 2. This example was obtained by considering the same gathering process as the one used in the example in Fig. 1. The only difference is that the input image contains an extra object. However, the accumulator for the location parameters in Fig. 2(c), contains two well-defined peaks which suggest that both objects correspond to the model. Furthermore, the largest peak is associated with the letter "S", not the letter "I", so the wrong shape/transformation is actually selected. An analysis of the evidence gathering process shows that the incorrect location is due to the wrong values produced by the lack of discriminatory power of the invariance constraints. That is, points in the incorrect shape are identified with points in the model and wrong evidence is then generated. This can be seen in the example shown in Fig. 3. Fig. 3(a) shows the true primitive in the image whilst Fig. 3(b) shows the object that is wrongly identified with the model. Fig. 3(a) shows four points used to gather evidence. These points are identified with the points in the model shown in Fig. 3(c). The points in Figs. 3(a) and in (c) define the same invariance, then evidence of the true shape is generated. Fig. 3(b) shows four points whose invariance value is the same as that for the points in Fig. 3(c). Thus, the points generate wrong evidence. This evidence defines the model shape superimposed in Fig. 3(b). Actually, the generality of the transformation can produce more mismatching cases than that in Fig. 3(b). This problem can seriously interfere with the detection process and a shape that does not correspond to the model is then located.

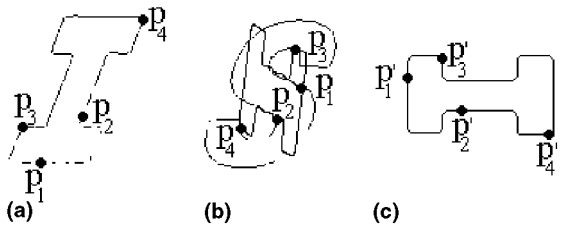

Fig. 3. Example of corresponding points: (a) points in an image primitive; (b) points in a background object; (c) corresponding points in the model shape for (a) and (b).

The lack of discriminatory ability is due to the myopic nature of local analysis. That is, the value of the invariance is aimed to identify equivalence of local shape information when it has suffered a transformation, and not to distinguish when two shapes are different. In consequence, Eq. (9) only defines conditions under which local image information corresponds to a transformed version of local information in the model without any consideration about an object's identity.

The constraint in Eq. (11) eliminates false evidence by considering gradient direction information. According to this equation, evidence is gathered only if the transformation is consistent with the image data. That is, only if the gradient direction computed at the image points is the same as the gradient direction computed at the model shape transformed according to the parameters. For the example in Fig. 3(b) this is equivalent to verifying that the gradient direction at the points $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$, and $\boldsymbol{p}_4$ for the two superimposed shapes is the same. In implementation, only an estimate of gradient direction can be obtained from the information in the image. Consequently, the constraint can only verify that the values of gradient direction in an image are similar to the gradient direction in the transformed model.

The effectiveness of this constraint is illustrated in the example in Fig. 4. The image in Fig. 4(a) contains four objects, two of which are instances of the model shape. Fig. 4(b) shows the edges used in the gathering evidence process, together with the (encircled) points of high curvature. Fig. 4(c) shows the accumulator obtained by using the HT without any constraint. Here we can see that the two of the four objects are incorrectly located. Fig. 4(d) shows the accumulator obtained by constraining the gathering process. The accumulator presents two well-defined peaks that provide an accurate estimate of the position of the

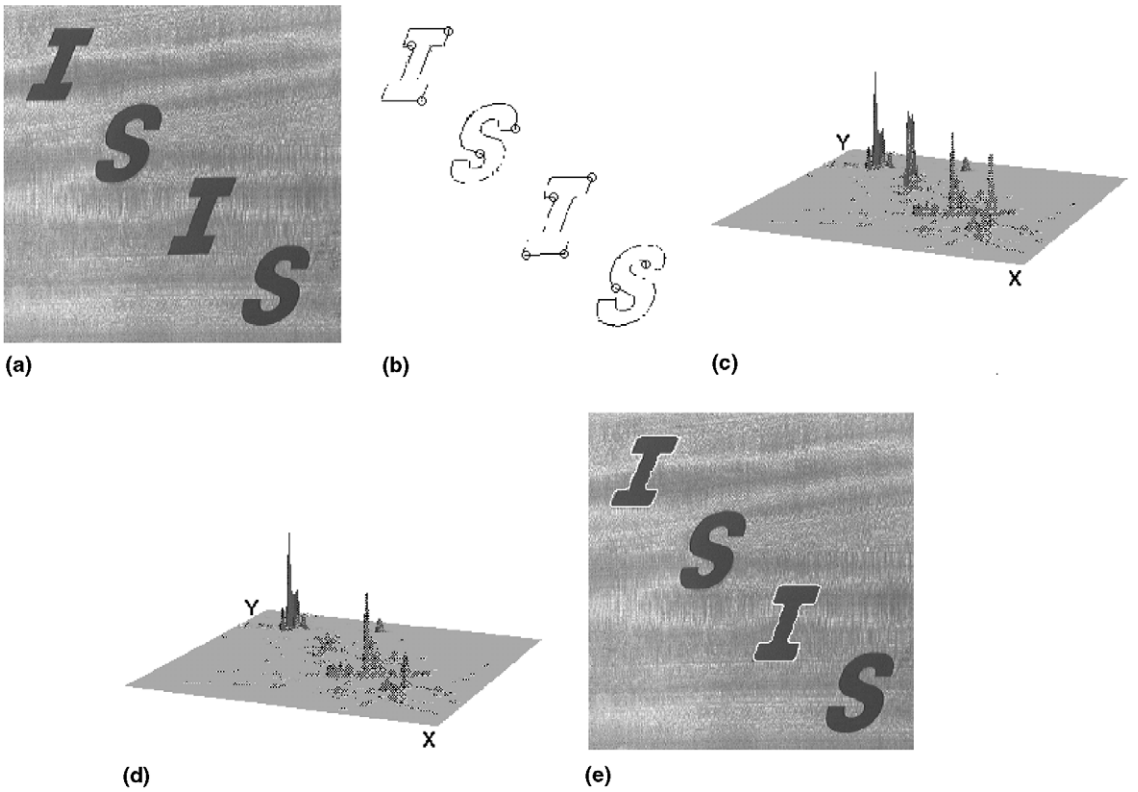

Fig. 4. Example of the extraction process: (a) raw image; (b) edge points; (c) translation parameter accumulator; (d) improved accumulator; (e) extraction result.

instances of the primitive. Fig. 4(e) shows the result of the complete extraction process superimposed on the original image.

Fig. 5 shows another example of the improvement achieved in the gathering process when the constraint in Eq. (11) is used. The edges in Fig. 5(b) were obtained from the image in Fig. 5(a) and used to gather evidence of the model shape in Fig. 5(c). This model was actually defined by a continuous curve with 15 Fourier coefficients. Fig. 5(d) shows the accumulator obtained according to Eq. (9). The accumulator in Fig. 5(e) was obtained by including the constraint in Eq. (11). Comparison of the accumulator arrays accumulator in Fig. 5(d) and (e) shows that prominent wrong peaks are completely eliminated, producing a single well-defined peak. Fig. 5(f) shows the result (superimposed in black) obtained by the extraction technique.

7. Conclusions

The analytic formulation of the HT can be extended to extract arbitrary shapes under general transformations. The generality of this extension increases significantly the amount of false evidence. Geometric invariant features can be included in the formulation as an effective way of reducing the dimensionality of the transformation and to reduce the amount of false evidence gathered. However, the generality of the transformation can still produce an excessive amount of false evidence, potentially leading to incorrect results.

A significant reduction of false evidence can be obtained by considering a verification process that evaluates whether the parameters given by the HT mapping define a transformation that is congruent to other image data. In this paper, we

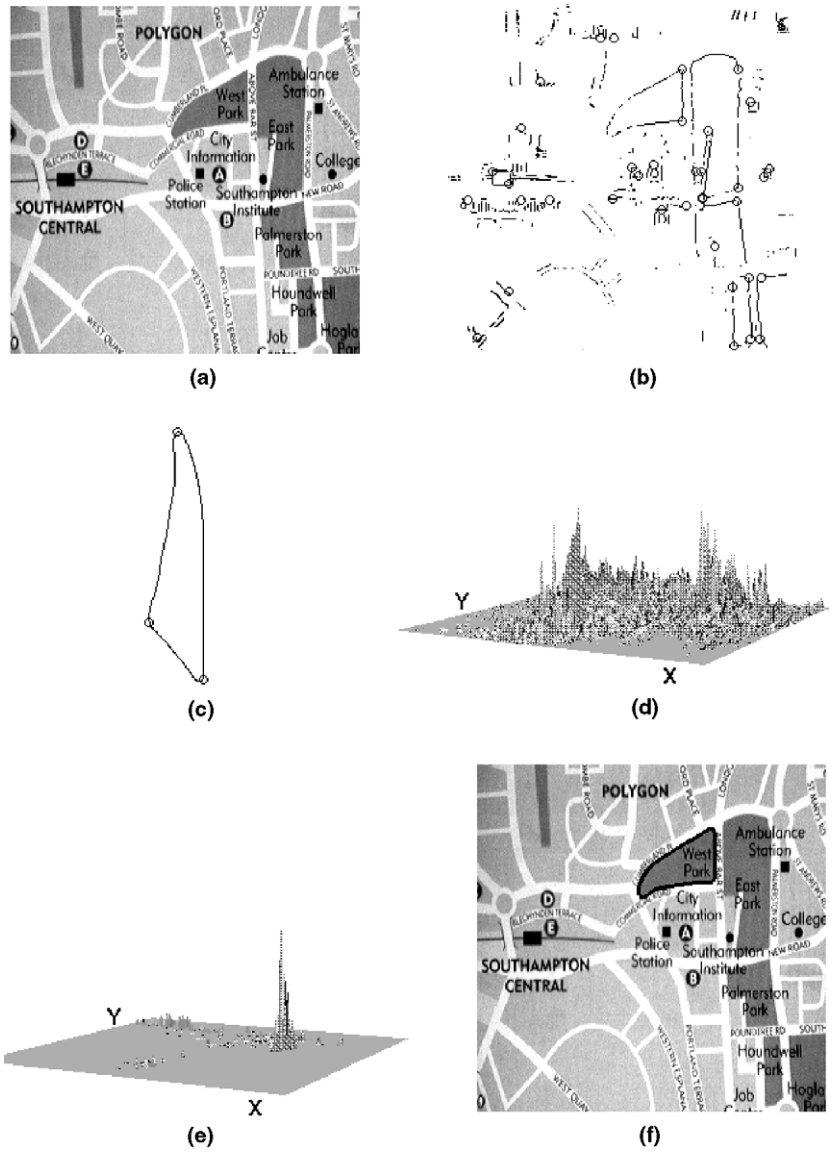

Fig. 5. Example of the extraction process: (a) raw image; (b) edges; (c) model shape; (d) translation parameters accumulator; (e) improved accumulator; (f) extraction result.

have used gradient direction information to reduce the false evidence in affine mappings. Experimental results show that this approach can improve the extraction process significantly leading to correct identification of the appearance parameters even when the original situation would have appeared ambiguous to an unconstrained approach.

References

Aguado, A.S., Nixon, M.S., Montiel, M.E., 1998. Parameterising arbitrary shapes via Fourier descriptors for evidence-gathering extraction. Comput. Vision Image Understanding 69 (2), 202–221.

Aguado, A.S., Montiel, M.E., Nixon, M.S., 1996. On using directional information for parameter space decomposition in ellipse detection. Pattern Recognition 29 (3), 369–381.

Aguado, A.S., Montiel, M.E., Nixon, M.S., 1997. Arbitrary shape Hough transform by invariant geometric features. In: IEEE SMC'97 (Int. Conf. Syst., Man Cybernet.), vol. 3. pp. 2661–2665.

Aguado, A.S., Montiel, M.E., Nixon, M.S., 2000. Bias error analysis of the generalised Hough transform. J. Math. Imaging Vision 12 (1), 25–42.

Ballard, D.H., 1981. Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognition 13, 111–122.

Bhandarkar, S.M., Su, M., 1991. Sensitivity analysis for matching and pose computation using dihedral junctions. Pattern Recognition 24, 505–513.

Califano, A., Mohan, R., 1994. Multidimensional indexing for recognizing visual shapes. IEEE Trans. Pattern Anal. Mach. Intell. 16 (4), 373–392.

Chakravarthy, C.S., Kasturi, R., 1991. Pose clustering on constraints for object recognition. In: CVPR'91. pp. 16–21.

Dufresne, T.E., Dhawan, A.P., 1995. Chord-tangent transformation for object recognition. Pattern Recognition 28 (9), 1321–1331.

Grimson, W.E.L., Huttenlocher, D.P., 1990. On the sensitivity of the Hough transform for object recognition. IEEE Trans. Pattern Anal. Mach. Intell. 12 (3), 255–274.

Grimson, W.E.L., Huttenlocher, D.P., 1991. On the verification of the hypothesized matches in model-based recognition. IEEE Trans. Pattern Anal. Mach. Intell. 13, 1201–1213.

Illingworth, J., Kittler, J., 1993. A survey of the Hough transform. Comput. Vision, Graphics Image Process. 48, 87–116.

Lamdan, Y., Wolfson, H.J., 1991. On the error analysis of 'Geometric Hashing'. In: CVPR'91. pp. 22–27.

Leavers, V.F., 1993. Which Hough transform?. CVGIP: Image Understanding 58 (4), 250–264.

Lo, R.-C., Tsai, W.-H., 1997. Perspective-transformation-invariant generalized Hough transform for perspective planar shape detection and matching. Pattern Recognition 30 (3), 383–396.

Roth, G., Levine, M.D., 1993. Extracting geometric primitives. CVGIP: Image Understanding 58, 1–22.

Ser, P.-K., Siu, W.-C., 1995. A new generalized Hough transform for the detection of irregular objects. J. Visual Commun. Image Representation 6 (3), 256–264.

Sklansky, J., 1978. On the Hough technique for curve detection. IEEE Trans. Comput. 27, 923–926.

Stockman, G.C., Agarwala, A.K., 1977. Equivalence of Hough curve detection to template matching. Commun. ACM 20, 820–822.

Tsuji, S., Matsumoto, F., 1978. Detection of ellipses by a modified Hough transform. IEEE Trans. Comput. 27, 777–781.

Wu, W.-Y., Wang, M.-J.J., 1993. Elliptical object detection by using its geometric properties. Pattern Recognition 26 (10), 1499–1509.

Yip, R.K.K., Tam, R.K.S., Leung, D.N.K., 1995. Modification of the Hough transform for object recognition using a 2-dimensional array. Pattern Recognition 28 (11), 1733–1744.

Yoo, J.H., Sethi, I.K., 1993. An ellipse detection method from the polar and pole definition of conics. Pattern Recognition 26 (2), 307–315.

Yuen, S.Y., Ma, C.H., 1997. An investigation of the nature of parameterization for the Hough transform. Pattern Recognition 30 (6), 1009–1040.