Invariant Characterization of the Hough Transform for Pose Estimation of Arbitrary Shapes

Alberto S. Aguado[1], Eugenia Montiel[2] and Mark S. Nixon[3]
[1]Electronic Engineering, University of Surrey, UK
[2] iMAGIS, INRIA Rhône-Alpes, France
[3]University of Southampton, UK

Abstract

We develop a new formulation for including invariance in a general form of the Hough transform. We first develop a formal definition of the Hough transform mapping for arbitrary shapes and general transformations. We then include an invariant characterization of shapes and we develop and apply our technique to extract shapes under similarity and affine transformations. Our characterization does not require the computation of properties for lines or other primitives that compose a model, but is based solely on the local geometry given by points on shapes. Experimental results show that the new technique is capable of extracting arbitrary shapes under occlusion and when the image contains noise.

1 Introduction

In shape extraction, it is important to be able to handle difference in an object's appearance due to change in camera position [1]. Some techniques, including cluster methods, pose clustering, evidence gathering, geometric hashing and hypothesis accumulation have handle changes in object's appearance by verifying the consistency between structures in the image and in the target shape [2-6]. Our approach circumvents this problem by including invariant properties within the evidence gathering procedure of the Hough transform (HT). Our formulation is focused on similarity and affine transformations and it avoids problems due to the uncertainty in the location or lines and other image primitives [7] by considering only edge points in a shape. We take as our starting point the evidence gathering of Fourier parameterized shapes [8]. The use of parameterized shapes ameliorates difficulties inherent in use of tabular curve descriptions (as in the Generalized Hough transform (GHT)).

2 Evidence gathering of arbitrary shapes

Model shape extraction is fundamentally a problem of analysis of regression [9]. Here, we consider as a fitting curve a parametric model defined by a shape model $\upsilon(\boldsymbol{s})$ and a parametric transformation $\boldsymbol{f_a}$. Thus, a parametric model is defined as $\boldsymbol{w}(s,\boldsymbol{a}) = f(\boldsymbol{a},\upsilon(s))$, where $\boldsymbol{a}$ is a vector that contains the transformation parameters and s defines the points in the model. If $\boldsymbol{f_a}$ does not contain any translation term, the model shape $\boldsymbol{w}(s,\boldsymbol{a})$ is relative to the origin. Therefore, a primitive in an image is represented by a translation of the model. That is,

$$\boldsymbol{z}(s,\boldsymbol{a}) = \boldsymbol{w}(s,\boldsymbol{a}) + \boldsymbol{b} \tag{1}$$

where the point $\boldsymbol{b} = (a_0, b_0)$ defines the position of the primitive in an image. The shape extraction problem is solved by determining the parameters in $\boldsymbol{a}$ and the position $\boldsymbol{b}$ such that they define a curve $\boldsymbol{z}(s,\boldsymbol{a})$ that best fits the image data. Deviations due to the use of minimization of quadratic errors combined with the change in the position of the objects have been discussed in [10,11].

The Hough transform (HT) obtains a robust mode fitting by gathering evidence of all the potential 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 $\boldsymbol{z}(s_0,\boldsymbol{a})$, then the location parameter as a function of λ_0 and $\boldsymbol{a}$ is given by

$$\boldsymbol{b}(\lambda_0,\boldsymbol{a}) = \boldsymbol{b}(\boldsymbol{a}_0,\boldsymbol{b}_0) \tag{2}$$

for $\boldsymbol{b}(\lambda_0,\boldsymbol{a}) = \lambda_0 - \boldsymbol{w}(\boldsymbol{s}_0,\boldsymbol{a})$. That is, the location can be determined given a point in the image and the parameters of the transformation. Each combination of the parameters defines a potential value for the location. That is, the match of a point in the model and a point in the image defines a hyper-surface in the parameter space. This hyper-surface defines the point spread function (psf) of the point λ_0. In the HT, the parameters are computed by increasing the elements of an accumulator space that forms the trace of each psf and then searching for a maximum. The elements that are incremented in the accumulator space are given by the mapping

$$\{(\boldsymbol{b},\boldsymbol{a}) | \boldsymbol{b} = \lambda_0 - \boldsymbol{w}(\boldsymbol{s}_0,\boldsymbol{a}), \boldsymbol{w}(\boldsymbol{s}_0,\boldsymbol{a}) = \boldsymbol{f}(\boldsymbol{a},\upsilon(\boldsymbol{s}_0))\} \forall \upsilon_0 \in \boldsymbol{I}, \forall \upsilon(\boldsymbol{s}_0) \tag{3}$$

This equation defines a general HT mapping for arbitrary shapes and transformations. The remainder of this paper will focus on reducing the computational requirements in this equation and on providing an analytic formulation for models represented by curves under similarity and affine transformations.

3 Generalized invariant HT

Invariance provides a general approach for reducing the computational requirements of the HT. In order to characterize invariance, we define a function Q that computes a feature for a point in a curve. The function Q is invariant with respect to $\boldsymbol{f}_a$. That is, $\boldsymbol{Q}(\boldsymbol{f}(\boldsymbol{a},\upsilon(\boldsymbol{s}_0))) = \boldsymbol{Q}(\upsilon(\boldsymbol{s}_0))$. If Q is invariant under translation, then according to the definitions in the previous section, it is possible to establish the relationship $\boldsymbol{Q}(\boldsymbol{z}(\boldsymbol{s}_0,\boldsymbol{a})) = \boldsymbol{Q}(\boldsymbol{f}(\boldsymbol{a},\upsilon(\boldsymbol{s}_0))) = \boldsymbol{Q}(\upsilon(\boldsymbol{s}_0))$. Thereby, for a point λ_0 there exists a model point $\upsilon(\boldsymbol{s}_0)$ such that,

$$\boldsymbol{Q}(\lambda_0) = \boldsymbol{Q}(\upsilon(\boldsymbol{s}_0)). \tag{4}$$

In order to gather evidence, we can constrain the elements of the accumulator space in eq. (3) by considering only the elements for which eq. (4) holds. To constrain eq. (3) we determine, by eq. (4), the potential points in the curve for a given image point. These points can be represented as,

$$\boldsymbol{W}(\lambda_0) = \{\upsilon(\boldsymbol{s}_j) | \boldsymbol{Q}(\lambda_0) - \boldsymbol{Q}(\upsilon(\boldsymbol{s}_j)) = 0\} \tag{5}$$

Thus, instead of considering all the points of $\upsilon(\boldsymbol{s})$ for each point λ_0 in I (i.e., $\forall \lambda_0 \in \boldsymbol{I}, \forall \upsilon(\boldsymbol{s}_0)$), evidence can be gathered by considering only the points in $\boldsymbol{W}(\lambda_0)$ for each point λ_0 in I (i.e., $\forall \lambda_0 \in \boldsymbol{I}, \forall \upsilon(\boldsymbol{s}_0) \in \boldsymbol{W}(\lambda_0)$). A more significant simplification can be achieved if we consider that the matching process in eq. (5) is an invariant mapping. Thus, we can find the transformation by solving for $\boldsymbol{a}$ in $\boldsymbol{Q}(\lambda_0)=\boldsymbol{Q}(\boldsymbol{f}(\boldsymbol{a},\upsilon(\boldsymbol{s}_0)))$. In general, the solution is multi-valued. That is, for each point $\lambda_0 \in \boldsymbol{I}$, and each point $\upsilon(\boldsymbol{s}_0)\in \boldsymbol{W}(\lambda_0)$ we can determine a different transformation that maps the model point into the image point. We define the collection of solutions as,

$$\boldsymbol{f}^{\Delta}(\lambda_0,\upsilon(\boldsymbol{s}_0))=\{\boldsymbol{a}|\boldsymbol{Q}(\lambda_0)=\boldsymbol{Q}(\boldsymbol{f}(\boldsymbol{a},\upsilon(\boldsymbol{s}_0)))\} \qquad (6)$$

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

$$\{\boldsymbol{b}|\boldsymbol{b}=\lambda_0-\boldsymbol{f}(\boldsymbol{a},\upsilon(\boldsymbol{s}_0)),\boldsymbol{a}\in \boldsymbol{f}^{\Delta}(\lambda_0,\upsilon(\boldsymbol{s}_0))\}\,\forall\lambda_0\in \boldsymbol{I},\forall\upsilon(\boldsymbol{s}_0)\in \boldsymbol{W}(\lambda_0) \qquad (7)$$

Consequently, if we establish an invariant function Q for a family of transformations $\boldsymbol{f}_{\boldsymbol{a}}$, we can characterize 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,

$$\{\boldsymbol{a}|\boldsymbol{a}=\boldsymbol{f}^{\Delta}(\lambda_0,\upsilon(\boldsymbol{s}_0))\}\,\forall\lambda_0\in \boldsymbol{I},\forall\upsilon(\boldsymbol{s}_0)\in \boldsymbol{W}(\lambda_0) \qquad (8)$$

Alternatively, we can exploit the values of the location parameters to solve for some of the parameters in $\boldsymbol{a}$, reducing the computational burden. This approach will be considered in further detail in section 6.4.

4 General algorithm

According to eqs. (7) and (8) the extraction process defined via the invariant form of the HT can be implemented in three steps. First, for each point in the image it is necessary to identify a potential set of points in the model by matching invariant features according to eq. (5). Secondly, the image transformation that maps the point in the image and the point in the model is determined by eq. (6). Finally, evidence of the location position and transformation parameters is gathered by the HT mappings in eqs. (7) and (8). The reminder of this paper will focus on characterizing the function Q and solving for the analytic expression in (5) and (6) when $\boldsymbol{f}_{\boldsymbol{a}}$ is defined by similarity and affine transformations.

5 Similarity transformations

5.1 Parametric model

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

$$\boldsymbol{w}(s,\boldsymbol{a})=w_{\boldsymbol{x}}(s,\boldsymbol{a})\boldsymbol{U}_x+w_{\boldsymbol{y}}(s,\boldsymbol{a})\boldsymbol{U}_y \qquad (9)$$

where $\boldsymbol{U}_x=[1,0]$, $\boldsymbol{U}_y=[0,1]$, $w_x(s,\boldsymbol{a})=lR_x(s,\rho)$, $w_y(s,\boldsymbol{a})=lR_y(s,\rho)$ and $[R_x(s,\rho)\;\; R_y(s,\rho)]$ is the result of multiplying the model vector $[\upsilon_x(s)\;\; \upsilon_y(s)]$ by a rotation matrix.

5.2 Geometric invariance

For similarity transformations, the function Q can be obtained by considering the concept of angle. An angle is defined by three points. Here, the point that is characterized by the invariant is denoted as s_0, while the points used to define the angle are denoted by the sub-indices s_1 and s_2. An invariant characterization of the point $\boldsymbol{w}(s_0,\boldsymbol{a})$ is given by

$$Q_{sim}(w(s_0,\boldsymbol{a}),w(s_1,\boldsymbol{a}),w(s_2,\boldsymbol{a}))=\frac{V_{x_2}V_{y_1}-V_{x_1}V_{y_2}}{V_{x_2}V_{x_1}+V_{y_1}V_{y_2}}, \tag{10}$$

for $V_{x_i}=w_x(s_i,\boldsymbol{a})-w_x(s_0,\boldsymbol{a}), V_{y_i}=w_y(s_i,\boldsymbol{a})-w_y(s_0,\boldsymbol{a})$. Note that although the number of parameters in Q has been increased with respect to the definition presented in sec. 4, the function has the same meaning as in eq. (7). That is, it provides a characterization of a single point in the model (i.e., for $\boldsymbol{w}(s_0,\boldsymbol{a})$).

Although eq. (10) defines a unique invariant for the points in the model, there exist alternative ways in which the points can be chosen. Here we use the geometry shown in figure 1. 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 characterizes a point by only one other point in the image and this point does not require to have 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 [12]. In [13,14] this relationship is exploited to define indexed tables suited to invariant extraction of shapes by the GHT. These tables store the position of the center of a shape as a function of the invariant properties in the pole-polar form.

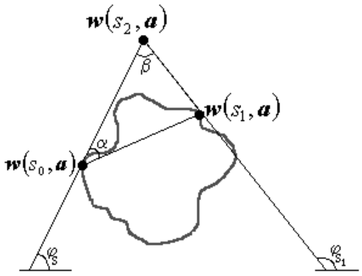

Fig. 1. Arrangements of points defined by a pole-polar relationship.

5.3 Local matching

According to eq. (5), we need to solve for the point that satisfies the relationship $\boldsymbol{Q_{sim}}(\lambda_0,\lambda_1)=\boldsymbol{Q_{sim}}(\upsilon(\boldsymbol{s}_0),\upsilon(\boldsymbol{s}_1))$. The form of each invariance in this relationship can be obtained by simplifying eq. (10) 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. The invariance defined in eq. (10) defines two features in the arrangement shown in Fig. 1(b): (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

$$Q_{sim}(\lambda_0,\lambda_1)=\frac{G(\lambda_0)-\gamma(s_0,s_1)}{1+G(\lambda_0)\gamma(s_0,s_1)},\quad Q_{sim}(\upsilon(s_0),\upsilon(s_1))=\frac{G(\upsilon(s_0))-\varphi(s_0,s_1)}{1+G(\upsilon(s_0))\varphi(s_0,s_1)} \tag{11}$$

for $\gamma(s_0,s_1)=(\lambda_{y_1}-\lambda_{y_1})/(\lambda_{x_1}-\lambda_{x_0})$ and $\phi(s_0,s_1)=(\upsilon_y(s_1)-\upsilon_y(s_0))/(\upsilon_x(s_1)-\upsilon_x(s_0))$. For the second case, as a measure of the angle β the invariant arrangement of points is given by

$$Q'_{sim}(\lambda_0,\lambda_1)=\frac{G(\lambda_0)-G(\lambda_1)}{1+G(\lambda_0)G(\lambda_1)},\quad Q'_{sim}(\upsilon(s_0),\upsilon(s_1))=\frac{G(\upsilon(s_0))-G(\upsilon(s_1))}{1+G(\upsilon(s_0))G(\upsilon(s_1))} \tag{12}$$

From eq. (5) we observe that the problem of obtaining a definition of $\boldsymbol{W}(\lambda_0)$ can be formulated as the problem of obtaining the points $\upsilon(\boldsymbol{s}_0)$ such that the invariant features $\boldsymbol{Q_{sim}}(\lambda_0,\lambda_1)$ and $\boldsymbol{Q'_{sim}}(\lambda_0,\lambda_1)$ computed from an image are equal to the invariant features $\boldsymbol{Q_{sim}}(\upsilon(\boldsymbol{s}_0),\upsilon(\boldsymbol{s}_1))$ and $\boldsymbol{Q'_{sim}}(\upsilon(\boldsymbol{s}_0),\upsilon(\boldsymbol{s}_1))$ in the model. That is, for two points in an image, we have a pair of equations that define a pair of points in the model with the same invariant characterization. More formally,

$$W_{sim}(\lambda_0,\lambda_1)=\{\upsilon(s_0),\upsilon(s_1)|Q_{sim}(\lambda_0,\lambda_1)=Q_{sim}(\upsilon(s_0),\upsilon(s_1)),Q'_{sim}(\lambda_0,\lambda_1)=Q'_{sim}(\upsilon(s_0),\upsilon(s_1))\} \tag{13}$$

The pair of simultaneous equations can be written as,

$$\begin{aligned}\boldsymbol{Q_{sim}}(\lambda_0,\lambda_1)&=(\boldsymbol{G}(\upsilon(\boldsymbol{s}_0))-\phi(\boldsymbol{s}_0,\boldsymbol{s}_1))/(1+\boldsymbol{G}(\upsilon(\boldsymbol{s}_0))\phi(\boldsymbol{s}_0,\boldsymbol{s}_1))\\ \boldsymbol{Q'_{sim}}(\lambda_0,\lambda_1)&=(\boldsymbol{G}(\upsilon(\boldsymbol{s}_0))-\boldsymbol{G}(\upsilon(\boldsymbol{s}_1)))/(1+\boldsymbol{G}(\upsilon(\boldsymbol{s}_0))\boldsymbol{G}(\upsilon(\boldsymbol{s}_1)))\end{aligned} \tag{14}$$

The solution of these pair of equations defines the points in eq. (5).

5.4 Parameters of the transformation

A shape's rotation and scale can be obtained by matching two points of the curve to two points in the model. That is, for each pair of points $\lambda_0,\lambda_1\in\boldsymbol{I}$, and each pair of points $\upsilon(\boldsymbol{s}_0)$ and $\upsilon(\boldsymbol{s}_1)$ such that $\boldsymbol{Q_{sim}}(\lambda_0,\lambda_1)=$ $\boldsymbol{Q_{sim}}(\upsilon(\boldsymbol{s_j}),\upsilon(\boldsymbol{s}_1))$ and $\boldsymbol{Q'_{sim}}(\lambda_0,\lambda_1)=$ $\boldsymbol{Q'_{sim}}(\upsilon(\boldsymbol{s_j}),\upsilon(\boldsymbol{s}_1))$, we can define a function of the form of eq. (6) as,

$$f^{\Delta}(\lambda_0,\lambda_1,\upsilon(s_0),\upsilon(s_1))=\{l,\rho|Q_{sim}(\lambda_0,\lambda_1)=Q_{sim}(f((l,\rho),\upsilon(s_0)),f((l,\rho),\upsilon(s_1)))\} \tag{15}$$

By using the geometric properties of a pair of points, it is possible to obtain an explicit form of eq. (15) as,

$$\rho=tan^{-1}\left(\frac{\lambda_{y_1}-\lambda_{y_0}}{\lambda_{x_1}-\lambda_{x_0}}\right)-tan^{-1}\left(\frac{\upsilon_y(s_1)-\upsilon_y(s_0)}{\upsilon_x(s_1)-\upsilon_x(s_0)}\right)\qquad l=\frac{|\lambda_1-\lambda_0|}{|\upsilon(s_1)-\upsilon(s_0)|} \tag{16}$$

6 Affine transformations

6.1 Parametric model

A parametric model for an affine transformation can be obtained by multiplying a model shape $\upsilon(\boldsymbol{s})$ by a linear transformation. Thus, the parametric model $\boldsymbol{b}(\lambda_0,\boldsymbol{a})=\lambda_0-\boldsymbol{w}(s_0,\boldsymbol{a})$ can be defined by the transformation parameters $\boldsymbol{a}=(A,B,C,D)$ for the orthogonal components of the parametric model in eq. (9) defined as,

$$w_x(s,\boldsymbol{a})=A\upsilon_x(s)+B\upsilon_y(s)\quad w_y(s,\boldsymbol{a})=C\upsilon_x(s)+D\upsilon_y(s). \tag{17}$$

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. The distance ratio between two parallel lines can be defined by four points, thus an invariant characterization of a point $\boldsymbol{w}(s_0,\boldsymbol{a})$ in the parametric model can be obtained by considering three additional points. If the two points $\boldsymbol{w}(s_0,\boldsymbol{a})$ and $\boldsymbol{w}(s_1,\boldsymbol{a})$ define a parallel line to the line formed by the points $\boldsymbol{w}(s_2,\boldsymbol{a})$ and $\boldsymbol{w}(s_3,\boldsymbol{a})$, then the distance ratio invariant for a point $\boldsymbol{w}(s_0,\boldsymbol{a})$ can be defined as,

$$Q_{aff}(\boldsymbol{w}(s_0,\boldsymbol{a}),\boldsymbol{w}(s_1,\boldsymbol{a}),\boldsymbol{w}(s_2,\boldsymbol{a}),\boldsymbol{w}(s_3,\boldsymbol{a}))=\frac{w_x(s_1,\boldsymbol{a})-w_x(s_0,\boldsymbol{a})}{w_x(s_3,\boldsymbol{a})-w_x(s_2,\boldsymbol{a})}=\frac{w_y(s_1,\boldsymbol{a})-w_y(s_0,\boldsymbol{a})}{w_y(s_3,\boldsymbol{a})-w_y(s_2,\boldsymbol{a})} \tag{18}$$

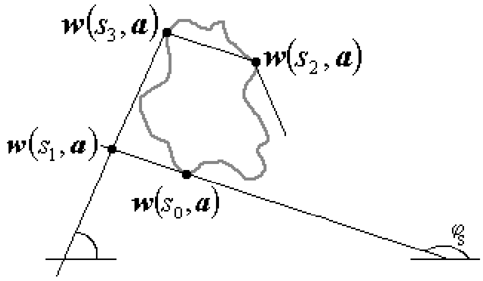

Fig. 2. Arrangement of points to define two parallel lines.

Fig. 2 shows a geometric arrangement that defines invariant features in a manner analogous to Fig. 1. In this case, only three points are on the shape, whilst the fourth one is determined indirectly by using gradient direction information. In this case, the point $\boldsymbol{w}(s_1,\boldsymbol{a})$ is defined by the intersection of the tangent lines to the points $\boldsymbol{w}(s_0,\boldsymbol{a})$ and $\boldsymbol{w}(s_3,\boldsymbol{a})$.

6.3 Local matching

According to eq. (5), the collection of points $\boldsymbol{W}(\lambda_0)$ is given by finding the points $\upsilon(\boldsymbol{s}_0)$ that satisfy the relationship $\boldsymbol{Q_{aff}}(\lambda_0,\lambda_2,\lambda_3)=\ \boldsymbol{Q_{aff}}(\upsilon(\boldsymbol{s}_0),\upsilon(\boldsymbol{s}_2),\upsilon(\boldsymbol{s}_3))$. If we consider the geometry in figure 2, thus this relationship can be written as

$$Q_{aff}(\lambda_0,\lambda_2,\lambda_3)=\frac{\lambda_{x_1}-\lambda_{x_0}}{\lambda_{x_3}-\lambda_{x_2}},\quad Q_{aff}(\upsilon(s_0),\upsilon(s_2),\upsilon(s_3))=\frac{\upsilon_x(s_1)-\upsilon_x(s_0)}{\upsilon_x(s_3)-\upsilon_x(s_2)} \tag{19}$$

where

$$\begin{aligned}\lambda_{x_1}&=\frac{\lambda_{y_0}-\lambda_{y_3}+G(\lambda_3)\lambda_{x_3}+G(\lambda_0)\lambda_{x_0}}{G(\lambda_3)+G(\lambda_0)},\\ \lambda_{y_1}&=\frac{G(\lambda_3)G(\lambda_0)(\lambda_{x_3}-\lambda_{x_0})+G(\lambda_3)\lambda_{y_0}-G(\lambda_0)\lambda_{y_3}}{G(\lambda_3)+G(\lambda_0)}\end{aligned} \tag{20}$$

and where $\upsilon(\boldsymbol{s}_1)$ is defined in a similar fashion. By considering the second part of eq. (18), we obtain the invariance definition

$$Q'_{aff}(\lambda_0,\lambda_2,\lambda_3)=\frac{\lambda_{y_1}-\lambda_{y_0}}{\lambda_{y_3}-\lambda_{y_2}},\quad Q'_{aff}(\upsilon(s_0),\upsilon(s_2),\upsilon(s_3))=\frac{\upsilon_y(s_1)-\upsilon_y(s_0)}{\upsilon_y(s_3)-\upsilon_y(s_2)}. \tag{21}$$

Thus, $\boldsymbol{W}(\lambda_0)$ is given by all the points in the model such that the invariant features $\boldsymbol{Q_{aff}}(\lambda_0,\lambda_2,\lambda_3)$ and $\boldsymbol{Q'_{aff}}(\lambda_0,\lambda_2,\lambda_3)$ computed from an image are equal to the invariant features $\boldsymbol{Q_{aff}}(\upsilon(\boldsymbol{s}_0),\upsilon(\boldsymbol{s}_2),\upsilon(\boldsymbol{s}_3))$ and $\boldsymbol{Q'_{aff}}(\upsilon(\boldsymbol{s}_0),\upsilon(\boldsymbol{s}_2),\upsilon(\boldsymbol{s}_3))$ on the model shape. That is,

$$\begin{aligned}W_{aff}(\lambda_0,\lambda_2,\lambda_3)=\{&\upsilon(s_0),\upsilon(s_2),\upsilon(s_3)\,|\,Q_{aff}(\lambda_0,\lambda_2,\lambda_3)=\\ &Q_{aff}(\upsilon(s_0),\upsilon(s_2),\upsilon(s_3)),Q'_{aff}(\lambda_0,\lambda_2,\lambda_3)=Q'_{aff}(\upsilon(s_0),\upsilon(s_2),\upsilon(s_3))\}\end{aligned} \tag{22}$$

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

$$\begin{aligned}(\boldsymbol{Q_{aff}}(\lambda_0,\lambda_2,\lambda_3))(\upsilon_x(\boldsymbol{s}_3)-\upsilon_x(\boldsymbol{s}_2))-\upsilon_x(\boldsymbol{s}_1)+\upsilon_x(\boldsymbol{s}_0)&=0,\\ (\boldsymbol{Q_{aff}}(\lambda_0,\lambda_2,\lambda_3))(\upsilon_y(\boldsymbol{s}_3)-\upsilon_y(\boldsymbol{s}_2))-\upsilon_y(\boldsymbol{s}_1)+\upsilon_y(\boldsymbol{s}_0)&=0,\\ \boldsymbol{G}(\upsilon(\boldsymbol{s}_0))(\upsilon_x(\boldsymbol{s}_3)-\upsilon_x(\boldsymbol{s}_2))-\upsilon_y(\boldsymbol{s}_3)+\upsilon_y(\boldsymbol{s}_2)&=0.\end{aligned} \tag{23}$$

The third equation ensures that the arrangement of points forms two parallel lines. That is, that the gradient direction at the point λ_0 equals to 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. (6), and can be developed in a manner analogous to eq. (15). In this case, the parameters are obtained by a system of four equations that define a gathering process in a 4D parameter space. However, the parameter space can be reduced by using the information of the shape's position. After the location parameters have been obtained, we can define two independent systems of two equations. As such, the gathering process can be performed in two 2D accumulators by solving for,

$$f^{\Delta}(\lambda_0, \lambda_2, \upsilon(s_0), \upsilon(s_2)): \begin{array}{ll} \lambda_{x_0} - a_0 = A\upsilon_x(s_0) + B\upsilon_y(s_0) & \lambda_{y_0} - b_0 = C\upsilon_x(s_0) + D\upsilon_y(s_0) \\ \lambda_{x_2} - a_0 = A\upsilon_x(s_2) + B\upsilon_y(s_2) & \lambda_{y_2} - b_0 = C\upsilon_x(s_2) + D\upsilon_y(s_2) \end{array} \quad (24)$$

7 Implementation and examples

To locate a model shape under a similarity transformation, we substitute the definitions in eqs. (13) and (15) into eqs. (7) and (8). Thus, evidence is gathered by using a pair of independent 2D accumulator spaces. The evidence gathering implementation is divided into four steps: (1) for each point λ_0 in the image chose another point λ_1 and compute $\boldsymbol{Q}_{sim}(\lambda_0, \lambda_1)$ and $\boldsymbol{Q}'_{sim}(\lambda_0, \lambda_1)$ according to eqs. (11) and (12); (2) use these values in eq. (14) to find all the points $\upsilon(\boldsymbol{s}_0)$ and $\upsilon(\boldsymbol{s}_1)$ that satisfy eq. (13); (3) use the points λ_0, λ_1, $\upsilon(\boldsymbol{s}_0)$ and $\upsilon(\boldsymbol{s}_1)$ in eq. (16) to find the parameters of the transformation and increment the associated element in the accumulator space; (4) compute the location parameter according to eq. (7). In this implementation, it is important to ensure that both points belong to the same primitive. Thus, 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 $\upsilon(\boldsymbol{s})$ as an orthogonal Fourier expansion [8]. In this representation, derivatives are easily computed. The system in eq. (14) is solved by using a successive approximation method.

Fig. 3 shows an example of the accumulation process for similarity transformations. This figure contains a synthetic shape with random noise. In this example, 65% of the data corresponds to noise. Fig. 3(b) shows the result of the extraction process superimposed on the original image. Fig. 3(c) and 3(d) show that the location and parameters accumulators contain well-defined peaks.

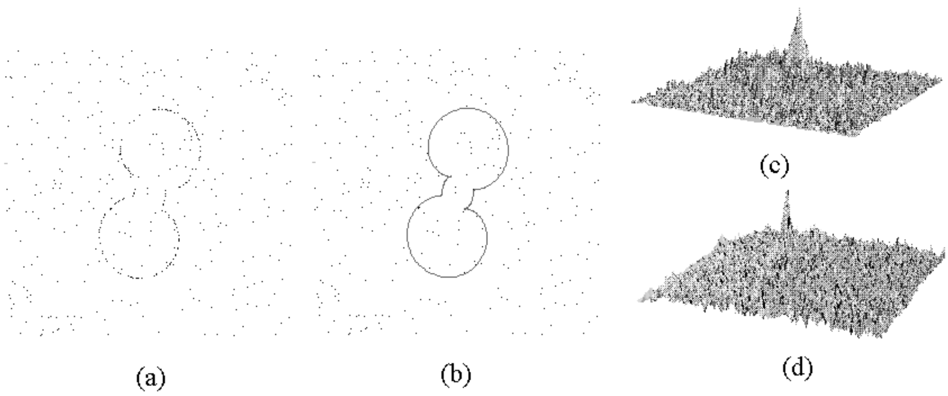

Fig. 3. (a) Example of an image with 65% of noise. (b) Result of the extraction process. (c) Location accumulator. (d) Scale and rotation accumulator.

Fig. 4 shows an example of the accumulation process on a real image. The model shape in Fig. 4(a) was obtained from a binary image of 128×128 pixels. The result of the extraction process is presented in Fig. 4(d) superimposed as a thick border. Fig. 4(e) and 4(f) show the final accumulator for the position and for the rotation and scale parameters.

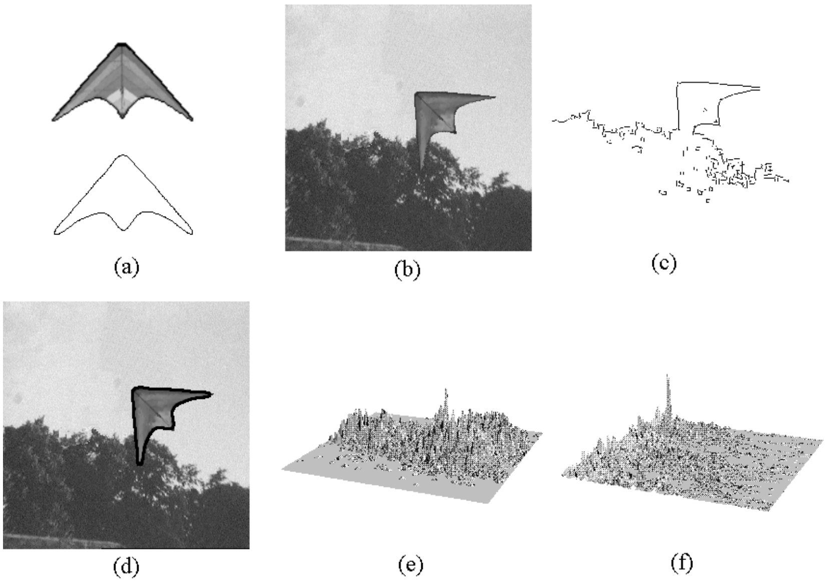

Fig. 4. Shape extraction. (a) Model shape. (b) Raw image. (c) Image edges. (d). Result. (e) Location accumulator. (f) Scale and rotation accumulator.

In order to develop an evidence gathering process for affine transformations, we substitute the definitions in eqs. (22) and (24) into eqs. (7) and (8). The evidence gathering process is divided into 5 steps: (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 than the gradient direction at the point λ_0. From these, compute the value of $\boldsymbol{Q}_{aff}(\lambda_0,\lambda_2,\lambda_3)$ and $\boldsymbol{Q}'_{aff}(\lambda_0,\lambda_2,\lambda_3)$ according to eqs. (19) and (21); (2) use these values in the system of equations in (23) to find all the points $\upsilon(\boldsymbol{s}_0)$, $\upsilon(\boldsymbol{s}_2)$ and $\upsilon(\boldsymbol{s}_3)$; (3) use the points λ_0, λ_2, λ_3, $\upsilon(\boldsymbol{s}_0)$, $\upsilon(\boldsymbol{s}_2)$ and $\upsilon(\boldsymbol{s}_3)$ to solve for the parameters A,B,C,D; (4) compute the location parameter according to eq. (6) and increment the associated element in the corresponding position in the accumulator space; (5) after all evidence has been gathered and the location parameter is known, repeat step (1) and use the points in the system of equations in (24) to gather evidence for the parameters of the transformation.

In this implementation, the solution for eq. (23) imposes significant computational load. 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 the slope of the point λ_0.

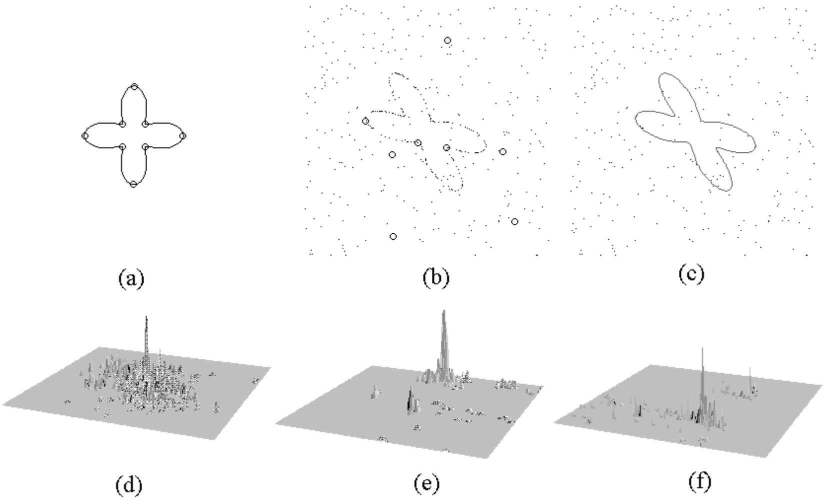

Fig. 5. (a) Model shape. (b) Synthetic image where 35% of the data belongs to the primitive. (c) Result of the extraction process. (d) Location accumulator. (e) Result for the parameters *A* and *C*. (f) Result for the parameters *B* and *D*.

Fig. 5 shows an example of the evidence gathered for affine transformations. Fig. 5(a) shows the model shape. 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 to locate the shape, we generated a collection of synthetic images containing 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 3 points of high curvature in correct positions (those which lie on the shape), whilst the other five were positioned at random. In the example 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. The accumulator presented corresponds to the parameters of the transformation. An example of the extraction process applied to a real image is presented in Fig. 6.

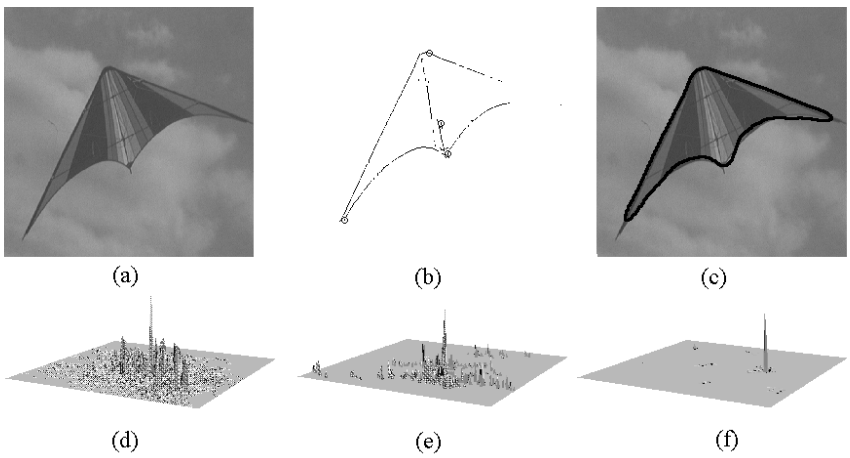

Fig. 6. Shape extraction. (a) Raw image. (b) Image edges and high curvature points. (c) Result. (d) Location accumulator. (e) Accumulator for the parameters *A* and *C*. (f) Accumulator for the parameters *B* and *D*.

8 Conclusions

We have developed a technique for including invariance in a general form of the HT for parametric models defined by similarity and affine transformations. The advantage of this characterization is that it significantly reduces the uncertainty associated with the use of higher level primitives such as lines or curves. Based on invariance characterization it is possible to locate of a shape using a 2D accumulator space. 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. We have included an effective strategy that reduces the number of correspondences between the model and the image. This is indispensable since the generality of the transformation increases the geometric complexity of the features identified in the shape. We have shown that the technique can obtain adequate results when some features are generated by background objects or are missed.

References

[1] R.T. Chin, C.R. Dyer, "Model-based recognition in robot vision", ACM Computing Surveys **18**:67-108 (1986).

[2] G. Stockman, "Object recognition and localization via pose clustering", CVGIP, **40**:361-387 (1987).

[3] Y. Lamdan, J Schawatz, H. Wolfon, "Object recognition by affine invariant matching", Proc. IEEE Conf. on CVPR, pp. 335-344 (1988).

[4] S. Umeyama, "Parameterised point pattern matching and its application to recognition of object families", IEEE Trans PAMI **15**(1):136-144 (1993)

[5] A.Califano, R. Mohan, "Multidimensional indexing for recognition of visual shapes", IEEE Trans. PAMI **16**:373-392 (1994).

[6] D.H. Ballard, D. Sabbah, "Viewer independent shape recognition", IEEE Trans. PAMI **6**:653-659 (1983).

[7] W.E.L. Grimson, D.P. Huttenglocher, "On the sensitivity of the Hough transform for object recognition", IEEE Trans. PAMI **12**:255-274 (1990).

[8] A.S. Aguado, M.S. Nixon and M.E. Montiel, Parameterizing Arbitrary Shapes via Fourier Descriptors for Evidence-Gathering Detection, Computer Vision and Image Understanding **69**(2): 202-221 (1998).

[9] G. Roth, M. D. Levine, Extracting geometric primitives", CVGIP: Image Understanding **58**:1-22 (1993).

[10] F. Bookstein, "Fitting conic sections to scattered data", Computer Graphics and Image Processing **9**:56-91 (1979).

[11] D. Forsyth, J.L. Mundy, A. Zisserman. C.M. Brown, "Projectively invariant representations using implicit algebraic curves", Lecture Notes in Computer Science **427**:427-436 (1990).

[12] H.K. Yuen, J. Illingworth, J. Kittler, "Detecting partially occluded ellipses using the Hough transform", Image and Vision Computing **7**(1):31-37 (1989).

[13] P.-K. Ser, W.-C. Siu, "A new generalized Hough transform for the detection of irregular objects", Journal of Visual Communication and Image Representation **6**(3):256-264 (1995).

[14] T.E. Dufresne, A.P. Dhawan, "Chord-tangent transformation for object recognition", Patt. Recog. **28**(9):1321-1331 (1995).