Invariant Shape Object Recognition Using B-Spline, Cardinal Spline, and Genetic Algorithm

PISIT PHOKHARATKUL
Department of Computer Engineering, Faculty of Engineering, Mahidol University
Salaya, Phuthamonthon, Nakhorn Pathom, 73170, THAILAND

SKUL KAMNUANCHAI[1] and CHOM KIMPAN[2]
Faculty of Information Technology, Rangsit University
Muang-Ake, Paholyotin Road, Patumtani 12000, THAILAND

SUPACHAI PHAIBOON
Department of Electrical Engineering, Faculty of Engineering, Mahidol University
Salaya, Phuthamonthon, Nakhorn Pathom, 73170, THAILAND

Abstract: - Object recognition is an essential part of the computer vision system. This paper proposes a genetic algorithm to search the features of model shapes of the object from model-base, to identify input shapes of the object. The dominant points are extracted from the edge of binary images using Gaussian filtering. There are two methods to compute the output features. The first, B-Spline, used the dominants to compute the control points. The second, Cardinal Spline computes the data points form the dominant points. The control points, and the data points are built a model shapes for searching by genetic algorithms to identify the input images. Then, we are compared the two method. Training data composes of original object, its translation, its rotation and its scaling. The recognition results of B-Spline implementation are 97% for rotated object, 94% for rotated and scaling object. The recognition results with Cardinal Spline feature are 97% for rotated, 95.2% for rotated and scaling object.

Key-Words: - Object recognition, Centroidal profiles, Gaussian filtering, B-Splines, Cardinal splines, Genetic algorithm

1 Introduction

Object recognition is an important goal for computer vision systems to enable the understanding of images [1]. The use of curvature or dominant points [2-5] on an unknown object boundary has been proven to be an effective means for recognition of object shapes. Most of the current object recognition involves matching the input image with a set of predefined models of object. If the images are invariant to various fluctuations of input image, this leads to the problem of recognition. It is necessary to use many of the known rotated objects precompiled, creating a model database, and this database to large to use.

In this research, in order to solve the problem as mentioned above, we use Gaussian filtering which are extracts the dominant points from the edge of binary images. Then we use the B-Spline or Cardinal Spline to compute the control points or data points for creating the model shapes. The matching process uses the data from training database to measure the similarity between the unknown images with random data from the training database which determine by genetic algorithm. The research shows the effectiveness of the B-Spline and Cardinal Spline to build the database and the selecting points using the GA for matching. The detail of recognition procedure will be illustrated in the following sections.

2 Centroidal Profile Representation

The centroidal profile is characterized by an ordered sequence that represents the distance from the digitized boundary of the object to its centroid as a function of distance along the boundary. A simple object is shown in figure 1(a) and its corresponding centroid profile is illustrated in Fig. 1(b). The centroid (X_c , Y_c) is estimated using the following formula:

$$X_c = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (1)$$

$$Y_c = \frac{1}{N}\sum_{i=1}^{N} y_i \qquad (2)$$

where N is the number of the boundary points in an object. Next, the distances $d\ (i)$, i=1,2…N, from centroid to the boundary points $\{x_i,\ y_i,\ i=1,2...N\}$, are computed starting from an arbitrary position of the boundary and tracking the boundary in a counterclockwise direction. Using as distance measure the Euclidean norm we have:

$$d(i) = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \qquad (3)$$
$$i = 1,2,...,N.$$

After, the distance function is calculated by convolving the distance function of the boundary point $d\ (i)$ with the Gaussian filtering [3] :

$$g(t,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-t^2}{2\sigma^2}} \qquad (4)$$

In this paper, we selected the scale-space filtering σ=3. The points of local maxima and minima are determined from the smooth distance function are illustrated in Fig. 1(d) and are assigned as the reference points (Fig. 1(e)) to break the boundary of the shape as boundary points. Next, we divide a half of points between these reference points. If the points that are divided have lengths over 10% of the number of boundary points, point division begin again as illustrated in Fig. 1(f). Thus, the B-splines control points of object are illustrated Fig. 1(g), and as features present of object.

3 B-Splines Representation

B-splines are piecewise polynomial curves that are guided by a set of points called the control points (CPs). The CPs are blended with a set of functions called the blending functions. Any point on the curve segment is given by a parametric from as

$$P(u) = \sum_{i=0}^{n} v_i N_{i,k}(u) \qquad (5)$$

with u as the parameter, v_i, i=0,1,…n are the n+1 control points, and $N_{i,k}(u)$, i=0,1,…,n are the (k-1)th degree blending functions. They can be evaluated recursively from

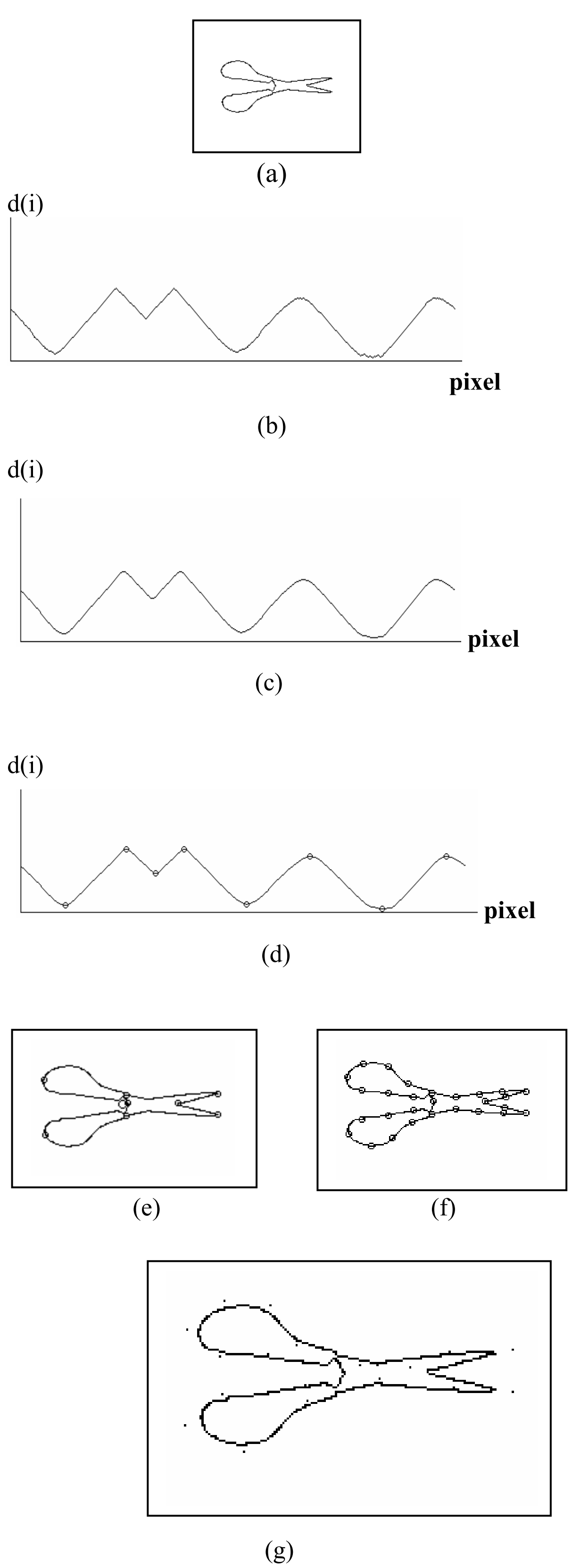

Fig.1.A simple object and boundary representations: (a)a clipper; (b) centroidal representation; (c)smooth using Gaussian filter; (d-e) local maxima and minima;(f) reference points of object; (g) control points.

$$N_{i,1}(u) = \begin{cases} 1, & \text{for } t_i < u < t_{i+1} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

$$N_{i,k}(u) = N_1 + N_2, \quad (7)$$

where

$$N_1 = \frac{(u - t_i)N_{i,k-1}(u)}{(t_{i+k-1} - t_i)},$$

$$N_2 = \frac{(t_{i+k} - u)N_{i+1,k-1}(u)}{(t_{i+k} - t_{i+1})}, \quad k = 2,3,4,\ldots.$$

The blending functions are non-zero only for an interval given by the degree of the polynomial. Cubic polynomials are most often preferred because they also preserve continuity of curvature at the point on the curve. Since the CPs are regenerative, they yield a large compression of the boundary data. Scale, translation and rotation of the shape result in similar transformation of the CPs.

Consider figure 1 which shows a piecewise curve joined with curvature continuity, i.e. cubic polynomial. For the segment i the blending functions that take non-zero values are, respectively, $N_{i-1,4}(u)$, $N_{i,4}(u)$, $N_{i+1,4}(u)$, $N_{i+2,4}(u)$.

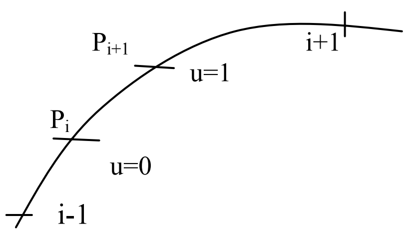

Fig.2. piecewise continuous curve.

We let the parameter u take on values between 0 and 1 for each of the spans and let the points,

$$p_i(0) = p_i \text{ for } i = 0,1,\ldots n-1$$

$$\text{and } p_{n-1}(1) = p_n. \quad (9)$$

For closed curves, $v_0 = v_{n+1}$, v_n and the B-spline equations can be reformulated in a matrix [1] form as,

$$1/6 \begin{bmatrix} 4 & 1 & . & 1 \\ 1 & 4 & . & 1 \\ . & & & 1 \\ 1 & . & 1 & 4 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \\ v_n \end{bmatrix} = \begin{bmatrix} p_0 \\ p_1 \\ \\ p_n \end{bmatrix} \quad (10)$$

4 Cardinal Splines Representation

The principle of the Cardinal splines is based on calculation with 4 points that are together, 2 points, that define the beginning and the end, and 2 other point that define the slope. The curve P_k and P_{k+1} is defined by $\overline{P_{k-1}P_{k=1}}$ and $\overline{P_k P_{k=1}}$ through the parameter t which is called the tension parameter. This parameter controls the shape of the curve.

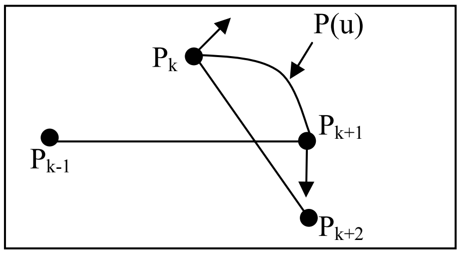

Fig.3. point of cardinal spline

When set p_{k-1}, p_k, p_{k+1} and p_{k+2} are control point. Any point on the curve segment is given by a parametric form as

$$P(u) = P_{k-1}A(u) + P_kB(u) + P_{k+1}C(u) + P_{k+2}U(u) \quad (11)$$

When

$$A(u) = -Su^3 + 2Su^2 - Su$$

$$B(u) = (2-S)u^3 + (S-3)u^2 + 1$$

$$C(u) = (S-2)u^3 + (3-2S)u^2 + Su$$

$$D(u) = Su^3 - 2Su^2$$

U = (position point of picture between P_k *and* P_{k+1}*)/15*

When S = (1.0-t)/2.0 ; $0 \le t \le 1$ This paper t is 0.00

5 Shape Representation

The boundary points from the centroidal profile of the object represented by a set of points p_is is chosen on the boundary of the object from which the computed points. As we discussed in this Section, control points are extracted from the boundary points. The centroid was used as the central reference point, and the distance between each control point and the centroid is computed. The ten length maxima are selected stored in a privileged segment. In the following, we assume privileged segments that are both the model, and the input (scene) description. These are given the form: (l_i, θ_l, p_i), for i=1,2,..,10 where l is the length between control point and centriod, θ is the control point orientation measured relative to the horizontal axis, and p is the position of the control point from the control points segment .

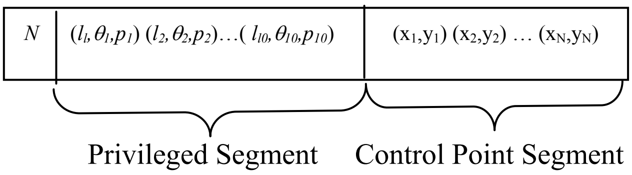

Fig. 4. Shape description.

Figure 4 shows examples of the shape description of the model shapes and input shapes. From figure 4, control points are stored in a control point segment. The number of control points involved in these description ranges between 18 and 40.
Genetic algorithms (GAs)[6,7,8] are general purpose search techniques based on principles inspired from the genetic, and evolution mechanisms observed in natural systems, and populations of living begins. A GA is usually implemented according to the following steps:

5.1 Chromosome Representation

In this paper, a binary string built from the matching model shape, the privileged of the model shape, and the privileged of the input shape represents the chromosome. Each part is encoded into a four-bit representation, resulting in a chromosome with 12 bits. A population of 20 chromosomes was employed in the algorithm.

Example: The following is an example individual:

[(0001 0001 0011), (1001 0101 0101), . . .]

This individual to the first feature of the first model shape mapped the third feature of the input shapes.

5.2 Fitness Function

The model position was defined by a transformation T, the product of a rotation, a scaling, and a translation. The transformation T described by a parameter vector $v = (k.cos\theta, k.sin\theta, tx, ty)$, such as the image (x^*, y^*) of an arbitrary point (x, y) of the model description was given by the set of equations

$$x^* = tx+x.k.cos\theta - y.k.sin\theta \qquad (12a)$$
$$y^* = ty+x.k.sin\theta + y.k.cos\theta . \qquad (12b)$$

Fitness was calculated by testing the compatibility of the input shapes features, and the corresponding model shape feature. The difference between input shapes I_j feature, and model feature M_i was measured by mean of a difference function d_r, as defined below:

$$d_r = \sqrt{(X_{I_j} - X)^2 + (Y_{I_j} - Y)^2} \quad , \qquad (13)$$

where (X_{I_j}, Y_{I_j}) is the I_jth feature of the input shape and (X , Y) was calculated using the following formula:

$$X = tx_0 + X_i.s_0.\cos\theta_0 - y_i.s_0.\sin\theta_0 \qquad (14a)$$
$$Y = ty_0 + X_i.s_0.\sin\theta_0 + y_i.s_0.\cos\theta_0 \qquad (14b)$$

and

$$s_0 = (l_j)/(l_i) \qquad (15a)$$
$$\theta_0 = \theta_j - \theta_i \qquad (15b)$$
$$tx_0 = X_j - s_0(X_i \cos\theta_0 - Y_i \sin\theta_0) \qquad (15c)$$
$$ty_0 = Y_j - s_0(X_i \sin\theta_0 + Y_i \cos\theta_0) \qquad (15d)$$

where (X_i , Y_i) is the i th feature of the M th model shape. Fitness function was calculated using the following formula:

$$Fitness = -(\sum_{r=1}^{n} d_r + (\sum input\ or\ model\ features\ remain\ unassigned\ )) \qquad (16)$$

6 Experimental

In the experiments, we tested this proposed method to the invariant objects recognition. The object shapes were represented by the computed points, which were computed by the B-Spline or Cardinal Spline. The shape of object was composed the length distances between the centroid and computed points, the angles of computed point orientation, and the coordinate of computed points respectively. The learning set consists of 15 reference objects (Fig.12) that have 300x300 pixels. The parameters of all model shapes were stored in the database. In the recognition stage, the genetic algorithm searches models in the database to select a model that has the best-match sequence with input image. The feature of the purposed method consists of 12 bit strings. The population size is 20, and the number of selected strings for mutation probability for each bit was 0.033. The number of iterated generations was

150, with a sufficient number for obtaining the solution.
The proposed methodology was used to invariant object recognition, such as the examples in figure .
The input image that has different orientation was captured by digital camera.
The system was tested with objects that have 7 different rotations (totaling 105 images of objects), 4 different sizes (totaling 60 images of objects), and 32 different rotations and sizes (totaling 480 images of objects),. These objects were obtained under the same illumination conditions

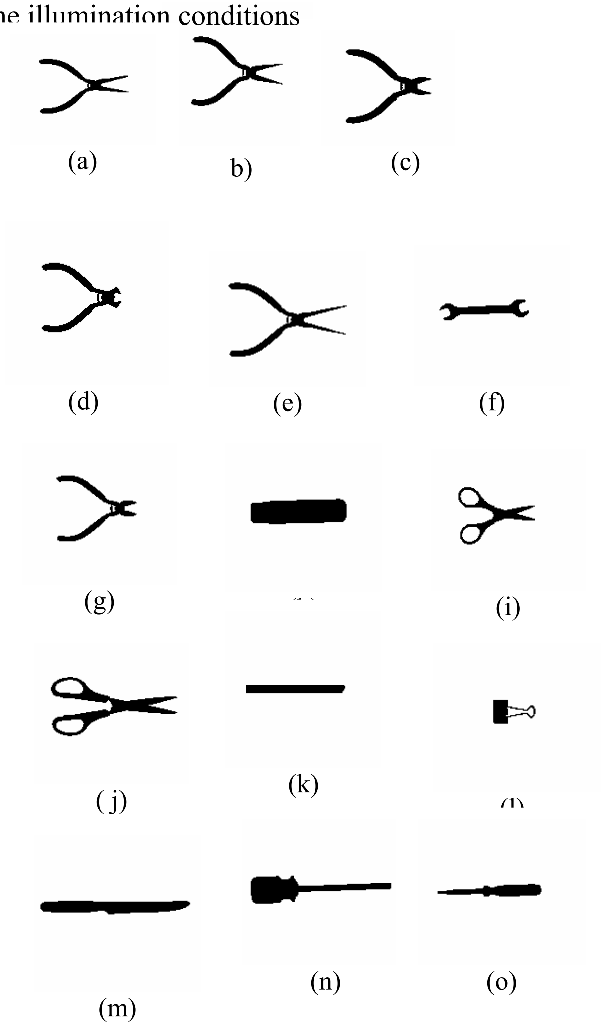

Fig.5 Example of Reference objects.

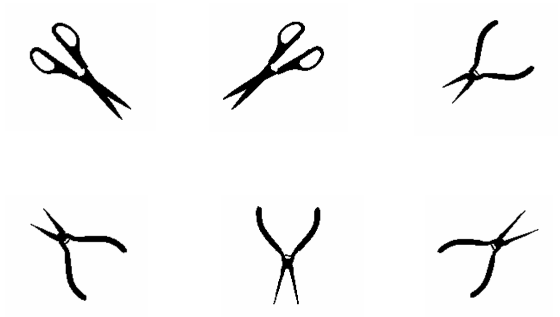

Fig.6 Example of rotated objects.

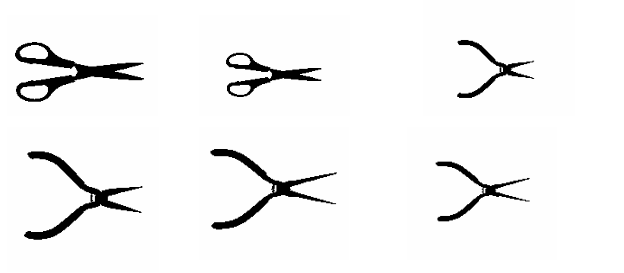

Fig.7 Example of scaled objects.

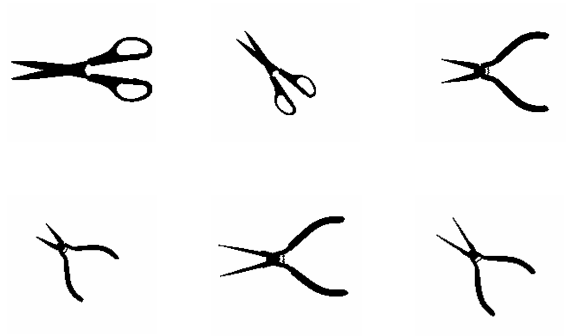

Fig. 8 Example of both scaled and rotated objects.

6.1 The B-Spline recognition results are shown in Table 1.

Tables 1: Recognition accuracy (%).

Object	Rotated Objects	Scaled Objects	Rotated and scaled objects
a	100	100	94
b	100	100	94
c	100	75	81
d	71	100	81
e	100	100	100
f	100	100	100
g	100	75	94
h	100	100	100
i	100	100	100
j	100	100	100
k	86	75	91
l	100	100	100
m	100	100	100
n	100	100	97
o	100	100	97
Average	97	95	94

6.2 The Cardinal-Spline recognition results are shown in Table 2.

Table 2 Recognition accuracy (%).

Object	Rotated Objects	Scaled Objects	Rotated and scaled objects
a	100	100	93.7
b	100	100	100
c	100	100	93.7
d	71.4	75	81.3
e	100	100	100
f	100	100	100
g	100	75	93.7
h	100	100	100
i	100	100	100
j	100	100	100
k	100	100	93.7
l	100	100	100
m	100	100	100
n	85.7	100	90.5
o	100	100	95.9
verage	97.1	97	95.2

7 Conclusion

In this particular research, we studied invariant object recognition using the control points of B-spline and data points of Cardinal spline with genetic algorithms. Genetic Algorithms have been proven to be powerful methods in search, optimization and machine learning. Experimental results show that the genetic algorithm has been successful in shape-matching experiments attempted so far. The algorithm is fast, and explores a relatively small number of elements of the search space. However, a 2-D object can have different shapes depending upon the position, orientation, and size. The boundary shapes of the object were composed of the control and data points. The method is adjusted to be independent of translation, scaling, and rotation, by using relative distances of computed points to the centroid. The advantage of the centroidal profile is reference points tend to be a stable point of reference for the object.

In the case of features being data points, using cardinal spline extracting directly from the boundary points without using the inverse matrix, the experiment shows that this method is better than using control points of the B-Spline method.

To compare other methods, for example the Gaussian Filter [3] which has a rating recognition less than the 2 other mentioned, because the number of dominant points recognized by the Gaussian Filter is less. Many of the points are lost. As for Convex Factor [4] and the Corner Detector [5] these have a recognition less than the Gaussian Filter. The corner Detector is the fastest of them all, yet has the worst recognition.

The main advantages of the Cardinal spline and B-spline recognition system are as follows: Feature points stored in a database are between 18-40 features for Cardinal spline and not over 60 features for B-spline; the recognition method was adjusted to be independent of transformation such as translation, scaling, and rotation; it used privileged segments to calculate orientation, scaling, and translation.

References:

[1] D. Ballard and C. Brown, *Computer Vision*, Prentice Hall, Englewood Cliffs, 1982.

[2] R. Chin, and C. Dyer, “Model-Based recognition in robot vision”, *Compuingt.Surv*eys, Vol. 18, No. 1, 1986, pp. 67-108.

[3] S. Foitong, P. Phokharatkul, O. Pingern, and C. Kimpan, “Invariant Object Recognition Using Dominant Points and Genetic Algorithms”, *Proceedings of the International Conference on Robotics, Vision*, Information and Signal Processing, 2003, pp. 532-537.

[4] P. Phokharatkul, S. Foitong, and C. Kimpan, “Object Recognition Using Characteristic Component and Genetic Algorithms”, *IEEE Region 10 International Conference on Electrical and Electronic Technology*, 2001, pp. 345-349.

[5] S. Foitong, C. Kimpan, and P. Phokharatkul, “ Invariant Object Recognition Using Shape Descriptors and Genetic Algorithms”, *World Multiconference on Systemics, Cybernetics and Informatics*, Vol.VI, 2001, pp. 274-278.

[6] Endor Ozcan, and Chilukuri K.Mohan, “Shape recognition using Genetic algorithms.” *Proceeding of IEEE International Conference on Evolutionary Computation*, 1966, pp. 414-416.

[7] P.W.M. Tsang, “A Genetic for Affine Invariant Object Shape Recognition.”, *Genetic Algorithm in Engineering System: Innovations and Applications*, 1995, Conference Publication No. 414.

[8] D.E. Glodberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley, 1989.