An Elliptic Approach to Fuzzy Pattern Recognition

Fatiha KARBOU(*). Fatima KARBOU(**). M. KARBOU(***)

(*). ENSEM. Dept. E.G, B. P. 8118. Oasis. Route d'El Jadida. Casablanca. Morocco
(**). IAV HASSAN II. dept. Topography. B.P. 6202. Rabat-instituts. Rabat. Morocco
(***). Sophia Concept.47. rue Planquette Belvédère. Casablanca . Morocco

Abstract

If we want to compare the form of two objects. the human vision takes into account the parameter's width/length/height at the same time. However. the machine needs to compare widths then lengths and finally height. In each comparison the machine considers only one character. The goal of this paper is to imitate the human manner of comparison and recognition by using two or three characters instead of one during the comparison. The ellipse is a first approach of comparison because it provides us a general and a simple relation that can link two parameters that are the half axis of the ellipse. Indeed, we assimilate each character to a half axis of the ellipse and the result is a geometrical figure that varies according to values of the two characters.

Keywords: Fuzzy sets. Pattern recognition. clustering. ellipse

1. Introduction

In fact. the pattern recognition takes generally in consideration only one character at each comparison. In deed. the comparison is made in two stages: first. we compare the characters one by one and we make the summons of all results in order to use the whole of the characters. Nevertheless. the human vision. takes in consideration the totality of characters at the same time. Consequently. the decision is taken globally without dividing the comparison. For example. if we want to compare the form of two objects. the human vision takes into account the parameter's width/length/height at the same time. However. the machine needs to compare widths then lengths and finally height. In each comparison the machine considers only one character.

The goal of this paper is to imitate the human manner of comparison and recognition by using two or three characters instead of one during the comparison. "The generalization of this approach for any number of characters will make the object of a next work". So. we try to seek a relation in order to link the characters to make comparison using the whole of the characters as one entity. Indeed. human vision has no need to make the processing of comparison separately. character by character: because the parameters width/length or width /width or all other relation. are information that appears clearly.

2. Characterization of the relation between two characters by an ellipse

The ellipse is a first approach of comparison because it provides us a general and a simple relation that can link two parameters that are the half axis of the ellipse. Indeed, we assimilate each character to a half axis of the ellipse and the result is a geometrical figure that varies according to values of the two characters. Thus. we take into account the two characters as an alone entity. If we take the example of the next figure that is the generalization of the function XOR.

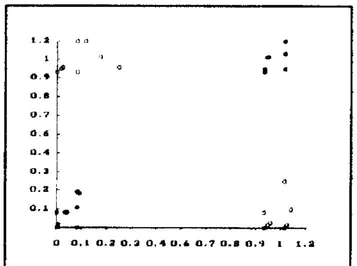

Fig. 1 An example of Xor Problem

We can see that a regular metric would not allow us to solve the problem of XOR. However. if one represents each point by an ellipse we will see that close elements to (0, 0) and (1,1) would be ellipses whose eccentricity is neighbor 0. And close elements to (0,1) and (1,0) will have ellipses whose eccentricity is neighbor 1. Thus. we have succeeded to separate two clusters only by using properties of ellipses.

3. Description of the method

Instead of separating the two characters. we use an ellipse and we assimilate each character to a half axis of the ellipse. In the case of intervals. the center of the ellipse is the point formed by the two inferior interval extremities. and in the case of values. the center is arbitrary and is common to all ellipses as illustrated in the following figure.

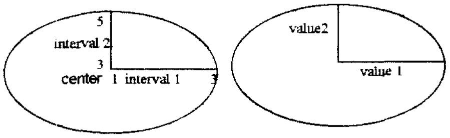

Fig.2 The ellipse formed between the two attributes

Moreover, in order to recognize a flat or a round form or another form, it is not necessary to take into account the size of elements but only the form. What is easily obtained with concentric ellipses (see the next figure).

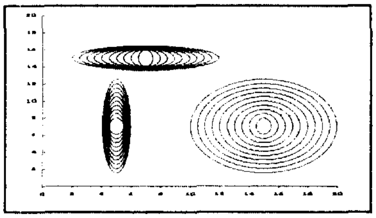
Fig.3 An example of concentric ellipses

4. Comparison of ellipses

The comparison of ellipses can be made using several methods. Indeed, let consider E1=(O1,a, b) the ellipse describing the relation between the two character C1 and C2 for an element of reference VR, and E2 = (O2, x, y) representing the same relation between the same characters for an other element V. where O1=(x10,y10) is the center of E1, and O2=(x20,y20) the center of E2. b is the small (respectively the great) axis of E1 and x (respectively the great) is the small axis of E2.

The similarity between VR and V, necessitate the measure of the degree of inclusion of E2 in E1 or the inverse. We can have several equations of comparison that use equations of the ellipse or the description by intervals or others relationships[].

4.1. Comparison by perimeter or surface

If the size of the ellipse is determinant in the decision, then the use of the perimeter or the surface is justified. Indeed, we can propose a degree of inclusion based on the perimeter and/or the surface of ellipses. Let consider for example the degree of inclusion following, that takes account into the perimeter and the surface at the same time; k1, k2, and k3 are coefficients of weighting.

$$D_{inclusion}(E_1,E_2)=\frac{1}{1+kD(O_1,O_2)+(k_2D_p(E_1,E_2)+k_3D_s(E_1,E_2))^{k1}}$$

where

$D(O_1,O_2)=(x_{10}-x_{02})^2+(x_{10}-x_{02})^2$:

compares the remoteness of centers,

$D_p(E_1,E_2)=(\pi(a+b)-\pi(x+y))^2$:

compares perimeters.

$D_s(E_1,E_2)=(\pi ab-\pi xy)^2$:

compare surfaces.

Remark: If we do not want to hold the remoteness of ellipse's centers, one can take k=0. We can do the same thing if we want only to compare the perimeter (k=k2=0) or the surface (k=k1=0).

4.2. Comparison by taking into account the deformation of the ellipse

If the form of the ellipse is determinant in the decision, (for example, if we want to know if objects have almost the same form) then an equation of inclusion degree taking into account the form is justified. Indeed, we can propose a degree of inclusion based on the relation linking the half small and the half great axis. Let take for example the following degree of inclusion:

$$D_{inclusion}(E_1,E_2)=\frac{1}{1+kD(O_1,O_2)+(k_2D_e(E_1,E_2))^{k1}}$$

where

$D_e(E_1,E_2)=(e_1-e_2)^2$:

e_1 (respectively e_2), represents the eccentricity of E_1 (respectively E_2). k, k_1 and k_2 are coefficients of weighting. This function calculates the degree of deformation of the ellipse (in addition to the removing of the two ellipses).

4.3. Comparison generalizing the degree of inclusion using intervals

An ellipse is characterized by its center and its two half axis. Therefore to be able to separate two ellipses it will be necessary to compare the removing of their center and the width of its half axis that form two intervals.

The degree of inclusion of E1 in E2, or the inverse, is a function that depends on the distance D(O1,O2) that compares the removing of centers and the degree of inclusion between intervals formed by the half axis. Indeed, we can propose the following degree of inclusion:

$$D_{inclusion}=\mu_1 e^{-k1D(O_1-O_2)}+\mu_2 D_{inclusion}(\mathrm{int}(a),\mathrm{int}(x))+\mu_2 D_{inclusion}(\mathrm{int}(b),\mathrm{int}(y))$$

where

- μ_i are coefficients of weighting whose sum is equal to 1.

- int(t) is the interval formed by the half axis t. This interval is equal to [O - x/2, O+x/2] where O is the center of the ellipse.

- $D_{inclusion}(int(),int())$ is the interval fuzzy inclusion degree [1,2]

Case of concentric ellipses: In the example illustrated in the figure (3.1.1) we present some different concentric ellipses that move following one axis or two axes.

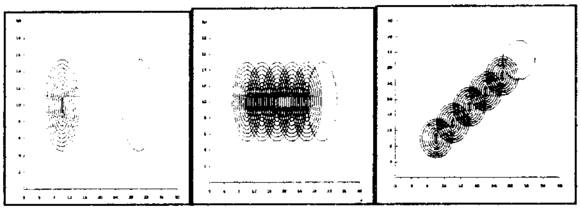

Fig.4 ellipses moving following one axis or two axes.

5. Results

The fuzzy inclusion degree of E1 in E2 is greater once E1 nears E2. Its maximal value (≤ 1) is obtained when the centers O1 and O2 of the two ellipses coincides. It depends on sizes, forms and removing of the two ellipses. It is equal to 1 when the two ellipses are exactly identical and confused, and the more the difference of the two ellipses is important, the more this value decreases.

5.1. Results for the relative inclusion degree to the deformation of ellipses

If we make a displacement following the axis of the X. Or if we displace following the two axes.
Although we con see that the degree of inclusion of the ellipses is close because the ellipses have the same deformation.

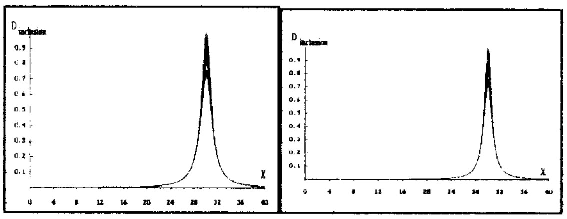

Fig.5 Results relative to the inclusion degree for the deformation of ellipses

5.2. Results for the relative inclusion degree to sizes of ellipses

If we displace following one axis. If we displace following the two axes.

We can see that it has no difference between the two figures. And if the sizes of the ellipses are determinant, the axis of displacement is not important. Thus, to displace following the axis of X or the axis of the Y, is the same thing. We can also see that the inclusion degree increase with the size.

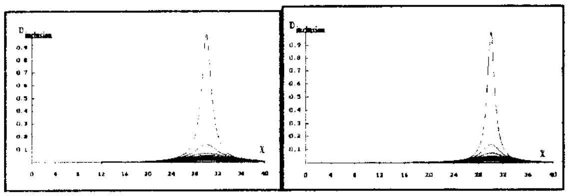

Fig.6 Results relative to the inclusion degree for the size of ellipses

5.3. Result for the degree of interval inclusion

If the totalities of ellipses displace horizontally. An alone parameter following the axis of the Y that does not change. The inclusion degree let us see that the displacement is made following one axis. Indeed, the termination of inclusion curves is a right. If the ellipses displace following the two axes, the inclusion degree decreases quickly.

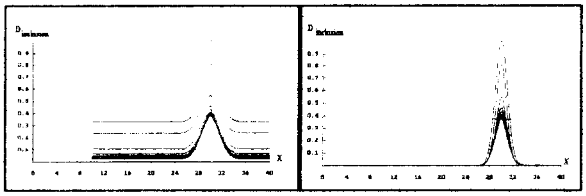

Fig.7 Results relative to the inclusion degree for interval inclusion

6. Application to the case of Iris data of Anderson

We have already, put in obviousness during the using of fuzzy curves [1,2] that there is a relation that links elements of this set.

The Iris data of Anderson[3] has often been used as a standard set for testing the performances of data analysis's algorithms and discrimination's criteria. Here, there are 50 plants in each of two varieties of Iris represented in the data: Virginia iris and Versicolor iris.

So, This distinction appeared from curves must be lucid visually. But the simple measurements of length and of width is insufficient to express similar relation. Thus, we try to express these relations simply and generally with an ellipse.

We have chosen this set to emphasize the fact that the combination of two characters or three can informs us more on the real structure of the set. Thus, in the next figure, we have illustrated the result obtained for some combination's width/width, length/length or width/length. The combination (i-j) indicate the relation between the character the and a character j. We can see the there is a separation between the two varieties.

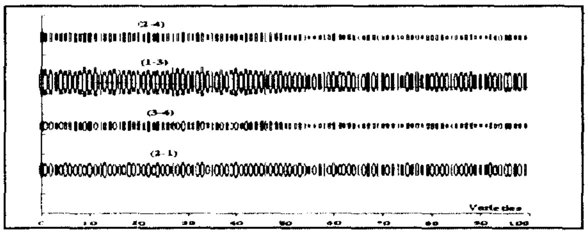

Fig.8 Ellipses formed between two attributes

The most interesting of this method is the possibility to realize a separation using one element only. That means that one element can represent its cluster. We can assimilate this relation at human vision.

To show the power of the separation, we take one element of these Iris and we try to give the separation that it provides. Following the diagram of ellipses, the two clusters are separated by using only one relation: the ellipse representing width/length (the relation 3-4 that characterizes the characters 4 and 3). It similar to the human vision.

Results

The separation of the classes is very visible only with one element (figure 1.(a)).

We have see in the previous figure that the size of the ellipse is determinant so the separation with perimeter is confirmed figure (1. (b)).

We have seen that the Iris data of Anderson had the same deformation of the ellipses, that is confirmed with the inclusion degree by the form (figure 1.(c)), because all the elements had nearly the same inclusion degree.

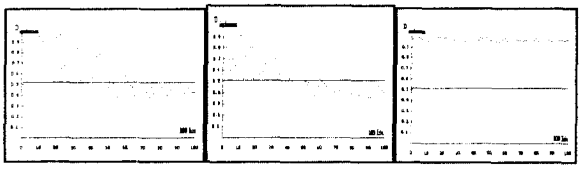

Fig.9 Results of the degree : (a)by interval, (b) by perimeter, (c) by the form

7. Generalization to the case of a three characters: a cylindrical approach

The generalization to the case of a relation binding three characters is made easily by considering a cylinder whose basis is the ellipse of the two first characters and the height is the value of the third character (as illustrated in the following figure) . The generalization of the inclusion degree is easy.

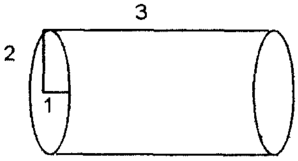

Fig.10 A cylinder representing a combination of three attributes

7.1. by using the degree of interval inclusion

We are interesting only for determination of the most discriminate attributes, so, have used just combination of two characters to seek the best combinations .And it is necessary to target the determinant characters. This will identify most discriminate characters. We have determined that the character 4 is the most discriminate character because the combination (4-4-4) give the most separation. It necessary to note that we use only one element for separation.

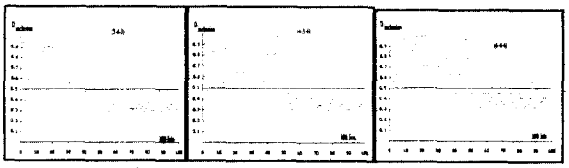

Fig.11 Results relative to the interval inclusion degree

7.2. Using the surface

The separation by surface is also confirmed. But it's not so illustrating.

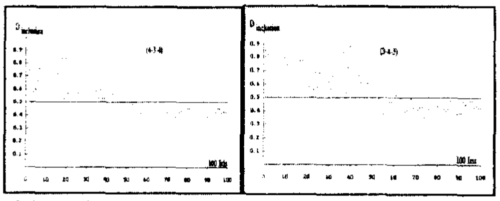

Fig.12 Results relative to the surface inclusion degree

Conclusion

the results obtained with this new approach is encouraging. They open a domain of very hopeful research.

8. Cases of circular clusters

The resolution of a problem depends strongly on the manner followed to resolve it. Thus, we can say that this resolution depends on the angle with which we consider the problem. The manner of resolution can be analytic, graphic or physic.

We are concerned by the graphic resolution of some problems of classification in the plan. So, to be able to classify a certain number of elements, we try to construct prototypes that take into account the totality of data using their properties.

In this paper, we try to detect circular clusters. Therefore, Data is the coordinate (x, y) of each point referring to (OXY); and we try to attribute elements to a cluster or to another.

However, we consider that it is more judicious to use the datum (x, y) in stead of using the datum x and the datum y separately. Consequently, if M is a point represented by the couple (x, y), we will represent it by the ellipse of half axis x and y. This representation is the graphic approach that we are going to use in order to realize the classification. Indeed, it is simpler to compare two ellipses than two points especially if this comparison is made graphically.

8.1 Analysis of bowls formed by ellipses of circular clusters

To resolve different problems, we are usually tempted to find the relation that links parameters concerning the problem considered. However, in a first approach, we are not going to seek those relations but only to valorize them. A such valorization can guarantee a simple resolution of the problem considered without having to support heavy and complex mathematical equation. To realize this goal, we should approach the problem using another method.

The human vision arrives to rencogniez or to join points of an even circle by following an order and an orientation well defined in the space. (see the following figure)

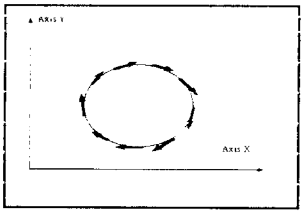

Fig.13 Example of order and orientation

We would like to imitate the human manner in order to recognize a circular cluster or another regular form.

The imitation that we plan to realize is based on this famous order and orientation characteristic of human. Thus, for a well definite circle, abscissas and the orderly follow an order and a well precise orientation, namely:

- x and y varies similarly, (increase or decrease)
- x and y vary differently, one increases and the other decreases

It is necessary to fix a starting point and an orientation.

Thus, if we have a number n of points then we seek a starting point and an orientation, and we begin by ordering these next points following their abscissa and according to the orientation chosen. Finally, we follow the evolution of bowls of classification formed by the consecutive juxtaposition of ellipses that represent points of the plan.

If we examine these following graphs obtained by piling ellipses following x increases, we notice that the order and the orientation in which ellipses are piled are very important in the training of bowls of classification. Each element has a well precise place and in a known order in the bowl.

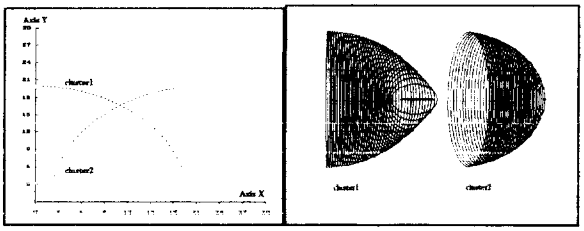

Fig.14 Two circular clusters and their bowls

Indeed, if x increases and y increases (case of cluster2), we see although that the bowl is opened and grown following the two axes. This evolution is made little by little using consecutive ellipses or consecutive points, so, we try to put in order elements before classifying them. (x increases gradually and y increases gradually.) Moreover, when x increases and y decrease, the bowl is closed at the beginning. (see cluster 1)

Examine now the case of an element that does not belong to the two clusters. We find that its ellipse is not included in any bowl of classification, because it does not respect the order of magnitude between x and y. I.e. it does not find a place in the bowl Indeed, y is greater than the order of the two bowls for the same abscissa.

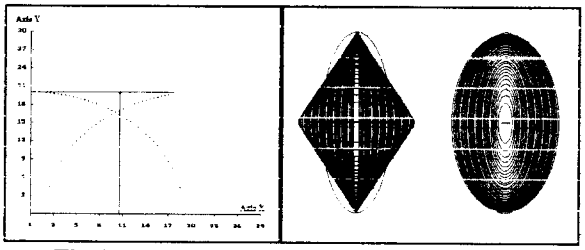

Fig.15 A point no belonging to two circular clusters with their bowls

Let now examine the case of an element belonging to a cluster 2. We can say that we have a perfected concordance between the ellipse representing the point and the bowl of the cluster 2, however, the discordance with the bowl of the cluster 1 is visually too clear. Indeed, the ellipse representing the point cut several ellipses in the bowl 1formed only by concentric ellipses. (ellipses do not cut.)

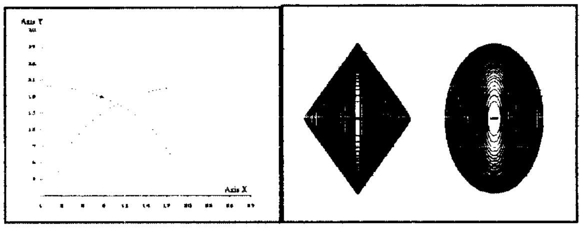

Fig.16 A point belonging to one of the two circular clusters with their bowls

3. General illustration

- With half concentric circle, the two generated bowls are as concentric and well ordered but different, one is included strictly in the other (they do not touch). Indeed, an element belonging to a half circle is not going belonging to the other.
- With two half circle with similarly rays and different centers, each generated bowl is the continuation of the other in magnitude of size since abscissas are distanced.

If the two half circles are tangent, the two generated bowls will be tangent as well. (See figure)

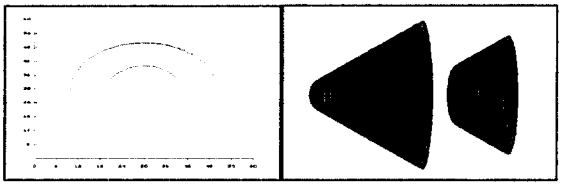

Fig.17 Two half concentric circles with their bowls

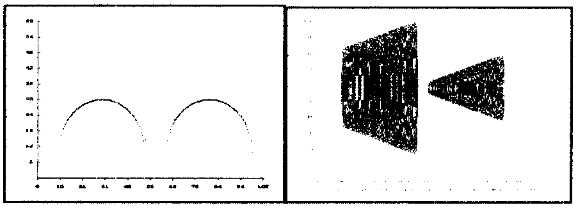

Fig.18 Two half circles with similarly rays and their bowls

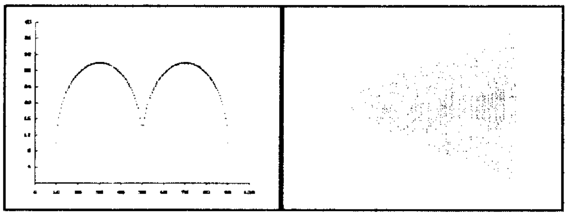

Fig.19 Two half secant circles with their bowls

4. Method's description

If we examine the position of elements of a circle in the plan we have the next cyclic consecutive cases:

1. x increases and y increases,
2. x increases and y decreases,
3. x decreases and y decreases,
4. x decreases and y increases,

Therefore to regroup elements of a circle, It is necessary to respect the order and the orientation stated.

Thus, to be able to detect points of a circle, it is necessary:

1. to order elements according to their abscissa following an order growing.
2. to seek closest ellipses in position (x, y) and in size.
3. to pile ellipses to form the bowl.

This algorithm is in its simplest form. Thus, the piling of ellipses facilitates the training of classification. Indeed, we will have any need to adjust the prototype of each cluster since each element classified to a position and a well precise rank. The generalization of such algorithm will be given later.

Conclusion

The results obtained are very encouraging because they are simpler to use than other mathematical relations. We will generalize the method to the case of the presence of noise.

9. References and bibliography

1. F. KARBOU, New Fuzzy Concepts, FUZZ-IEEE'97, Six IEEE international Conference on Fuzzy Systems, Barcelona, Spain, 1-5 JULLY, 1997.
2. F. KARBOU & all , An Interval Approach to clustering FUZZ-IEEE'98, Anchorage, Alaska, USA, May 4-9, 1998
3. J. M. KELLER, M. R. GRAY, J. A. GIVENS, A fuzzy K-Nearest Neighbor Algorithm, IEEE, Vol. SMC-15, N° 4, Juillet/Aout (1985), pp 580-585
4. Michael P. WINDHAM, Cluster Validity for Fuzzy Clustering Algorithm, Fuzzy sets and systems 5 (1981) 177-185, North-Holland Publishing Company