A New Method for Fuzzy Estimation of the Fractal Dimension and its Applications to Time Series Analysis and Pattern Recognition

Oscar Castillo and Patricia Melin
Computer Science Department, Tijuana Institute of Technology
P.O. Box 4207 Chula Vista CA 91909, USA.
ocastillo@tectijuana.mx, pmelin@tectijuana.mx

Abstract:
We describe in this paper a new method for the estimation of the fractal dimension of a geometrical object using fuzzy logic techniques. The fractal dimension is a mathematical concept, which measures the geometrical complexity of an object. The algorithms for estimating the fractal dimension calculate a numerical value using as data a time series for the specific problem. This numerical (crisp) value gives an idea of the complexity of the geometrical object (or time series). However, there is an underlying uncertainty in the estimation of the fractal dimension because we use only a sample of points of the object, and also because the numerical algorithms for the fractal dimension are not completely accurate. For this reason, we have proposed a new definition of the fractal dimension that incorporates the concept of a fuzzy set. This new definition can be considered a weaker definition (but more realistic) of the fractal dimension, and we have named this the "fuzzy fractal dimension".

1. Introduction

Traditionally, we would assign a particular geometrical object a crisp value of the fractal dimension, and this numerical value was considered as a specific label for the object. However, this numerical value is difficult to use in classification or recognition applications because calculated values won't match these crisp values. We have experienced this problem when we used this idea for classifying bacteria with the fractal dimension [2, 7]. We have found particular numerical labels for each of the bacteria, but when we try to use these values for recognizing specific bacteria in samples we have run into problems because of uncertainties. For this reason, we have proposed the following scheme for estimating the fuzzy fractal dimension of a set of geometrical objects. First, we calculate the standard fractal dimension of the objects, using the box counting algorithm with samples of points from the objects. Second, with the crisp values for the fractal dimensions of the objects build linguistic values for the dimensions, these will be the fuzzy fractal dimensions of the objects. Third, using these linguistic values of the fractal dimensions build a set of fuzzy rules that relate each object with each rule. This set of fuzzy if-then rules can be considered a classification scheme for the set of geometrical objects, and can be used for recognizing these objects because a particular value is mapped to an object. We can apply this method either for pattern recognition or for time series analysis as follows. First, we need to build the specific classification rules for the application using the fractal dimension. Then, we need to implement a method for sampling the object to obtain the data needed to calculate the crisp value of the fractal dimension. Finally, we use this crisp value as input in the set of fuzzy rules to obtain as output the specific classification for the object. For real image processing this can be used to recognize a particular object as needed for robotic applications [3]. For time series analysis, this can be used for modeling and forecasting purposes. In any case, the generalization of the mathematical concept of the fractal dimension [6], to include now the ideas of fuzzy logic [9] is also important from the theoretical point of view because is only the initial point in the fuzzy generalization of Fractal Theory.

2. Fractal Dimension of an Object

Recently, considerable progress has been made in understanding the complexity of an object through the application of fractal concepts [6] and dynamic scaling theory. For example, financial time series show scaled properties suggesting a fractal structure [2, 4]. The fractal dimension of a geometrical object can be defined as follows:

$$d = \lim_{r \to 0} [\ln N(r)] / [\ln(1/r)] \qquad (1)$$

where N(r) is the number of boxes covering the object and r is the size of the box. An approximation to the fractal dimension can be obtained by counting the number of boxes covering the boundary of the object for different r sizes and then performing a logarithmic regression to obtain d (box counting algorithm). In Figure 1, we illustrate the box counting algorithm for a hypothetical curve C. Counting the number of boxes for different sizes of r and performing a logarithmic linear regression, we can estimate the box dimension of a geometrical object with the following equation:

$$\ln N(r) = \ln\beta - d \ln r \qquad (2)$$

this algorithm is illustrated in Figure 2.

The fractal dimension can be used to characterize an arbitrary object. The reason for this is that the fractal dimension measures the geometrical complexity of objects. In this case, a time series can be classified by using the numeric value of the fractal dimension (d is between 1 and 2 because we are on the plane xy). The reasoning behind this classification scheme is that when the boundary is smooth the fractal dimension of the object will be close to one. On the other hand, when the boundary is rougher the fractal dimension will be close to a value of two.

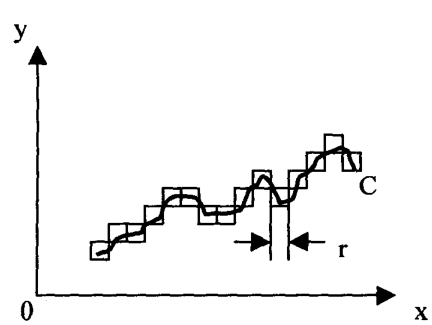

Figure 1 Box counting algorithm for a curve C.

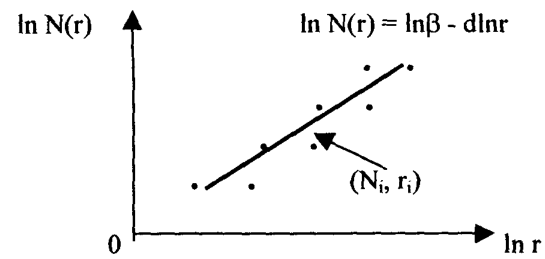

Figure 2 Logarithmic regression to find dimension.

3. Fuzzy Logic for Object Classification

We can use a fuzzy rule base as a classification scheme if we can make a suitable partition of the input space such that we are able to distinguish different geometrical objects by their characteristics. We will consider geometrical objects in the plane for simplicity. We can now use fuzzy clustering techniques [1, 8] to cluster the data, and then build a fuzzy rule base that will actually be a classification scheme for the particular application.

We will consider that we have n objects O1, O2, ..., On, and that we are able to apply fuzzy clustering techniques to obtain n pairs (Xi, Yi) i=1,...,n, which are the respective centers of the n clusters. Then a fuzzy rule base can be constructed as follows:

If X is x_1 and Y is y_1 then Object is O_1
If X is x_2 and Y is y_2 then Object is O_2 (3)
... ...
If X is x_n and Y is y_n then Object is O_n

These rules can be used for pattern classification or time series analysis because in both cases the data has the same general structure. For applications of higher dimensionality this approach can be generalized in a straightforward manner, but of course the problem is that the number of rules increases dramatically (which is known as the curse of dimensionality).

To illustrate these ideas we will consider a particular application. Lets consider the problem of forecasting the time series of the exchange rate US dollar/MX peso. We used the time series of average weekly rates for 36 weeks to find the fuzzy model as explained above. We then used the fuzzy model to predict future values of the exchange rate and compare these to the actual values to validate this approach.

We show in Figure 3 the time series of exchange rates for 36 weeks of US dollar/MX peso from August 1999 to April 2000. We can notice from this figure the cyclical behavior of the time series over this short period of time.

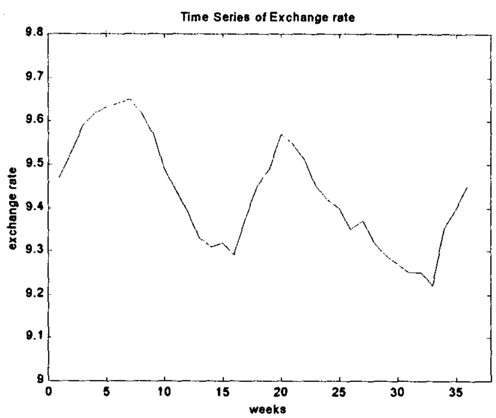

Figure 3 Time series of exchange rates US/Mexico.

We used the Fuzzy Logic Toolbox of MATLAB for fuzzy clustering of this data, and then for implementing the fuzzy rule base using the recognized clusters. In this case, five rules of the form shown in Equation 3 were used. We show in Figure 4 the general architecture of the fuzzy system. In this case, the Mamdani fuzzy reasoning procedure was used due to its simplicity.

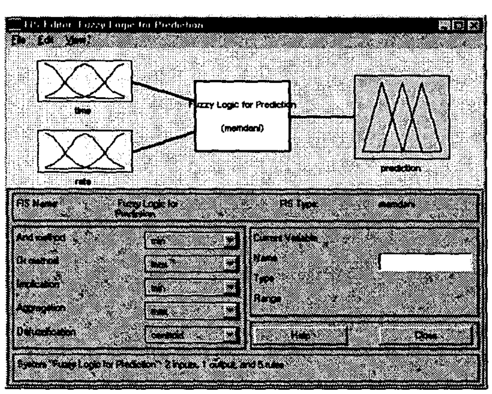

Figure 4 General Architecture of the fuzzy system.

We also show in Figure 5 how we can use this fuzzy system to predict a particular value of the exchange rate base on the actual value of the exchange rate and the value of future time. We validated these results against the real values that occurred during the following four weeks.

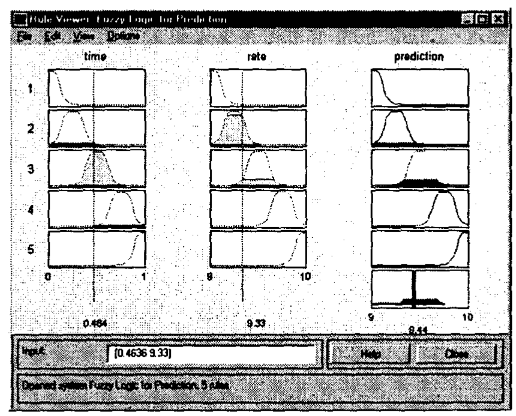

Figure 5 Fuzzy Prediction of the exchange rate.

4. Fuzzy Estimation of the Fractal Dimension

The fractal dimension of a geometrical object is a crisp numerical value measuring the geometrical complexity of the object. However, in practice it is difficult to assign a unique numerical value to an object. It is more appropriate to assign a range of numerical values in which there exists a membership degree for this object. For this reason, we will assign to an object O a fuzzy set μ_o, which measures the membership degree for that object.

Lets consider that the object O is in the plane xy, then a suitable membership function is a generalized bell function:

$$\mu_o = 1 / [1 + |(x-c) / a|^{2b}] \qquad (4)$$

where a, b and c are the parameters of the membership function. Of course other types of membership functions could be used depending on the characteristics of the application.

By using the concept of a fuzzy set we are in fact generalizing the mathematical concept of the fractal dimension. In fact, our definition of the fuzzy fractal dimension is as follows.

Definition 1: Let O be an arbitrary geometrical object in the plane xy. Then the fuzzy fractal dimension is the pair: (d_o, μ_o) where d_o is the numerical value of the fractal dimension calculated by the box counting algorithm, and μ_o is the membership function for the object.

With this new definition we can account for the uncertainty in the estimation of the fractal dimension of an object. Also, this new definition enables easier pattern recognition for objects, because it is not necessary to match an exact numerical value to recognize a particular object.

5. Fuzzy Fractal Approach for Time Series Analysis and Prediction

Let us consider now the problem of time series analysis and prediction. Let $y_1, y_2, \ldots, y_n$ be an arbitrary time series. If we want to be able to forecast this time series, we need to analyze the data and extract the trends and periodicities of the series. Assuming that the time series can be clustered into n objects $O_1, O_2, \ldots, O_n$ as shown in Figure 6, then we can build a fuzzy rule base as in Section 3 of this paper. However, we now also want to consider the geometrical complexity of the objects $O_1, O_2, \ldots, O_n$ as measured by their fractal dimensions $d_{o1}, d_{o2}, \ldots, d_{on}$ respectively. Then a fuzzy rule base for time series prediction can be expressed as follows.

If dim is d_{o1} and pos is x_1 then prediction is O_1
If dim is d_{o2} and pos is x_2 then prediction is O_2
... ... (5)
If dim is d_{on} and pos is x_n then prediction is O_n

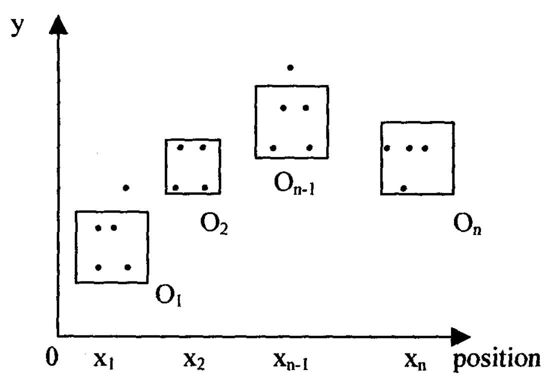

Figure 6 Fuzzy clustering of the time series.

In this case, we need to define membership functions for the fractal dimension, position, and for the geometrical objects. This fuzzy rule base can be used with the Mamdani reasoning method, and center of gravity as defuzzification method. However, it is also possible to use a Sugeno type fuzzy system in which the consequents can be linear functions. This is illustrated in Equation (6).

If dim is d_{o1} and pos is x_1 then $y=a_1x_1+b_1d_{o1}+c_1$
If dim is d_{o2} and pos is x_2 then $y=a_2x_2+b_2d_{o2}+c_2$
... ... (6)
If dim is d_{on} and pos is x_n then $y=a_nx_n+b_nd_{on}+c_n$

In this case, we can use a neuro-fuzzy approach for adapting the parameters of the fuzzy system using real data of the problem. We can use, for example, an ANFIS approach [5] to learn from real data the best values for the coefficients of the linear functions and for the membership functions.

We show in Figure 7 an implementation of the Mamdani type fuzzy system in the Fuzzy Logic Toolbox of MATLAB for the time series of exchange rates of US dollars/MX pesos. In this figure, we can see the non-linear surface for the fuzzy inference system of prediction.

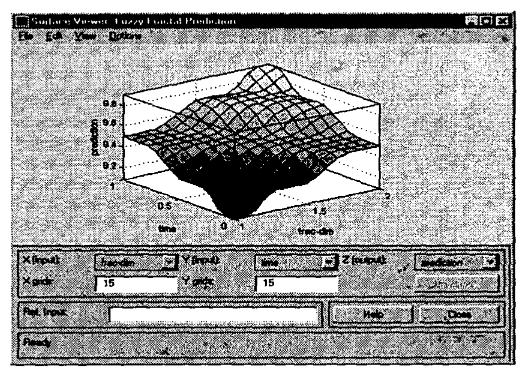

Figure 7 Non-Linear Surface for time series.

We also show in Figure 8 the implementation of the Sugeno type fuzzy system for the same time series. In this figure, we can see the non-linear surface for the Sugeno fuzzy system for time series prediction. This surface represents the fuzzy model for the problem of predicting the exchange rate of the US dollar/MX peso.

Finally, we show in Figure 9 the fuzzy reasoning for prediction for particular values of the fractal dimension and time.

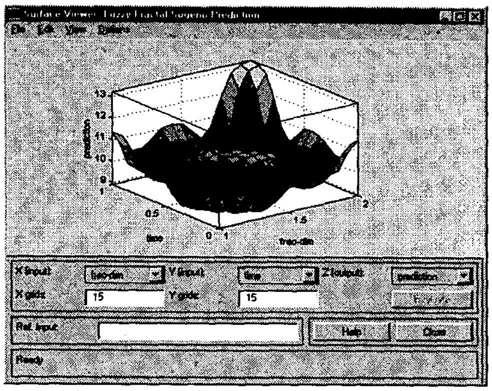

Figure 8 Surface for Sugeno type fuzzy system.

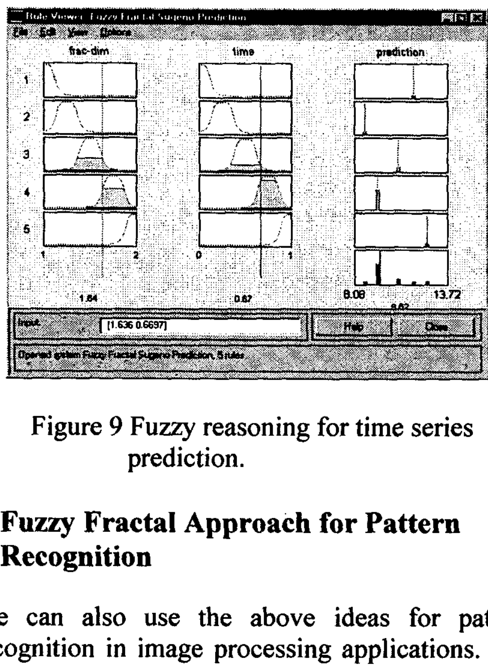

Figure 9 Fuzzy reasoning for time series prediction.

6. Fuzzy Fractal Approach for Pattern Recognition

We can also use the above ideas for pattern recognition in image processing applications. The method is very similar to the one for time series prediction, the only difference is that the data is not directly related to time. For pattern recognition only real geometrical variables are used. In this case, we also consider n objects $O_1, O_2, \ldots,$ On with n corresponding cluster centers (x_i, y_i), $i=1,\ldots,n$. Then the fuzzy rule base can be stated as follows.

$$\begin{array}{l} \text{If dimension is } d_{o1} \text{ then Object is } O_1 \\ \text{If dimension is } d_{o2} \text{ then Object is } O_2 \\ \ldots \qquad \ldots \qquad \ldots \\ \text{If dimension is } d_{o2} \text{ then Object is } O_n \end{array} \qquad (7)$$

To completely define this fuzzy system for pattern recognition, we will need to define the membership functions for the fractal dimensions and the objects. The method for calculating the fractal dimension is the same as before and for fuzzy reasoning we can use Mamdani or Sugeno type.

7. Conclusions

We have presented a new approach for time series analysis and pattern recognition combining fuzzy logic and fractal theory. With the new approach, we can always build a set of fuzzy rules using the fractal dimension of the objects to solve the problem of forecasting or recognition. We have very good results in predicting the exchange rate of the US dollar/MX peso with this new approach.

References:

[1] Bezdek, J.C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press.

[2] Castillo, O. and Melin, P. (1994). Developing a New Method for the Identification of Microorganisms for the Food Industry using the Fractal Dimension, *Journal of Fractals*, 2, No. 3, pp. 457-460.

[3] Castillo, O. and Melin, P. (1998). "A new Fuzzy-Fractal-Genetic Method for Automated Mathematical Modelling and Simulation of Robotic Dynamic Systems", Proceedings of FUZZ'98, IEEE Press, Alaska, USA, Vol. 2, pp. 1182-1187.

[4] Castillo, O. and Melin, P. (1999). "A new Fuzzy Inference System for Reasoning with Multiple Differential Equations for Modelling Complex Dynamical Systems", IOS Press, Vienna Austria, pp.224-229.

[5] Jang, J.R., Sun, C.T. and Mizutani E. (1997). Neuro-Fuzzy and Soft Computing, Prentice Hall.

[6] Mandelbrot, B. (1987). "The Fractal Geometry of Nature", W.H. Freeman and Company.

[7] Melin, P. and Castillo, O. (1998). "An Adaptive Model-Based Neuro-Fuzzy-Fractal Controller for Biochemical Reactors in the Food Industry", Proceedings of IJCNN'98, IEEE Computer Society Press, Alaska, USA, Vol. 1, pp. 106-111.

[8] Yager, R. and Filev, D. (1994). Generation of Fuzzy Rules by Mountain Clustering, Intelligent and Fuzzy Systems, Vol. 2, No. 3, pp. 209-219.

[9] Zadeh, L. A. (1975). "The Concept of a Linguistic Variable and its Application to Approximate Reasoning", *Information Sciences*, 8, pp. 43-80.