Proceedings of the V Brazilian Conference on Neural Networks - V Congresso Brasileiro de Redes Neurais
pp. 211–216, April 2–5, 2001 - Rio de Janeiro - RJ - Brazil

CALIBRA for Satellite Image

Aldrovando Luís Azeredo Araújo[1], Maria Bernadete Zanusso[2]
Katiuscia Alexandra de Castro[3]
Newton Graça Reis de Carvalho[4]
[1]Univ. Federal de Santa Catarina, Depto. de Matemática-Florianópolis-SC, Cep. 88040-900
[2]Universidade Federal de Mato Grosso do Sul, Depto. de Computação e Estatística(DCT)
Campo Grande-MS-Brasil, Cidade Universitária
Cep. 79070-900, Fone/Fax(0XX67)-7875450
[3,4]Universidade Federal de Mato Grosso do Sul-DCT
E-mails: azeredo@mtm.ufsc.br, mzanusso@dct.ufms.br, kac@dct.ufms.br

Abstract

This paper presents a system for callibration of the involved parameters in the classification and characterization process of fuzzy concepts that use t-norm differentiable and related membership families for the definition of characterization functions. An empirical evaluation using a satellite image database showed the comparability with the classical methods of classification.

1. Introduction

Since the theory of fuzzy sets was published by Zadeh[1] it has seen innumerable and varied applications in the many different areas of knowledge. It also reached a complete mathematical characterization of the fuzzy intersection and union through conjunctive (t-norm) and disjunctive (t-conorm) operators, respectively. In the fuzzy sets applications, the problem of choosing the membership functions and operators has been the theme of many research papers in the area.

The objective of this work is to present the CALIBRA system for satellite image. It uses the differentiable t-norms and t-conorms proposed in Zanusso[2] and Zanusso[3] for realize the conjunctions and disjunctions of fuzzy sets. The system got this name because by minimizing the mean squared error, it systematizes the parameter's callibration of the membership functions that represent elementary concepts over each attribute of the object to be characterized or classified. Its output is a characterizing expression of functions that represent fuzzy concepts obtained by conjunctions of the elementary concepts and disjunctions of these conjunctions. One expects to functions that will represent the classes of soil utilization well and that the system will attain a good classification.

In the second section of the paper, a summary of the t-norms and membership functions families is presented. The third section describe the satellite image database. The fourth presents a brief description of the system and the last section the preliminary training results.

2. T-Norms on Membership Functions Space

In this section we present the family of membership functions to represent fuzzy sets. This family is part of an space where the conjunctive and disjunctive operators will act on it. Differentiability and a conjecture on the relationship between the family of operators and the family of membership functions was also established in Zanusso[2] and Zanusso[3].

2.1. T-norms and t-conorms families

It seems that the first time that triangular norms and conorms appeared were with Schweizer and Sklar[4]. He established the basic definitions and the fundamental theorems. Based on these, as can be seen in Klir[5], the theory of fuzzy sets reached a complete characterization of fuzzy intersection and union by means of conjunctive (t-norm) and disjunctive(t-conorm) operators, respectively. Many authors present different t-norms to realize these operations but most of them depend on maximum and minimum.

Definition 1 (T-Norms Family) *By definition, if* $a, b \in (-1, 1]$

$$U_c^{\alpha,\beta}(a,b) \stackrel{\text{def}}{=} f_{\alpha,\beta}^{-1}\left(f_{\alpha,\beta}(a) + f_{\alpha,\beta}(b)\right), \alpha, \beta > 0, \quad (1)$$

$$a = 1 \Leftrightarrow U_c^{\alpha,\beta}(a,b) = b$$

$$a = -1 \Leftrightarrow U_c^{\alpha,\beta}(a,b) = -1$$

where $f_{\alpha,\beta}\colon (-1,1] \to [0,+\infty)$ *is defined by*

$$f_{\alpha,\beta}(x) = \alpha\left(\frac{1-x}{1+x}\right)^{\beta} \quad (2)$$

Definition 2 (T-Conorms Family) *By definition, if* $a, b \in [-1, 1)$

$$U_d^{\alpha,\beta}(a,b) \stackrel{\text{def}}{=} g_{\alpha,\beta}^{-1}\left(g_{\alpha,\beta}(a) + g_{\alpha,\beta}(b)\right), \alpha, \beta > 0, \quad (3)$$

$$a = 1 \Leftrightarrow U_d^{\alpha,\beta}(a,b) = 1$$

$$a = -1 \Leftrightarrow U_d^{\alpha,\beta}(a,b) = b$$

where $g_{\alpha,\beta}: [-1,1) \rightarrow [0,+\infty)$ *is defined by*

$$g_{\alpha,\beta}(x) = \alpha \left(\frac{1+x}{1-x}\right)^{\beta} \quad (4)$$

2.2. The family of Membership functions

The following family of membership functions to represent fuzzy sets was proposed

$$\mu_{p,k}(x) = \frac{x^k - p^k}{x^k + p^k} \quad (5)$$

where $p > 0$ e $k \neq 0$. These functions model *elementary concepts* like, *tall people*, *heavy people*, etc, for attributes height and weight, respectively.

Figure 1 shows a membership function of a concept, or class, G defined only over an attribute X, where $[I, S]$ is the interval of the variation of X. It is the conjunction(intersection) of two elementary concepts Z_{Xi} and Z_{Xs} as shown in Figure 2 for $I = 3$ and $S = 13$.

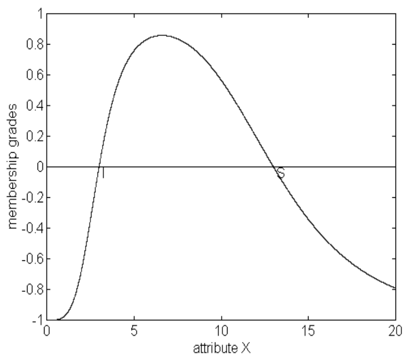

Figure 1: The membership of G over the X is given by Z_G

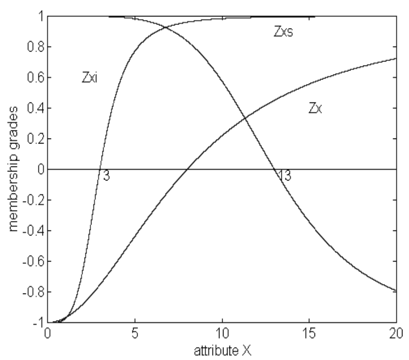

Figure 2: Elementary concept Z_{Xi} and Z_{Xs} which have Z_G as a conjunction

3. The Satellite Image Database

The satellite image database was taken from the following address "UCI Repository of Machine Learning Database and Domain Theories": ftp:ics.uci.edu:pub/machine-learning-databases. It was in use in the European Statlog project, which involves comparing the performances of machine learning, statistical, and neural network algorithms on data sets from real-world areas. This database was generated from Landsat Multi-Spectral Scanner image data. The original data for this database was generated from data purchased from NASA by the Australian Centre for Remote Sensing.

The Landsat satellite data is one of many sources of information available for a scene. This data satisfies the important requirements of being numerical and at a single resolution, and standard maximum-likelihood classification performs very well. Consequently, for this data, it should be interesting to compare the performance of other methods against the statistical approach.

One frame of Landsat MSS imagery consists of four digital images of the same scene in different spectral bands. Two of these are in the visible region(corresponding approximately to green and red regions of the visible spectrum) and two are in the (near) infra-red. Each pixel is a 8-bit binary word, with 0 corresponding to black and 255 to white. The spatial resolution of a pixel is about 80mx80m.

This database is a sub-area of a scene, consisting of 82 x 100 pixels. The binary values were converted to their present ASCII form and the classification for each pixel was performed on the basis of an actual site visit by Ms. Karen Hall, when working at the Centre for Remote Sensing at the University of New South Wales, Australia.

Each line of data corresponds to a 3x3 square neighborhood of pixels completely contained within the 82x100 sub-area. Each line contains the pixel values in the four spectral bands of each of the 9 pixels in the 3x3 neighborhood and a number indicating the classification label of the central pixel. The aim is to predict this classification, given the multi-spectral values.

The database contains thus 6435 patterns with 36 attributes (4 spectral bands x 9 pixels in neighborhood) plus the class label. There are six classes of soil utilization: red soil, cotton crop, grey soil, damp grey soil, soil with vegetation stubble and very damp grey soil. The attributes are numerical, in the range 0 to 255 (8 bits). The data was given in random order and certain lines of data have been removed so it cannot be reconstructed the original image from this database. In the work it was used only the four attributes for the central pixel. This avoids the problem which arises when a 3x3 neighborhood straddles a boundary. In this preliminary version of the system only three classes were considered.

4. CALIBRA for Fuzzy Concepts

Classifiers neural networks have been developed and are composed by logical neurons that realize AND/OR operations over the input that are membership grades. In particular the works of Brasil [6], Mitra[7], Fu [8], Halgamuge[9], Pedrycz[10] and Jang[11] make use of t-norms that depends on maximum and minimum that are not differentiable. The CALIBRA system reached comparable results with other classifiers when applied on the classification of species of iris flowers Zanusso[12]. Table 8 show this result. Now, it is applied to the satellite image database and in the subsequent sections it is showed preliminary results. In this section, a brief description of the system is presented.

4.1. Characterizing functions definitions

The membership function that characterize a class agree with the definitions of concept in the context of cognitive sciences when discussing the concept formation. How we can see in Uhr[13] a concept is a mental representation of a category of world objects given by a sequence of conjunctions or by a disjunction of conjunctions.

In first phase of training of the classifier it was chosen three classes: A (red soil), B (cotton crop) and C (grey soil). They were represented, or characterized, by functions of the four spectral values. For example, the characterizing function of class A is RA. It will be obtained using conjunctions and disjunctions of elementary fuzzy concepts defined over this four attributes. As the overlapping of classes was evident, that is the class A was mixed with B and C classes, it is considered $A = AA \cup AB \cup AC$ where AA would have the pixels that are only in A; AB would have the pixels that are classified in A by the experts, but are mixed with B too and, in the same way AC would have the pixels that are classified in A by the experts, but are mixed with C too.

Then

$$RA = f(x_1, x_2, x_3, x_4) = U_d(RAA, RAB, RAC) \tag{6}$$

where

$$\begin{aligned} RAA \quad = \quad & U_c(ZAA1i, ZAA2i, ZAA3i, ZAA4i, \\ & ZAA1s, ZAA2s, ZAA3s, ZAA4s). \end{aligned} \tag{7}$$

being that $ZAAxi$ is an inferior (i) elementary concept and $ZAAxs$ is an superior (s) elementary concept associated with the x-th attribute(spectral value), $x = 1, 2, 3$ and 4. The other two classes could be characterized in a similar way as A. The operators U_c and U_d are defined in terms of (1) and (2).

4.2. Fuzzy neural network

The calculation of the functions defined on the previous subsection can be easily put in a fuzzy neural network structure called CALIBRA system, as it is shown in Figura 3 and Figura 4.

4.3. The parameters

The parameters p's and k's is going to be adjusted by the training. To find the characterizing function RAA is necessary eight parameter p's and eight parameters k's. To RAB and RAC too.

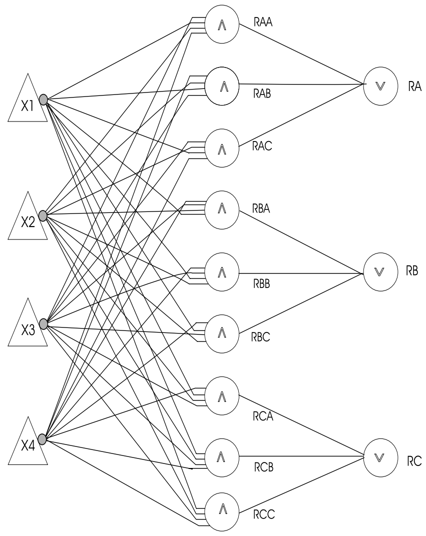

Figure 3: CALIBRA System

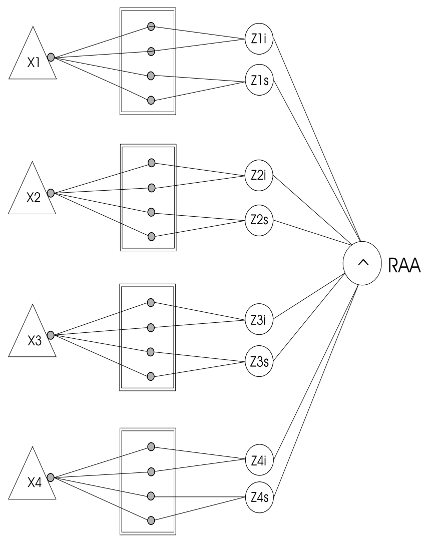

Figure 4: The class AA details

The parameters p's that define RA was put in a matrix pA and the parameters k's in a matrix kA.

$$pA = \begin{pmatrix} ...the\ parameters\ p's\ that\ define\ RAA... \\ ...the\ parameters\ p's\ that\ define\ RAB... \\ ...the\ parameters\ p's\ that\ define\ RAC... \end{pmatrix}$$

$$kA = \begin{pmatrix} ...the\ parameters\ k's\ that\ define\ RAA... \\ ...the\ parameters\ k's\ that\ define\ RAB... \\ ...the\ parameters\ k's\ that\ define\ RAC... \end{pmatrix}$$

Thus, to define RA is necessary to have 48 parameters, being 16 of RAA, 16 of RAB and 16 of RAC . To define the classes A, B and C is necessary to have 144 parameters.

The CALIBRA system has got this name because to minimizing the *mean squared error*, a callibration of the parameters of the characterizing functions RA, RB and RC has to take place for each context of application.

5. Training Results

In this section we presented the preliminary results of training with the satellite image database. The data set was divided 70% for training and 30% for test.

Table 1 shows some grey level of six test pixels on the four bands. We have drawn three pixels of each class. Table 2 shows the mean squared error for six epochs. We can see that the error to go down. Table 3 and Table 4 show the inferior and superior parameters p's and k's that define the elementary concepts that define the class A.

Table 1: Grey level of some test pixels

BANDS				
1	2	3	4	CLASS
72	115	120	102	A
53	79	96	78	A
47	34	114	126	B
52	43	92	92	B
86	104	108	85	C
87	103	105	86	C

Table 2: Error by epoch

EPOCH	ERROR
1	0.802935
2	0.488295
3	0.391975
4	0.32159
5	0.282739
6	0.279347

Table 3: The pA's parameters

inf.	40	52.2	74	45
sup.	54	26.7	74	103
inf.	39.9	81.3	91.7	93.4
sup.	81.7	130.4	131.7	65.4
inf.	83.4	139	69.96	31.1
sup.	77.5	138.8	147.6	102.4

Table 5 shows the membership grades of the six test pixels of Table 1 for the classes A, B and C given by the characterizing functions RA, RB and RC respectively. For example, the first pixel of Table 1, represented by the vector $(72, 115, 120, 102)$ has membership grade to class

Table 4: The kA's parameters

inf.	5.99	18.3	4.99	16
sup.	-5.9	-18.2	-5	-16
inf.	14.2	18.9	15.3	34.6
sup.	-30.4	-21.7	-9.8	-0.06
inf.	21.1	27.6	7.2	7.1
sup.	-44.2	-27.8	-32.4	-45

A equal to 0.256271, to class B equal to -0.9999 and to class C equal to -0.847963.

Table 5: Membership Grades

RA	RB	RC
0.256271	- 0.999909	- 0.847963
0.813607	- 0.941517	- 0.875795
- 0.999081	0.992846	- 0.989206
0.744403	0.935964	- 0.932550
- 0.921334	- 0.999491	0.222398
- 0.913810	- 0.999399	0.220758

Table 6 shows the ordinary (crisp) membership of the same pixels. It was considered that if the pixel had positive membership it would be given ordinary membership equal to one; the pixel that had negative membership would be given ordinary membership equal to zero. This criterion is still in study.

Table 6: Ordinary Membership

OA	OB	OC
1	0	0
1	0	0
0	1	0
1	1	0
0	0	1
0	0	1

Confusion matrix

One of the criteria to evaluate the performance of the classification methods is to use the comparison of the reference data with classification results to derive classification accuracies. For classified image (or a map), a confusion matrix (also called an error matrix or a contingency matrix) can be made. The Table 7 shows the this matrix. The major diagonal of the confusion matrix indicates the agreement between the two data sets.

The confusion matrix allows various accuracy indices to be derived. The Kappa coefficient K was given in Cohen[14]. It has been recommended as a suitable

Table 7: Confusion Matrix

	CLASSIFIER		
REFERENCE	A	B	C
A	97.3913	0	2.17391
B	9.95261	80.5687	0
C	1.2285	0	97.2973

Table 8: Classification Accuracy for Several Classifiers

Method of Classification	iris	satimage
K-Nearest Neighbor	92.00%	87.79%
Neural Network	95.33%	83.98%
C4.5 Decision Tree	92.67%	83.50%
Quadratic Bayes	95.33%	85.78%
Linear Bayes	97.33%	83.31%
CALIBRA	95.55%	91.75%

accuracy measure in thematic classification for representing the confusion matrix as a whole. It takes all the elements in the confusion matrix into consideration, rather than just the diagonal elements, which occurs with the calculation of overall classification accuracy.

The Kappa coefficient is defined by

$$k = \frac{p_0 - p_c}{1 - p_c} \quad (8)$$

where p_0 and p_c indicate the proportion of units which agree, and the proportion of units for expected chance agreement, respectively.

The Table 8 shows the comparison of results with others classifiers. The results of the other classifiers was published by Woods[15].

6. Conclusions

The results seems to indicate that the characterizing functions given in equations (6) and (7) represent each class well. On the other hand the CALIBRA has shown will a good performance of classification. The differentiability of t-norms and t-conorms can facilitate a more rigorous mathematical treatment of the system. One way to develop this system would be to create a *modus ponen* rule and other logical mechanisms for the inference building then an hybrid expert.

Aknowledge

The third and forth authors are grateful for to the CNPQ Scholarship Scientific Initiation and to UFMS by the support given for the development of the project.

References

[1] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[2] M. B. Zanusso. *Famílias de T-Normas Diferenciáveis, Funções de Pertinência Relacionadas e Aplicações*. PhD thesis, UFSC-EPS, Florianópolis-SC-Brasil, 1997.

[3] M. B. Zanusso. Differentiable t-norms and related membership functions families and their application. *IBERAMIA*, 2000.

[4] B. Schweizer and A. Sklar. Associative functions and statistical triangle inequalities. *Publ. Math. 8. Debrecem*, 1(1):169–186, 1961.

[5] G. J. Klir and b. Yuan. *Fuzzy Sets and Fuzzy Logic-Theory and Applications*. Prentice Hall, , New Jersey, 1995.

[6] L. M. Brasil and F. M. Azevedo. Fuzzyrulext: extraction technique of if/then rules for fuzzy neural nets. *Medical and Biological Engineering and Computing, Suplement Part I*, 1(1):354–373, 2000.

[7] S. Mitra and S. K. Pal. Fuzzy multi-layer perceptron, inferencing, and rule generation. *IEEE Transactions on Neural Networks*, 8(6):1338–1350, 1997.

[8] H. C. Fu and J. J. Shann. A fuzzy neural network for knowledge learning. *International Journal of Neural Systems*, 5(1):13–22, 1994.

[9] S. K. Halgamuge and M. Glesner. A fuzzy-neural approach of pattern classification with the generation of rules based on supervised learning. In *Proceedings of The Neuro-Nimes, Nanterre*, pages 167–173, 1992.

[10] W. Pedrycz and A. F. Rocha. Fuzzy-set based models of neurons and knowledge-based networks. *IEEE Transactions on Fuzzy Systems*, 1(4):254–266, 1993.

[11] J. S. R. Jang. Fuzzy modeling using generalized neural networks and kalman filter algorithm. In *Proceedings of The Ninth National Conference on Artificial Intelligence(AAAI-91),*, pages 762–767, 1991.

[12] M. B. Zanusso and A. L. A. Araújo. Calibra for fuzzy concepts. In *Vth Brasilian Symposium on Neural Networks. Belo Horizonte, Brasil, Dec.*, pages 130–134, 1998.

[13] L. Uhr. *Pattern Recognition Learning and Thought, Computers-Programed: Models of Higher Mental Processes*. Prentice-Hall, New Jersey, 1973.

[14] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 1(20):37–46, 1960.

[15] K. Woods and W. P. Kegelmeyer. Combination of multiple classifiers using local accuracy estimates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4):405–410, 1997.