Face Recognition with Choquet Integral in Modular Neural Networks

Gabriela E. Martínez, Patricia Melin, Olivia D. Mendoza and Oscar Castillo

Abstract In this chapter a new method for response integration, based on Choquet Integral is presented. A type-1 fuzzy system for edge detections based in Sobel and Morphological gradient is used, which is a pre-processing applied to the training data for better performance in the modular neural network. The Choquet integral is used how method to integrate the outputs of the modules of the modular neural networks (MNN). A database of faces was used to perform the pre-processing, the training, and the combination of information sources of the MNN.

1 Introduction

An integration method is a mechanism which takes as input a number n of data and combines them to result in a value representative of the information, methods exist which combine information from different sources which can be aggregation operators as arithmetic mean, geometric mean, OWA [1], and so on., In a modular neural network (MNN) is common to use some methods like fuzzy logic Type 1 and Type 2 [2–4], the fuzzy Sugeno integral [5], Interval Type-2 Fuzzy Logic Sugeno Integral [6], a probabilistic sum integrator [7], A Bayesian learning method [8], among others.

The Choquet integral is an aggregation operator which has been successfully used in various applications [9–11], but has not been used as a method of integrating a modular neural network, and in this chapter is proposed for achieving this goal 12.

G. E. Martínez · O. D. Mendoza
University of Baja California, Tijuana, Mexico

P. Melin (✉) · O. Castillo
Tijuana Institute of Technology, Tijuana, Mexico
e-mail: pmelin@tectijuana.mx

O. Castillo et al. (eds.), *Recent Advances on Hybrid Approaches for Designing Intelligent Systems*, Studies in Computational Intelligence 547,
DOI: 10.1007/978-3-319-05170-3_30,

This chapter is organized as follows: Sect. 2 shows the concepts of Fuzzy Measures and Choquet integral which is the technique that was applied for the combination of the several information sources. Section 3 presents Edge detection based in Sobel and Morphological gradient with type-1 fuzzy system. Section 4 shows the modular neural network proposed and in the Sect. 5 the simulation results are shown. Finally, Sect. 6 shows the conclusions.

2 Fuzzy Measures and Choquet Integral

Sugeno first defined the concept of "fuzzy measure and fuzzy integral" in 1974 [13]. A fuzzy measure is a nonnegative function monotone of values defined in "classical sets". Currently, when referring to this topic, the term "fuzzy measures" has been replaced by the term "monotonic measures", "non-additive measures" or "generalized measures" [14–16].When fuzzy measures are defined on fuzzy sets, we speak of fuzzified measures monotonous [16].

2.1 Fuzzy Measures

A fuzzy measure μ with respect to the dataset X, must satisfy the following conditions:

1. $$\mu(X) = 1;\ \mu(\emptyset) = 0$$

2. $$Si\ A \subset B,\ then\ \mu(A) \leq \mu(B)$$

In condition 2 A and B are subsets of X.

A fuzzy measure is a Sugeno measure or λ-fuzzy, if it satisfies the condition (1) of addition for some $\lambda > -1$.

$$\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\ \mu(B) \tag{1}$$

where λ can be calculated with (2):

$$f(\lambda) = \left\{ \prod_{i=1}^{n} (1 + M_i(x_i)\lambda) \right\} - (1 + \lambda) \tag{2}$$

The value of the parameter λ is determined by the conditions of the theorem 1.

Theorem 1 *Let* $\mu(\{x\}) < 1$ *for each* $x \in X$ *and let* $\mu(\{x\}) > 0$ *for at least two elements of X. Then* (2) *determines a unique parameter* λ *in the following way*:

If $\sum_{x \in X} \mu(\{x\}) < 1,$ *then* λ *is in the interval* $(0, \infty)$.

If $\sum_{x \in X} \mu(\{x\}) = 0,$ *then* $\lambda = 0$; *That is the unique root of the equation.*

If $\sum_{x \in X} \mu(\{x\}) > 1,$ *then* λ *se is in the interval* $(-1, 0)$.

The method to calculate Sugeno measures, carries out the calculation in a recursive way, using (3) and (4).

$$\mu(A_1) = \mu(M_i) \tag{3}$$

$$\mu(A_i) = \mu\big(A_{(i-1)}\big) + \mu(M_i) + \big(\lambda\mu(M_i) * \mu\big(A_{(i-1)}\big)\big) \tag{4}$$

where $1 < M_i \leq \cdots \leq$ n, and the values of $\mu(x_i)$ correspond to the fuzzy densities determined by an expert.

To perform this procedure $\mu(M_i)$ should be permuted with respect to the descending order of their respective $\mu(A_i)$.

There are 2 types of Integral, the integral of Sugeno and Choquet Integral.

2.2 *Choquet Integral*

The Choquet integral can be calculated using (5) or an equivalent expression (6)

$$\begin{aligned} &Choquet = \sum_{i=1}^{n} \left\{ \left[A_i - A_{(i-1)} \right] * D_i \right\} \\ &\text{With } A_0 = 0 \end{aligned} \tag{5}$$

Or also

$$\begin{aligned} &Choquet = \sum_{i=1}^{n} A_i * \left\{ \left[D_i - D_{(i+1)} \right] \right\} \\ &With\ D_{(n+1)} = 0 \end{aligned} \tag{6}$$

where A_i represents the fuzzy measurement associated with data D_i.

2.2.1 Pseudocode of Choquet Integral

INPUT: Number of information sources n; information sources x_1, x_2, …, x_n; fuzzy densities of information sources M_1, M_2, …, $M_n \in (0,1)$.
OUTPUT: Choquet integral ($\sigma(x_1)$, $\sigma(x_2)$, …, $\sigma(x_n)$).
STEP 1: Calculate λ finding the root of the function (2).
STEP 2: Fuzzify variable x_i.

$$D_i = \{x,\ \mu_{Di}(x) | x \in X\},\ \mu D_i(x) \in [0,\ 1]$$

STEP 3: Reorder M_i with respect to D (x_i) in descending order
STEP 4: Calculate fuzzy measures for each data with (3), (4).
STEP 5: Calculate Choquet integral with (5) or (6).
STEP 6: OUTPUT Choquet.
STOP

3 Edge Detection

Edge detection can be defined as a method consisting of identifying changes that exist in the light intensity, which can be used to determine certain properties or characteristics of the objects in the image.

We used the database of the ORL [17] to perform the training of the modular neural network, which has images of 40 people with 10 samples of each individual. To each of the images was applied a pre-processing by making use of Sobel edge detector and morphological gradient with type 1 fuzzy logic system [18] in order to highlight features, some of the images can be displayed in Fig. 4a.

3.1 The Morphological Gradient

To perform the method of morphological gradient is calculated every one of the four gradients as commonly done in the traditional method using (7–11), see Fig. 1, however, the sum of the gradients is performed by a fuzzy system type 1 [17] as shown in Fig. 2, and the resulting image can be viewed in Fig. 4b (Fig. 3).

$$D1 = \sqrt{(z5 - z2)^2 + (z5 - z8)^2} \tag{7}$$

$$D2 = \sqrt{(z5 - z4)^2 + (z5 - z6)^2} \tag{8}$$

$$D3 = \sqrt{(z5 - z1)^2 + (z5 - z9)^2} \tag{9}$$

$$D4 = \sqrt{(z5 - z7)^2 + (z5 - z3)^2} \tag{10}$$

$$G = D1 + D2 + D3 + D4 \tag{11}$$

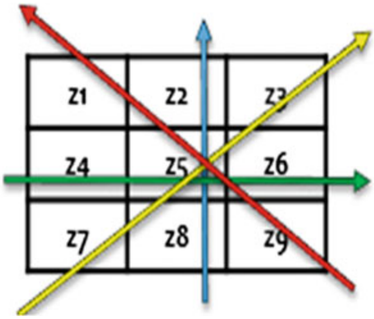

Fig. 1 Calculation of the gradient in the 4 directions

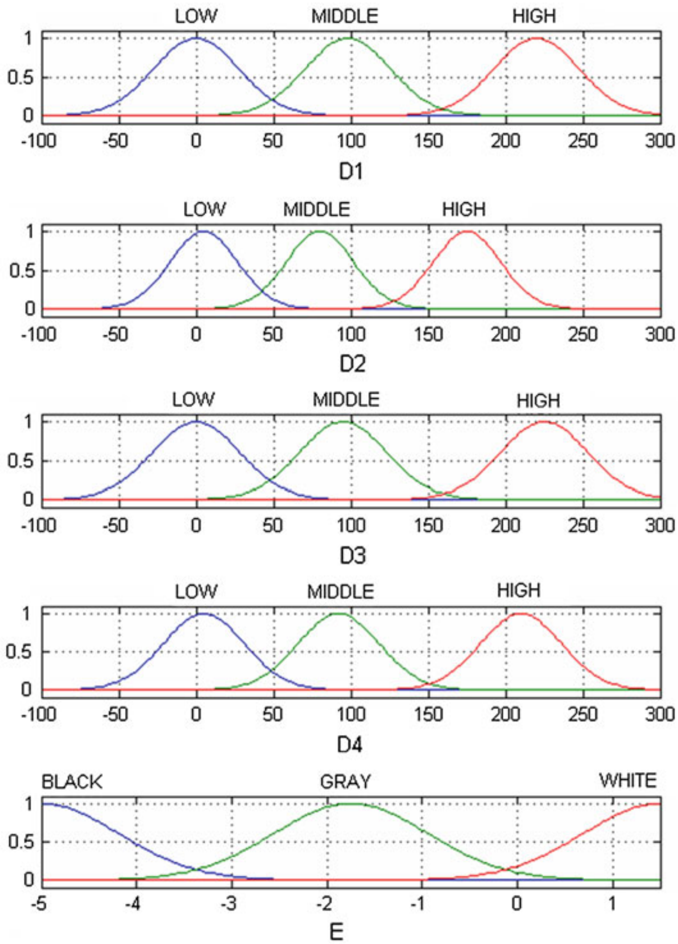

Fig. 2 Variables for the edge detector of morphological gradient type 1

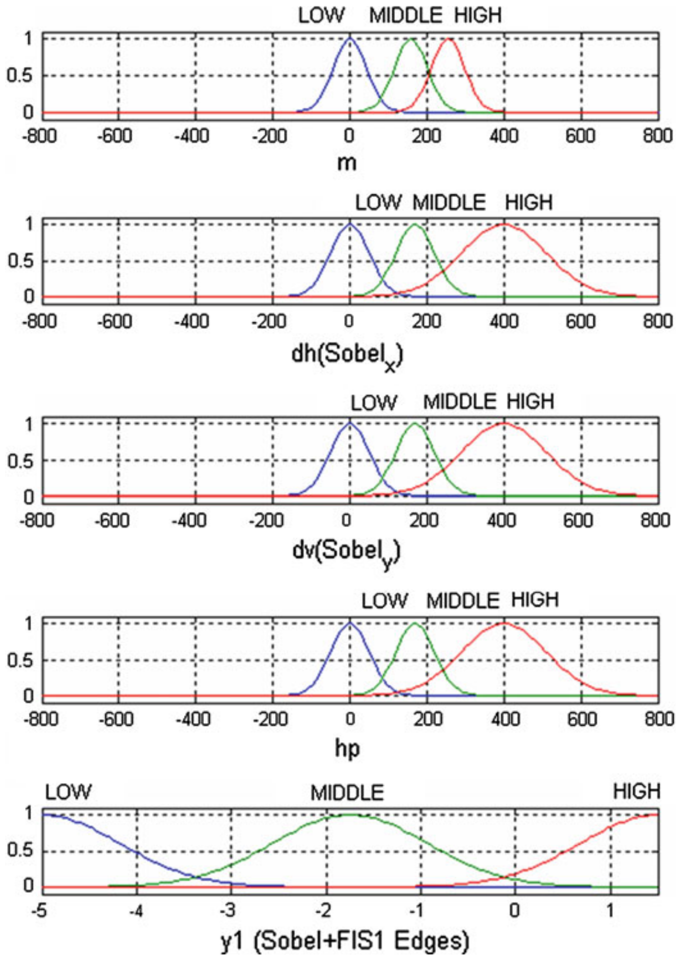

Fig. 3 Variables for the edge detector with the type-1 fuzzy Sobel

3.2 Sobel

The Sobel operator is applied to a digital image in gray scale, is a pair of 3 × 3 convolution masks, one estimating the gradient in the x-direction (columns) (12) and the other estimating the gradient in the y-direction (rows) (13) [19].

$$sobel_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{12}$$

$$sobel_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{13}$$

If we have $I_{m,n}$ as a matrix of m rows and r columns where the original image is stored, then g_x and g_y are matrices having the same dimensions as I, which at each element contain the horizontal and vertical derivative approximations and are calculated by (14) and (15) [19].

$$g_x = \sum_{i=1}^{i=3} \sum_{j=1}^{j=4} Sobel_{x,ij} * I_{r+i-2,c+j-2} \quad \begin{matrix} for = 1, 2, \ldots, m \\ for = 1, 2, \ldots, n \end{matrix} \tag{14}$$

$$g_y = \sum_{i=1}^{i=3} \sum_{j=1}^{j=4} Sobel_{y,i,j} * I_{r+i-2,c+j-2} \quad \begin{matrix} for = 1, 2, \ldots, m \\ for = 1, 2, \ldots, n \end{matrix} \tag{15}$$

In the Sobel method the gradient magnitude g is calculated by (16).

$$g = \sqrt{g_x^2 + g_y^2} \tag{16}$$

For the type-1 fuzzy inference system, 3 inputs can be used, 2 of them are the gradients with respect to the x-axis and y-axis, calculated with (14) and (15), which we call DH and DV, respectively. The third variable M is the image after the application of a low-pass filter hMF in (17); this filter allows to detect image pixels belonging to regions of the input were the mean gray level is lower. These regions are proportionally affected more by noise, which is supposed to be uniformly distributed over the whole image [19].

$$hMF = \frac{1}{25} * \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{17}$$

After applying the edge detector type 1 with Sobel, the resulting image can be viewed in Fig. 4c.

4 Modular Neural Networks

Were trained a MNN of 3 modules with 80 % of the data of ORL. Each image was divided into 3 sections horizontal and each of which was used as training data in each of the modules, as shown in Fig. 5.

The integration of the modules of the MNN was made with the Choquet integral.

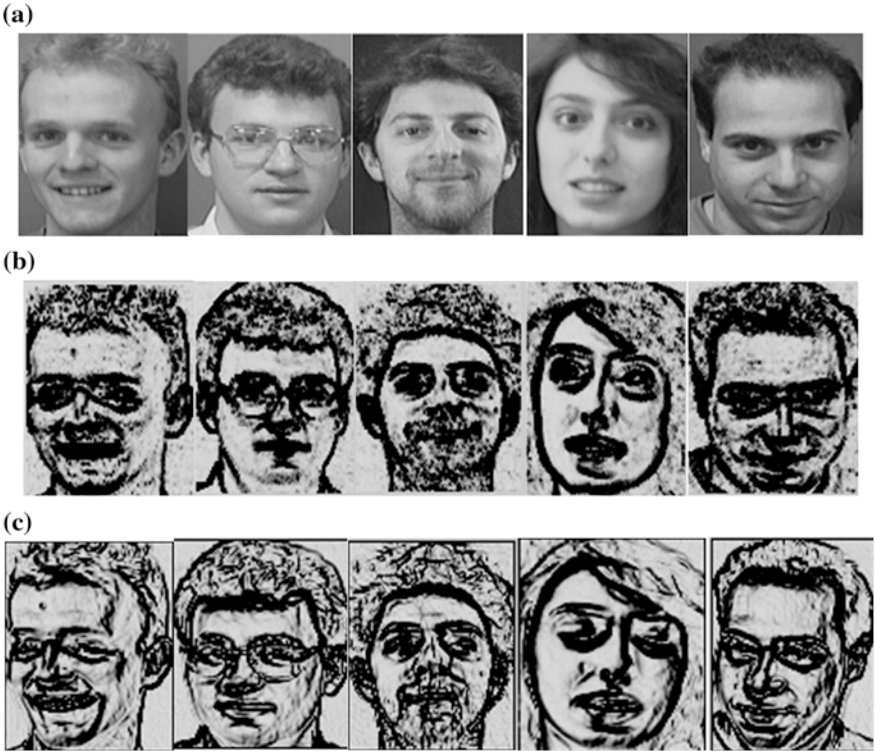

Fig. 4 **a** Original image, **b** image with morphological gradient, **c** image with Sobel

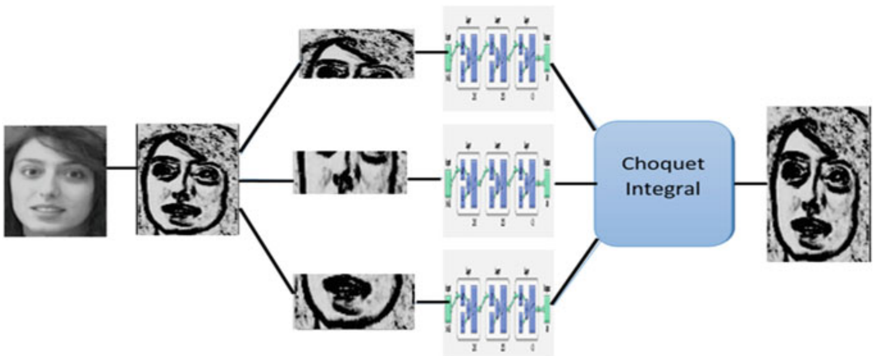

Fig. 5 Architecture proposal of the modular neural network

Table 1 Distribution of the training data

Training (%)	Validation (%)	Test (%)
70	15	15

Table 2 Procedure performed in the experiment

1.—Define the database of images
2.—Define the edge detector
3.—Detect the edges of each image
4.—Add the edges to the train set
5.—Divide the images in three parts
6.—Calculate the recognition rate using the k-fold cross validation method
(a) Calculate the indices for k folds
(b) Train the modular neural network k − 1 times for each training fold
(c) Simulate the modular neural network with the k test old
7.—Calculate the mean of rate for all the k-folds using Choquet integral

Training parameters were

Training method: gradient descendent with momentum and adaptive learning rate back-propagation (Traingdx).

Each module with two hidden layers [200 200].

Error goal: 0.00001

Epochs: 500

In Table 1 the distribution of the training data is shown.

4.1 The Experiment with a Modular Neural Network Recognition System and Choquet Integral for the Modules Fusion

The experiment consist on applying each evaluated edge detector to obtain a data set of same well know benchmark data base of images like ORL database of faces and then train a neural network to compare the recognition rate using the k-fold cross validation method [20], see Table 2.

5 Simulation Results

In the experiments we performed 27 tests in simulation of the trainings with each edge detectors making variations in fuzzy densities and performing the calculation

Table 3 Results of training with morphological gradient edge detector

Fuzzy densities			lambda	Training data			Test data		
				Morphological gradient and sobel			Morphological gradient		
				Mean rate	Mean rate	Std rate	Max rate	Mean rate	Std rate
0.1	0.1	0.1	−1.40E−16	1	0	1	0.8825	0.0456	0.9375
0.1	0.1	0.5	2.80E−16	1	0	1	0.8825	0.0456	0.9375
0.1	0.1	0.9	−0.5401	1	0	1	0.885	0.0379	0.9375
0.1	0.5	0.1	2.76E−16	1	0	1	0.8825	0.0381	0.925
0.1	0.5	0.5	−0.2918	1	0	1	0.8825	0.0381	0.925
0.1	0.5	0.9	−0.9107	1	0	1	0.8825	0.0338	0.925
0.1	0.9	0.1	−5.40E−01	1	0	1	0.88	0.036	0.925
0.1	0.9	0.5	−0.9107	1	0	1	0.8825	0.0371	0.925
0.1	0.9	0.9	−0.9891	1	0	1	0.8875	0.0319	0.925
0.5	0.1	0.1	2.76E−16	1	0	1	0.88	0.0527	0.9375
0.5	0.1	0.5	−0.2918	1	0	1	0.88	0.0527	0.9375
0.5	0.1	0.9	−0.9107	1	0	1	0.88	0.0489	0.9375
0.5	0.5	0.1	−0.2918	1	0	1	0.88	0.0456	0.925
0.5	0.5	0.5	−0.7639	1	0	1	0.88	0.0456	0.925
0.5	0.5	0.9	−0.9647	1	0	1	0.8775	0.0454	0.925
0.5	0.9	0.1	−0.9107	1	0	1	0.88	0.042	0.925
0.5	0.9	0.5	−0.9647	1	0	1	0.8825	0.0429	0.925
0.5	0.9	0.9	−0.9945	1	0	1	0.885	0.0428	0.925
0.9	0.1	0.1	−0.5401	1	0	1	0.875	0.05	0.9375
0.9	0.1	0.5	−0.9107	1	0	1	0.875	0.05	0.9375
0.9	0.1	0.9	−0.9891	1	0	1	0.875	0.0476	0.9375
0.9	0.5	0.1	−0.9107	1	0	1	0.875	0.0476	0.925
0.9	0.5	0.5	−0.9647	1	0	1	0.875	0.0476	0.925
0.9	0.5	0.9	−0.9945	1	0	1	0.875	0.0442	0.925
0.9	0.9	0.1	−0.9891	1	0	1	0.8775	0.0428	0.925
0.9	0.9	0.5	−0.9945	1	0	1	0.88	0.0438	0.925
0.9	0.9	0.9	−0.99	1	0	1	0.88	0.042	0.925
				1	**0**	**1**	**0.8800**	**0.0436**	**0.9292**

of the parameter λ with the bisection method. The results obtained with the morphological gradient are shown in Table 3 and with Sobel in Table 4.

In Table 5 the percentages of recognition of the Choquet integral with each edge detector are displayed. It can be noted that when using the morphological gradient it was obtained 94 % recognition of the MNN, while with Sobel a 93.125 % was obtained.

Table 4 Results of training with sobel edge detector

Fuzzy densities			lambda	Training data			Test data		
				Morphological gradient and Sobel			Sobel		
				Mean rate	Mean rate	Std rate	Max rate	Mean rate	Std rate
0.1	0.1	0.1	−1.40E−16	1	0	1	0.865	0.0056	0.875
0.1	0.1	0.5	2.80E−16	1	0	1	0.8625	0.0153	0.8875
0.1	0.1	0.9	−0.5401	1	0	1	0.8575	0.0143	0.875
0.1	0.5	0.1	2.76E−16	1	0	1	0.8675	0.0143	0.8875
0.1	0.5	0.5	−0.2918	1	0	1	0.8625	0.0234	0.9
0.1	0.5	0.9	−0.9107	1	0	1	0.86	0.0163	0.8875
0.1	0.9	0.1	−5.40E−01	1	0	1	0.8675	0.0143	0.8875
0.1	0.9	0.5	−0.9107	1	0	1	0.8675	0.0209	0.9
0.1	0.9	0.9	−0.9891	1	0	1	0.865	0.0137	0.8875
0.5	0.1	0.1	2.76E−16	1	0	1	0.86	0.0105	0.875
0.5	0.1	0.5	−0.2918	1	0	1	0.855	0.0143	0.875
0.5	0.1	0.9	−0.9107	1	0	1	0.8525	0.0105	0.8625
0.5	0.5	0.1	−0.2918	1	0	1	0.865	0.0105	0.875
0.5	0.5	0.5	−0.7639	1	0	1	0.8575	0.019	0.8875
0.5	0.5	0.9	−0.9647	1	0	1	0.8575	0.0112	0.875
0.5	0.9	0.1	−0.9107	1	0	1	0.87	0.0143	0.8875
0.5	0.9	0.5	−0.9647	1	0	1	0.87	0.0143	0.8875
0.5	0.9	0.9	−0.9945	1	0	1	0.865	0.0105	0.875
0.9	0.1	0.1	−0.5401	1	0	1	0.86	0.0105	0.875
0.9	0.1	0.5	−0.9107	1	0	1	0.8625	0.0088	0.875
0.9	0.1	0.9	−0.9891	1	0	1	0.86	0.0137	0.875
0.9	0.5	0.1	−0.9107	1	0	1	0.86	0.0056	0.8625
0.9	0.5	0.5	−0.9647	1	0	1	0.86	0.0105	0.875
0.9	0.5	0.9	−0.9945	1	0	1	0.86	0.0137	0.875
0.9	0.9	0.1	−0.9891	1	0	1	0.865	0.0137	0.8875
0.9	0.9	0.5	−0.9945	1	0	1	0.8675	0.0143	0.8875
0.9	0.9	0.9	−0.99	1	0	1	0.865	0.0105	0.875
				1	**0**	**1**	**0.8625**	**0.0131**	**0.8806**

Table 5 Results of training with Sobel edge detector

Method	% of recognition
Sobel	0.93125
Morphological gradient	0.94

6 Conclusions

The use of Choquet integral as a integration method answers of a modular neural network applied to face recognition has yielded favorable results when performing the aggregation process of the pre-processed images with the detectors of Sobel edges and morphological gradient, however it is still necessary to use a method that optimizes the value of the Sugeno measure assigned to each source of information because these were designated arbitrarily. Future work could be considering the optimization of the proposed method, as in [7, 21, 22].

Acknowledgments We thank the MyDCI program of the Division of Graduate Studies and Research, UABC, Tijuana Institute of Technology, and the financial support provided by our sponsor CONACYT contract grant number: 189350.

References

1. Zhou, L.-G., Chen b, H.-Y., Merigó, J.M., Anna, M.: Uncertain generalized aggregation operators. Expert Syst. Appl. **39**, 1105–1117 (2012)
2. Hidalgo, D., Castillo, O., Melin, P.: Type-1 and type-2 fuzzy inference systems as integration methods in modular neural networks for multimodal biometry and its optimization with genetic algorithms. Inf. Sci. **179**(13), 2123–2145 (2009)
3. Sánchez, D., Melin, P.: Modular neural network with fuzzy integration and its optimization using genetic algorithms for human recognition based on iris, ear and voice biometrics. In: Soft Computing for Recognition Based on Biometrics, pp. 85–102 (2010)
4. Sánchez, D., Melin, P., Castillo, O., Valdez, F.: Modular neural networks optimization with hierarchical genetic algorithms with fuzzy response integration for pattern recognition. MICAI **2**, 247–258 (2012)
5. Melin, P., Gonzalez C., Bravo, D., Gonzalez F., Martínez G.: Modular neural networks and fuzzy Sugeno integral for pattern recognition: the case of human face and fingerprint. In: Hybrid Intelligent Systems: Design and Analysis. Springer, Heidelberg, Germany (2007)
6. Melin, P., Mendoza, O., Castillo O.: Face recognition with an improved interval Type-2 fuzzy logic Sugeno integral and modular neural networks. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **41**(5) (2011)
7. Sánchez, D., Melin, P.: Optimization of modular granular neural networks using hierarchical genetic algorithms for human recognition using the ear biometric measure. Eng. Appl. Artif. Intell. **27**, 41–56 (2014)
8. Wang, P., Xu, L., Zhou, S.M., Fan, Z., Li, Y., Feng, S.: A novel Bayesian learning method for information aggregation in modular neural networks. Expert Syst. Appl. **37,** 1071–1074 (2010) (Elsevier)
9. Kwak, K.-C., Pedrycz, W.: Face recognition: a study in information fusion using fuzzy integral. Pattern Recogn. Lett. **26**, 719–733 (2005)
10. Timonin, M.: Robust optimization of the Choquet integral. Fuzzy Sets Syst. **213**, 27–46 (2013)
11. Yang, W., Chen, Z.: New aggregation operators based on the Choquet integral and 2-tuple linguistic information. Expert Syst. Appl. **39**, 2662–2668 (2012)
12. Meena, Y.K., Arya, K.V., Kala, R.: Classification using redundant mapping in modular neural networks. In: Second World Congress on Nature and Biologically Inspired Computing, Kitakyushu, Fukuoka, Japan, 15–17 Dec 2010
13. Sugeno, M.: Theory of fuzzy integrals and its applications. Thesis Doctoral, Tokyo Institute of Technology, Tokyo, Japan (1974)
14. Murofushi, T., Sugeno, M.: Fuzzy measures and fuzzy integrals. Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Yokohama, Japan (2000)
15. Song, J., Li, J.: Lebesgue theorems in non-additive measure theory. Fuzzy Sets Syst. **149**, 543–548 (2005)

16. Wang, Z., Klir, G.: Generalized Measure Theory. Springer, New York (2009)
17. Database ORL Face. Cambridge University Computer Laboratory. http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html. Nov 2012
18. Mendoza, O., Melin, P., Castillo, O., Castro, J.: Comparison of fuzzy edge detectors based on the image recognition rate as performance index calculated with neural networks. In: Soft Computing for Recognition Based on Biometrics. Studies in Computational Intelligence, vol. 312, pp. 389–399 (2010)
19. Mendoza, O., Melin, P., Licea, G.: A hybrid approach for image recognition combining type-2 fuzzy logic, modular neural networks and the Sugeno integral. Inf. Sci. Int. J. **179**(13), 2078–2101 (2009)
20. Mendoza, O., Melin, P.: Quantitative evaluation of fuzzy edge detectors applied to neural networks or image recognition. Advances in Research and Developments in Digital Systems (2011)
21. Sánchez, D., Melin, P.: Multi-objective hierarchical genetic algorithm for modular granular neural network optimization. In: Soft Computing Applications in Optimization, Control, and Recognition, pp. 157–185 (2013)
22. Sánchez, D., Melin, P., Castillo, O., Valdez, F.: Modular granular neural networks optimization with Multi-Objective Hierarchical Genetic Algorithm for human recognition based on iris biometric. In: IEEE Congress on Evolutionary Computation, pp. 772–778 (2013)