Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 10 (10): 848-851, 2010
ISSN 1812-5654

Arabic Vowels Fuzzy Neural Network Recognition

A. Taleb and A. Benyettou
SIgnal-IMage-PArole Laboratory, Department of Computer Science,
University of Sciences and Technology of Oran, Mohamed Boudiaf,
B.P. 1505, El Mnaouer 31000, Oran, Algeria

Abstract: In this study, We propose a fuzzy neural system containing inferred rules which are modelled separately by a three layer perceptron neural network giving the conclusion part according to the premise of the rule. Such a system is applied to different morphology words for Arabic vowels recognition as a two-dimensional fuzzy implication presented in the form of linguistic features values. The system has been implemented on a real-time mini-computer and is now operational, the results concerning a multi-speaker corpus of continuous speech are also promising.

Key words: Modern standard arabic, Linear Prediction Coding (LPC), fuzzy neural networks, speech recognition, Multi-Layer Perceptron (MLP)

INTRODUCTION

Suppose a two-dimensional system of fuzzy inferred rules:

$$\begin{cases} \text{if x is A then z is } C_1 \\ \text{and} \\ \text{f y is B then z is } C_2 \end{cases} \quad (1)$$

$C = C_1 \cap C_2$ is the consequent part of the reasoning. The case $C_1 = C_2 = C$ have been studied by Kumar and Jayati (1997) and Taleb and Benyettou (2003).

Our model gives a better adaptation to the pattern recognition since the intersection of the consequent parts of rules is taken as the conclusion part of the system. Thus, the linguistic connective and of our equations has more meaningful logical interpretation. We introduce the linguistic features related to F_1 and F_2, whose values are ones of the following fuzzy sets: small, medium, large as shown in Fig. 1.

Compared to the nature of our application, we take only space R^2, knowing that this theory adapts very well in space R^n, $n \geq 2$ (Lin and Lu, 1995).

The consequent part of the two-dimensional fuzzy implication represents the possibility of occurrences of the different classes: in the tip [ABCD] for $F_1 = F_2 =$ medium, C_i has a great possibility of occurrences in this tip that C_j.

The C_k class does not have any possibility of occurrence in the tip.

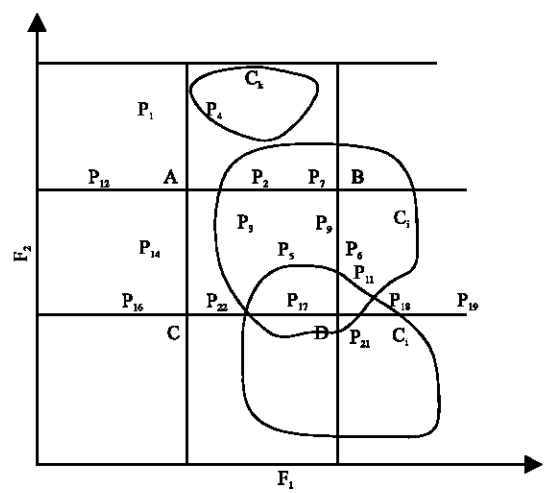

Fig. 1: Representation of a population distribution in the F_1-F_2 plan

Being an application of voice recognition, the formants F_1 and F_2 are represented by linguistic features such as F_1 is medium and F_2 is large, instead of the fixed values like F_1 = 400 Hz and F_2 = 1800 Hz. F_1 and F_2 are characterized by membership functions which return the borders between the fuzzy partitions as you can see it in Fig. 2.

In pattern recognition, on the training process, the extraction and selection of formant features deduced starting from 12 coefficients LPC are clustered in three classes relating to the Arabic vowels /æ/, /u/, /I/.

Corresponding Author: A. Taleb, Faculty of Science, University of Science and Technology of Oran, Mohammed Boudiaf USTO, P.O. Box B.P. 1505, El Menaour, Oran, Algeria Tel: +213770574242 Fax: +21341533496

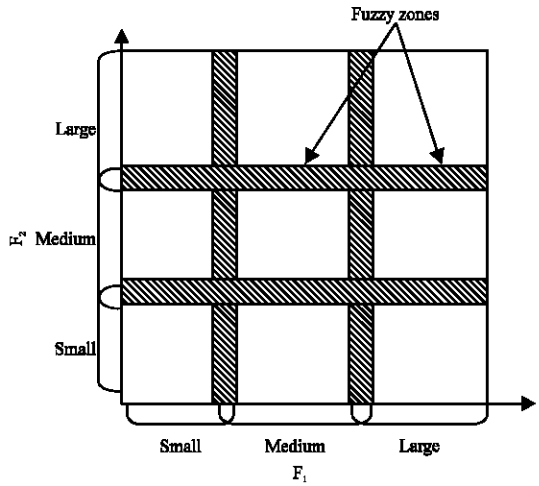

Fig. 2: Fuzzy partition of F_1-F_2 space

The decision functions are then determined by the suitable fuzzy If-Then rules.

Concerning classification process, first a set of predetermined features is extracted from the pattern then a set of If-Then rules determines the possibility of occurrence of each class in the feature space (Rahmani, 1998).

IMPLEMENTATION OF THE FUZZY RULES ON NEURAL NETWORKS

The fuzzy two-dimensional Eq. 1 is carried out through a two neural networks with standard retro-propagation or MLP's with three layers (Mesbahi and Benyettou, 2005). The input of neural network is the premise of the rule If-Then, which is represented by a fuzzy membership function. The output of the network is the consequent part of the rule, represented by a membership function, which determines the possibilities of occurrence of each class. The other parameters of the network are the same as the conventional MLP type. After the training of each network independently, we combine the output of the two networks by the intersection operator ∩ as shown in Fig. 3.

For the implementation of two fuzzy rules through two different neural networks, we use the linguistic features at the input of the networks with values: small, medium and large, with F_1 at the input of the first neural network and F_2 at the second network.

The field of values of F_1 varies from 200 to 900 Hz, that of F_2 varies from 500 to 2400 Hz. Therefore 'F_1 small' is not equivalent F_2 small and even F1 is large remains smaller than F_2 is small, which justifies the implementation of the fuzzy system on two neural networks.

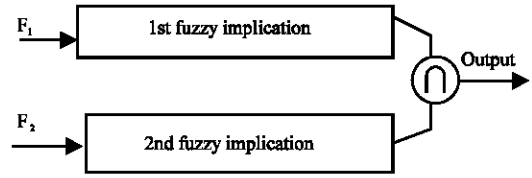

Fig. 3: Realization of two fuzzy rules through two neural networks

Table 1: Quantization of F_1 of the vowels /æ/, /u/, /I/

Interval	Concerned vowels	Small	Medium	Large
200≤F1<300	/i/	1.0	0.0	0.0
300≤F1<400	/u/, /i/	0.5	0.2	0.0
400≤F1<500	/æ/, /u/, /i/	0.2	1.0	0.0
500≤F1<600	/æ/, /u/, /i/	0.0	0.7	0.1
600≤F1<700	/æ/, /u/	0.0	0.1	0.4
700≤F1<800	/æ/	0.0	0.0	0.9
800≤F1<900	/æ/	0.0	0.0	1.0

Table 2: Quantization of F_2 of the vowels /æ/, /u/, /I/

Interval	Concerned vowels	Small	Medium	Large
500≤F2<710	/i/	1.0	0.0	0.0
710≤F2<920	/i/	0.7	0.0	0.0
920≤F2<1130	/æ/, /i/	0.4	0.2	0.0
1130≤F2<1340	/æ/, /i/	0.1	0.7	0.0
1340≤F2<1550	/æ/, /u/, /i/	0.0	1.0	0.0
1550≤F2<1760	/æ/, /u/, /i/	0.0	0.5	0.0
1760≤F2<1970	/æ/, /u/, /i/	0.0	0.2	0.3
1970≤F2<2180	/æ/, /u/	0.0	0.0	0.7
2180≤F2<2400	/æ/, /u/	0.0	0.0	1.0

Training phase: At the stage of the training, the fields of values of F_1 and F_2 are discredized. The discretization is quantified by values from 0 to 1. Each segment is labelled by generic elements. The notion of fuzzyfication is defined on the assignment of each generic element by a value which represents the degree of membership (Table 1, 2).

The membership function of the consequent part of the rule If-Then, represents the possibility of occurrence of each class in fuzzy partitioned space F_1-F_2.

EXPERIMENTAL RESULTS

We have proceeded to recording a corpus of 17 words and 3 vowels in Arabic language, pronounced by 16 speakers (10 men and 6 women), each word is uttered 3 times by each speaker. Each recorded word is in the form of a quantified and sampled signal, then safeguarded in a file *.wav. The speech waveform was sampled at Fe = 11025 Hz. A procedure of delimitation is used in order to remove silence to determine the beginning and the end of the word. To increase the amplitudes of the high frequencies, the signal is pre-accentuated, i.e., passed by a filter $1-\mu Z^{-1}$. This operation can simply made by evaluating:

$$s'(n) = s(n)-\mu s(n-1) \text{ for } n \geq 0 \text{ with } \mu = 0.94$$

Next the signal is convoluted by a Hamming window each 23 m sec (approximately 256 samples). A 6 m sec covering is carried out between two successive hamming windows in order not to lose information. Tweleve coefficients LPC are calculated in each window and are safeguarded in a file *.LPC. An algorithm of delimitation of the phonemes is used, then the extraction of formants F_1 and F_2 according to two methods: peak picking from spectral smoothing or polynomial root factorisation using the Z-transform (Taleb, 2000).

We represent in the F_1-F_2 plan, the delimitation of the vowel boundaries, after compression using vector quantification upon the extracted values form the LPC coefficients (Linde *et al*., 1980).

Table 3 presents the fuzzyfication desired of the neural network output while referring to the distribution of the vowels /æ/, /u/, /i/, in Fig. 4. The output linguistic features comprise the terms: Very High, High, Medium, Low, Very Low and Nil which are respectively assigned by values: 1, 0.8, 0.5, 0.3, 0.1 and 0.0.

In Fig. 4, only features High, Medium, Low and Nil are represented. This considered assignment is the defuzzyfication stage.

The reinforcement learning called sometimes learning with critic which is a slightly supervised training has been applied in the field of the reactive navigation of a mobile robot and can be used in order to adapt our system to make first a macro-classification dealing with each phonetic category (Dahmani and Benyttou, 2004a, b, 2005). The capacity to detect the acoustic features and

Table 3: Results of Fi. 4 assigned by linguistic features

	Conclusion of rules		
Premise of rules	/æ/	/u/	/i/
If F_1 is small then	Nil	Medium	High
If F_1 is medium then	High	Medium	Low
If F_1 is large then	Medium	Nil	Nil
If F_2 is small then	Nil	Nil	Medium
If F_2 is medium then	High	Low	Medium
If F_2 is large then	Medium	High	Nil
Recognition score (%)	92	84	80

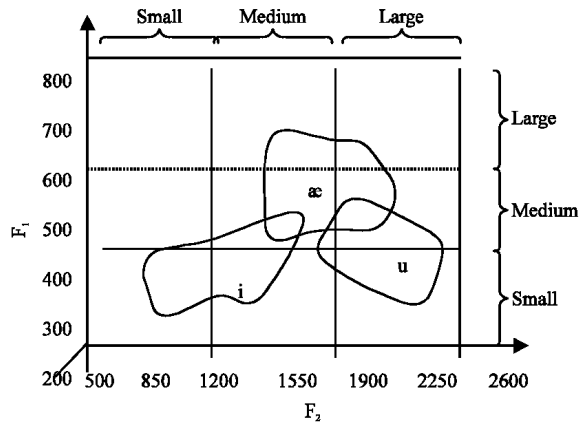

Fig. 4: Representation of Arabic vowels /æ/, /u/, /I/ in F1-F2 plan

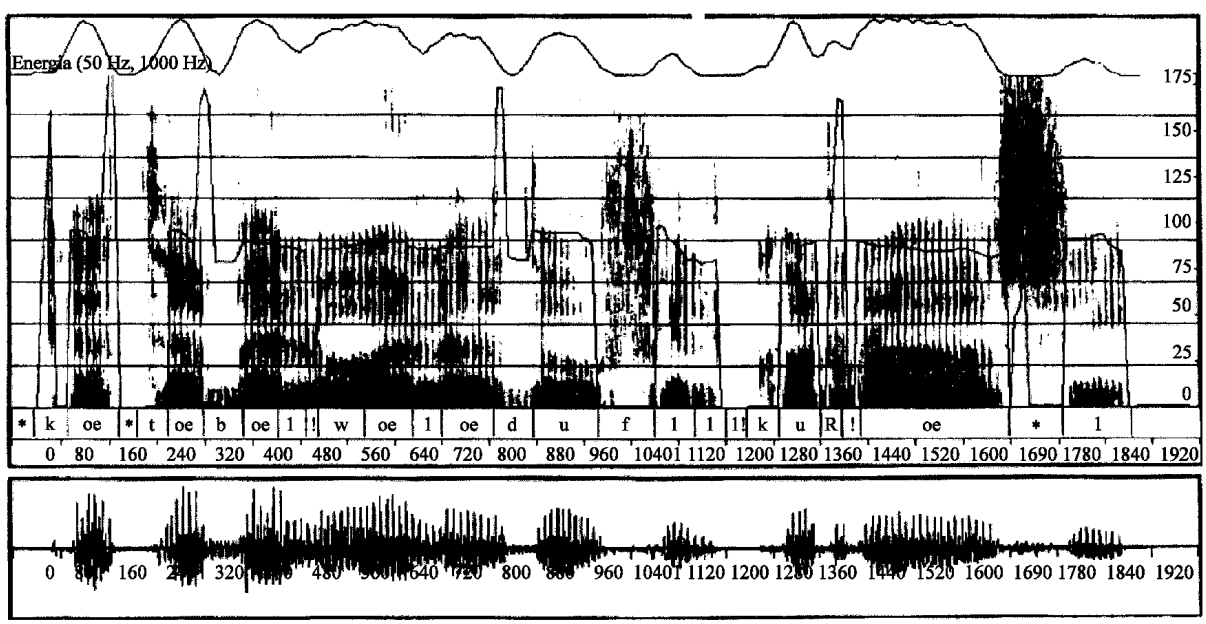

Fig. 5: The numerical spectrogram of the arabic sentence: koetoeboe-oel-woeloedu fi-oel-kurroe:si (translation of: the child has written on the notebook)

their independent temporal report of the temporal localisation may be used like in the adaptive temporal Radial Basis Function when we deal with the Arabic continuous speech recognition (Fig. 5) (Mesbahi and Benyettou, 2005). The integration of the Gamma neural Net to provide a good acoustic parameters for automatic speech recognition by taking into account the temporal aspect in the speech signal was considered by Louni and Taleb (2008).

CONCLUSION

The study presented here, actually being implemented for Arabic isolated words, may generalise to continuous voice recognition by associating a neural network for each class of consonants (fricatives, occlusives, sonorants). Combination of different formalisms, based on biological approaches on neural networks, enable us to tackle problems on the more advanced voice recognition (Neggaz and Benyettou, 2009) and is being investigated for the continuous Arabic speech recognition and will be presented in future work.

REFERENCES

Dahmani, Y. and A. Benyettou, 2004a. Parameter tuning of fuzzy subsets inertia influence in navigation case. J. Applied Sci., 4: 384-387.

Dahmani, Y. and A. Benyettou, 2004b. Fuzzy reinforcement rectilinear trajectory learning. J. Applied Sci., 4: 388-392.

Dahmani, Y. and A. Benyettou, 2005. Seek of an optimum way by Q-learning. J. Comput. Sci., 1: 28-30.

Kumar, S.R. and G. Jayati, 1997. Neuro fuzzy approach to pattern recognition. Neural Network, 10: 161-182.

Lin, C.T. and Y.C. Lu, 1995. A neural fuzzy system with linguistic teaching signals. IEEE Trans. Fuzzy Syst., 3: 169-189.

Linde, Y., A. Buzo and R.M. Gray, 1980. An algorithm for vector quantizer design. IEEE Trans. Commun., 28: 84-95.

Louni, A. and A. Taleb, 2008. Un codage neuro-prédictif pour lextraction des traits distinctifs appliqué à la reconnaissance des phonèmes arabes. Proceedings of the 10th Maghrebian Conference on Software Engineering and Artificial Intelligence, April 28-30, Oran, Algeria, pp: 132-138.

Mesbahi, L. and A. Benyettou, 2005. A new look to adaptive temporal RBF applied in speech recognition. J. Comput. Sci., 1: 1-6.

Neggaz, N. and A. Benyettou, 2009. Hybrid models based on biological approaches for speech recognition. Artificial Intel. Rev., 32: 45-57.

Rahmani, A., 1998. Automatic generation of optimal fuzzy expert systems by means of neuro-fuzzy approaches and genetic algorithms. Doctorat Thesis, University of Sidi Belabbes.

Taleb, A., 2000. Contribution des réseaux neuro-flous pour la reconnaissance de voyelles arabes. Master Thesis, Computer Science Dept., University of USTOran-MB.

Taleb, A. and A. Benyettou, 2003. A neural fuzzy system applied to arabic vowels identification. Proceedings of the 2nd International Workshop on Advanced Computation for Engineering Applications, Dec. 21-22, Cairo, Egypt.