International Journal of Mathematics Trends and Technology- Volume4 Issue 6 - July 2013

Ability of Evolutionary and Recurrent SOM model GA-RSOM in Phonemic Recognition Optimization

Mohamed Salah Salhi **;** Najet Arous; Noureddine Ellouze
Laboratory systems and signal processing (LSTS)
National engineering school of Tunis, BP 37,
The Belvedere, 1002 Tunis, Tunisia

Abstract— **The phoneme recognition aims to process a speech signal, characterized by a non-linearity with very high dynamics, allowing to perform various tasks on an information processing machine by an operator using orally address. This paper focuses on a proposed strategy, which implements an evolutionary recurrent self organizing map (SOM) model in phonemes recognition to improve their rates. It is a hybrid model (GA-RSOM) reflecting the approaches of K-means mobile centers, the evolutionary genetic algorithm (GA) principle and the recurrent temporal appearance of Kohonen map (RSOM) to be a powerful optimization tool for phonemic recognition, even in adverse environmental conditions.**

Keywords- The SOM model; the recurrent SOM; the K-means mobile centers; the genetic Algorithm GA; the hybrid model (GA-RSOM).

I. Introduction

Phoneme recognition is an approach of speech recognition using static entities with determined characteristics which can be checked, identified and used in various applications over voice communications. However, the identification of these features requires the right choice of the adequate tool for their recognition.

The phonemic recognition of speech is a discipline closely related to computing and pattern recognition. The first attempts to start in 1950 with the recognition of numbers and vowels. The appearance and ranges distribution of computers in the 1960s, even the advent of numerical methods give a new impetus for research in this area. The idea of integrating linguistic constraints has participated objective to improve phonemic recognition of speech in the 1970s. However, the recognition results remain dependent on the chosen tool. In 1982, the Finnish Teuvo Kohonen has made a self-organizing neural map said SOM stands for 'Self Organizing Map' inspired by the biology of the human brain, which was used as a selector features. With the various experiments and research over time, improvements and developments were made to this map giving birth for multiple versions. The phoneme recognition models have become increasingly reliable [1]. Our approach is the integration in the SOM maps the temporal aspect by recurrence and their hybridization with genetic algorithms AG.

Speech is a signal with very variable dynamic, which complicates the recognition state. A technique is highly relevant consists to decompose the signal into smaller atoms of the sound called phonemes that represent stationary states, spectrally distinguishable and separable, thus allowing to enhance the recognition rate. Several approaches and models were identified and served in this field along the last decade to improve the recognition rate, each of its way, its vision azimuth to his environment, to achieve what the biological nature do. These works are the subject of section 2. To overcome the shortcomings encountered recognition using statistical methods and probabilistic techniques in noisy conditions or in an insufficient amount of data for learning, we have adopted in the others sections a connectionist approach based on a dynamic evolutionary self organizing model in growth and including the time for phonemes recognition.

II. Phonemic Recognition Strategy by SOM

Based on a modeling network of biological inspiration, Teuvo Kohonen has produced a particular neural network map called SOM. This neural network consists of elementary processors set, designated by the neurons, connected to each other to exchange information. The implementation of a formal neurons number working in parallel and massively interconnected, giving them the learning and the decision capacity for the recognition [2], [3]. The activation function is generally nonlinear. Each function is appropriate for a well-defined application. It is said that the network consists of an enriched form of generalization after learning through a well-defined database.
The learning rule updates the neural weights in the vicinity of the activated neuron 'winner', by bringing them close to the input vector:

$$\Delta w_i = \gamma . h_{iv}(x(t) - w_i) \qquad (1)$$

γ is a learning ratio and h_{iv} is a neighborhood function, which decreases with the distance between units i and v on the map.

The nonlinearity of activation function gives opportunity to the map SOM network to be a universal tool of representation and recognition. In addition, this algorithm is unsupervised, and this specificity allows it to represent and be used in the recognition of a large volumes' input data, in contrast to supervised algorithms that can process only restricted and limited data volumes.

In 1999, A. Postigo Gardon, C. Ruiz Vázquez and Arruti Illarramendi have worked on the phonemic classification of the Spanish language in ' Spanish Phoneme Classification by Means of a Hierarchy of Kohonen Self-Organizing Maps '. Some results of Spanish phonemes classification, through self-organizing maps SOM of Kohonen, are presented. These results indicate that SOM can be very useful in the previous steps of continuous speech recognition, as it constitutes valuable assistance in the process of signal phonetic segmentation. This intermediate classification provides a new signal representation as the "phonetic". They found that by segmenting the signal into phonemes, the difference, between inputs and outputs of the recognition, is greatly reduced, thus simplifying the recognition task.

In 1991, Jari Kangas, of Computer & Computational Sciences Laboratory, Helsinki University of Technology, Espoo, Finland, in "Phoneme recognition using time-dependent versions of self-organizing maps," Acoustics, Speech, and Signal Processing, IEEE International Conference, has proposed two changes on the self-organizing map SOM for phonemic recognition. These two models of change are moving, unlike the original Kohonen algorithm, to the inclusion of time-dependent characteristics of the input signal. In the first modification, a time average of a SOM responses sequence is found, and this is recognized by another SOM. In the second modification, the input successive phonemes are concatenated together and will be recognized by the SOM. By comparing the results with those of a recognition system using the original SOM, it was found that could improve the isolated phonemes recognition from 10.4% error to 7.0% and 5.0% error for the temporal integration model and the concatenation model, respectively. The improvement of a full scale, where phonemes segments are to be located is 9.2% error to 8.2% and 7.6% of errors for the new methods, respectively.

In 2005, E. Gatt, J. Micallef, and E. Chilton, of Microelectronics Department, University of Malta Msida have presented at a conference conducted in Gammarth, Tunisia, the use of physical self-organizing maps SOM for the phonemes recognition. The SOM network has been implemented on a chip using 0.6 double Pym in metal with poly technology CMOS double. The results for self-organizing maps SOM are presented in phonemic recognition, and they are encouraging. They also presented the tested performance characteristics for the chip.

T. Kohonen and K. Torkkola in 'Microprocessor implementation of a large vocabulary speech recognizer and phonetic typewriter for Finnish and Japanese’, have applied successfully the strategy of phonotypic maps SOM called 'Feature Maps', in Finnish and Japanese recognition: mono speaker, in single words, for a vocabulary of 1000 words. They showed that phonetic recognition varies from 75% to 90%, the word recognition rate varies from 96% to 98%, and the word orthographic transcription varies from 90% to 97%, depending on the vocabulary and the speaker [4].

We propose to make preliminary experiments of phonemic recognition on TIMIT database with multi-speaker speech. We are following this approach:

- Reading and segmentation of TIMIT database sentences. The program performs the sentences segmentation into 10ms sound atoms, characterized by stationarity.

- Conversion of atoms in the MFCCs coefficients acoustic vectors for each signal window. Some approaches use the central window MFCCs of each phoneme, where maximum power is considered, however, our experiments results use the full windows of the phoneme.

- Classification of MFCCs coefficients in macro classes lists / vowels, semi vowels, nasals, fricatives, affricates, stops and the other.

- Phonemes identification of each TIMIT database class, backup and convert them to data structure 'sD' consistent with the SOM map structure.

- Training the map SOM, after considering the parameters dimension, size and network topology.

The phoneme recognition rates result on the SOM, are given in the table below:

TABLE I. PHONEMIC RECOGNITION OBTAINED BY STATIC SOM IN TIMIT DATABASE

Phonemes Rec. by SOM	Training		Test	
	Nbr Samples	Rate	Nbr Samples	Rate
Vowels	4432	74.95	1339	52.89
Semi-vowels	2717	86.57	844	70.99
Nasales	1109	94.32	332	64.45
Fricatives	1729	82.95	528	59.43
Stops	3946	75.95	1122	48.86
Others	239	75.95	75	52.75
Affriquates	172	90.03	42	82.29
Global Rates	14344	**82.96**	4282	**61.66**

The Kohonen algorithm, allows recognizing only static data with a fixed topology (the SOM outputs remain exactly the same positions of the input data). It is said that each neuron of the map SOM specializes in the recognition of a certain data. We propose in the following sections to make the self-organizing map a dynamic model by assigning loops to integrate the recurrence time.

Finally, in order to achieve better recognition rates, we propose to hybridize the SOM map with the genetic algorithm GA.

III. CHOSEN MODEL FOR THE TIME INTEGRATION

Our choice was made on the analysis of a model that integrates the time by application of loop induction on the map SOM; it is the model RSOM described by the following representation [5]:

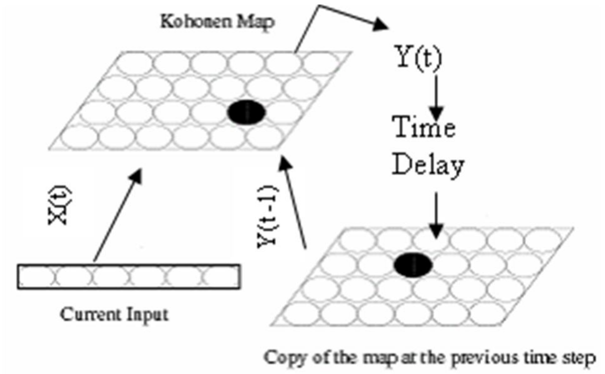

Fig. 1: Representation of the RSOM model

The acoustic vectors characterizing the phonemes were presented serially to the map RSOM. We applied this algorithm for the recognition of vowels and semivowels TIMIT database. The results are presented in the tables below:

TABLE II. COMPARISON OF VOWELS RECOGNITION RATES BY SOM AND RSOM

Vowels	SOM		RSOM	
	Train.	Test	Train.	Test
‘aa’	72.60	62.97	74.31	66.80
‘ae’	76.37	70.46	77.79	69.50
‘ah’	70.50	38.98	72.28	36.04
‘ao’	77.10	70.04	78.12	67.30
‘aw’	83.75	41.66	81.00	44.69
‘ax’	72.60	54.01	73.11	51.18
‘axr’	74.15	56.19	72.09	50.55
‘ay’	73.83	49.59	76.94	52.13
‘eh’	68.41	54.08	71.00	60.92
‘er’	82.71	50.35	81.90	51.66
‘ey’	78.10	51.28	78.15	56.00
‘ih’	63.91	43.80	64.23	50.72
‘ix’	54.23	60.41	53.17	57.99
‘iy’	75.48	74.20	71.46	64.39
‘ow’	72.86	46.78	75.68	48.07
‘uh’	85.28	30.08	85.21	37.40
‘uw’	90.40	47.54	89.85	54.61
‘ux’	83.25	49.76	79.96	46.01
Average rates	**74.95**	**52.89**	**75.34**	**53.66**
Allotted time	**t = 37 min for RSOM** **t = 35 min for SOM**			

TABLE III. COMPARISON OF SEMI-VOWELS RECOGNITION RATES BY SOM AND RSOM

Semi-vowels	SOM		RSOM	
	Train.	Test	Train.	Test
‘el’	89.22	37.00	94.05	41.12
‘hv’	91.13	52.00	94.14	58.03
‘r’	93.73	90.30	100	96.50
‘i’	75.07	83.79	86.60	83.40
‘y’	95.57	86.67	90.59	83.66
‘hh’	90.97	75.00	85.99	79.35
‘w’	70.32	72.18	78.60	78.17
Average rates	**86.57**	**70.99**	**89.99**	**74.31**
Allotted time	**t = 33 min for RSOM** **t = 29 min 30 s for SOM**			

We note that the execution time of the RSOM model is slower than the standard SOM model; this is due to the recurrence loops latency.

Similarly, we find that the obtained rates by working on all the windows of phoneme support, considering the total energy, are better than those obtained by locating the work on a single central window of the phoneme, where we consider that the phoneme energy is maximum, because there is no rule which confirms that the maximum energy of a phoneme is located in the center of its support. In addition, the obtained values will strongly depend on the width and the type of selected window.

IV. SOM EVOLUTIONARY PARAMETERS

The combined use of Artificial Neural Networks RNA and Genetic Algorithms GA for solving the same problem was mentioned by Mikhail Crucianu, in April 1997 in Algorithms for evolving neural networks [6]. Indeed, a RNA type model more precisely the self organizing maps SOM of Kohonen is defined by a relatively large number of parameters which can be optimized by the GA. The connections weights are usually variables whose numerical values are optimized by genetic algorithm. In some include works such Maniezzo, 1993, the precision for such weights values is also optimized by the GA. The precision deduced from a differential equation (Ec) modeling the weight is one of the control parameters of the network complexity, her optimization should improve the generalization ability of the network.

In 1999, Kim and al, suggest an adaptation scheme of SOM maps by evolution and training for image classification. Structure adapted by Kim is a chromosomal structure consisting of a successive real number, which generates the initial population randomly using a uniform distribution. Indeed, the said structure is composed of a prototype vectors number, the map dimension, and a vector prototype components number.

In 1994, McInerney and Dhawan have proposed a hybrid algorithm of GA with the Kohonen training. Their approach is that the chromosome supported on the concatenation of the prototype vectors and the training parameters such as the cycles number, the iterations number, the map size, the update equations, and the map topology and so on. It is not so easy to optimize simultaneously the prototypes and the training parameters. The connectivity of the Kohonen map whose the units number is given can be optimized by the algorithm GA. It presents an another control parameter of complexity and its reduction is coating of others desirable effects such as a faster learning especially for evolutionary algorithms, and easy knowledge retrieval from the network after learning.

In 1995, Hamalanian proposed a modeling method of Kohonen maps using the GA. In its approach, the adopted chromosomal structure is a binary string determined by a connection matrix where 1 is the existence of a connection between two or the same neuron and 0 indicates no connection. The upper triangle of the matrix will be used as a chromosome, the latter being illustrated in the following figure:

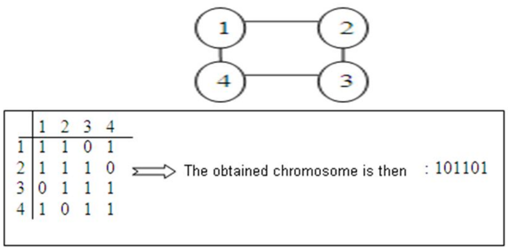

Fig. 2: Hamalanian chromosomal Code inferred from four neural units

V. ADOPTED STRATEGY FOR AN EVOLUTIONARY SOM

The adopted GA-SOM concept represents a strategy of phonemes recognition by the Kohonen SOM map evolution with a tendency to maintain the populations' genetic diversity [7],[8]. The SOM ability of approximation facilitates, in large part, the conservation of this genetic diversity. In the traditional evolution methodology based on tournament selection, the populations' genetic diversity will be degraded in earlier generations. However, the observed approximation property of SOM allows this algorithm to maintain the populations' genetic diversity across generations. Therefore, we can avoid premature convergence of GA and increase the space of our research towards an optimal solution. The development of a hybrid model, evolutionary self-organizing GA- SOM is introduced from the following outcome: an application of the algorithm GA on a given problem translates into a premature convergence and unaccepted results. As for the use of the SOM map, due to the treatment of this problem, we obtain the recognition rate relatively limited to local optima. Thus, during the course of optimization by applying the GA on the SOM recognition tool, a number of individuals made up of SOM maps, each representing a possible solution to the phonemic recognition problem, is produced, evaluated then recombined by following evolutionary operators adopted to produce progeny (offspring) through the generations. Only the information of previous generations, implicitly and partially, is preserved in the current genome. This strategy may raise the risk of individuals' regeneration (renewal) that has already been seen in the research process, in the case of population not rich in diversity. Our approach is to use an effective tool not only to monitor the whole process of the algorithm GA evolution, as in the case of using K-means in the individuals' generation of the GA population, but also to extract valuable data by ensuring the populations genetic diversity through the SOM map training which is used to guide the search process [9].

The proposed method for the GA-SOM hybridization is to introduce phonemic data to the map SOM for learning or test reasons. A single neuron of the SOM map will be activated and will be appointed BMU over each iteration. The winner neuron is the best neuron representative of data inputs to this iteration, among other neurons of the map. Thus, through the training iterations we will have each time a neuron BMU which specializes in an input and we will have a SOM map type assigned to this data which is treated as an individual (a chromosome) for the population reconstitution to handle by GA, according to the following diagram:

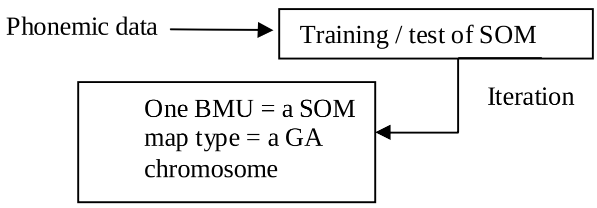

Fig. 3 : Explanatory diagram of the SOM hybridization

Each chromosome will be represented by a criteria matrix corresponding to the criteria matrix for each neuron of a SOM map type during a learning or test iteration [10].

In our approach, the new chromosomes forming the new population of the next generation are determined by the quality of SOM training, specifically, by the changes equation and the vectors weight updating. This update equation for the SOM training is modified by the addition of new coefficients respecting the chromosomes fitness values of current population. This idea is actually based on the fact that the neuron weight is the largest organelle in a unit. It simulates completely the ability of an input data. Thus, the evolution of data recognition result of the evolutionary process of units' weights related to the SOM topology as a diverse population.

The GA-SOM technique implementation is based on the chromosomes evolution. We start with the actual values of the current generation presented to the SOM, in the form of input vectors composed of phonemes MFCCs to be recognized.

After the SOM training, the weight vectors will be considered as the chromosomes of the next diversified generation. That means the next generation population is decided by the SOM training [10].

Then, the new chromosomes will be reproduced from the parents, as obtained prototype vectors, with a normalized function of a uniform geometric selection. In preparing for a new generation reconstruction, the chromosomes will be affected by traditional operators known to the GA, such as recombination (crossover) and mutation.

VI. EXPERIMENTAL RESULTS OF GA-SOM MODEL

The obtained results are shown in the following tables in a comparative form between the GA, the GA-SOM and SOM:

TABLE IV. RECOGNITION RATE OF VOWELS IN THE DR1 OF TIMIT DATABASE COMPARED APPLYING SOM, GA AND GA-SOM

Phonemes	GA		SOM		GA-SOM	
	Train. rates	Test rates	Train. rates	Test rates	Train. rates	Tes t rates
'aa'	95.51	95.52	72.60	62.97	96.17	74.54
'ae'	97.72	96.00	76.37	70.46	95.51	89.18
'ah'	96.76	94.31	70.50	38.98	62.00	57.00
'ao'	95.96	94.05	77.10	70.04	80.50	84.74
'aw'	98.33	95.63	83.75	41.66	71.82	87.25
'ax'	96.32	98.95	72.60	54.01	44.30	61.00
'axr'	97.04	96.58	74.15	56.19	82.58	89.52
'ay'	95.67	98.34	73.83	49.59	52.89	72.64
'eh'	96.17	94.86	68.41	54.08	84.85	65.48
'er'	97.04	92.36	82.71	50.35	84.00	79.20
'ey'	97.13	96.00	78.10	51.28	77.47	83.96
'ih'	97.15	94.07	63.91	43.80	80.43	79.61
'ix'	97.68	94.30	54.23	60.41	87.29	79.20
'iy'	98.68	92.60	75.48	74.20	93.17	97.16
'ow'	95.41	95.25	72.86	46.78	80.00	80.00
'uh'	97.17	95.36	85.28	30.08	63.16	60.00
'uw'	95.39	97.52	90.40	47.54	78.36	73.07
'ux'	95.58	93.41	83.25	49.76	80.64	80.19
Average	96.70	95.28	74.95	52.89	78.85	76.35
Allotted time		t = 14 h for GA-SOM (training case) t = 35 min for SOM (training case) t = 12 h 20 min for AG-SOM (test case) t = 25 min for SOM (test case)				

VII. Description of the Adopted GA-RSOM Model

First, we know that a phoneme is a sound entity characterized by some stationary. This specificity is a limit to the robustness of a recurrent model that includes time in phoneme recognition. To overcome this constraint, we tried to perform our experiments rather than on a central window of each phoneme, where maximum energy is considered, but on all the windows of each phoneme. The use of the entire signal range of a phoneme can lead us to explore some temporal variability that can benefit in relation to the results provided by the GA-recurrent SOM (GA-RSOM) compared with those of GA- Static SOM. This opens another research exit such as the whole words recognition, where the proposed model GA-RSOM is interesting.

The developed model GA-RSOM is a dynamic phoneme recognition model, and recurrent to take into account the speech variability [11]. Thus, it is evolving with the aim of optimizing the obtained recognition results. The following algorithm starts by creating an initial population constituted by the neural weights initialization of the developed map RSOM. The serial data Sequences to recognize will be presented to the map. Each vector element of the input sequence will be transmitted to each map neuron. While taking into account the memory effect of recurrent SOM map, we get at the end of each training' iteration a prototype vector for each neuron in the same size and the same distribution as that of the input sequence. Determining the difference between each of these prototypes and the input sequence leads to calculate the squared error and later to determine the scored effectiveness of each of these individuals. Once the individual scores are established, the reconstruction phase of the RSOM map will begin under GA. This step will create a new population of prototype vectors weight from the previous one.

It will first select two individuals from the previous generation based on their efficiency score. Then these two individuals will be crossed in order using local performance of each chromosome. This performance will be described by the weighting of each weight vector component, which is to say, by the genes criteria components the chromosome vector. Optionally, the newly created individual by crossing will be transferred with a ratio of Pm = 0.6, taking into account the interactions effects between individuals being created.

This process of selection, crossover, and mutation is performed as many times as needed to rebuild a new population by the winning neural prototypes 'BMU', at each training' iteration of the RSOM map. The stopping criterion of the algorithm is provided either after a certain generations' number, or by obtaining a performance criterion which results in an optimal recognition rate. The general flow of the adopted algorithm is presented as follows:

a. creating an initial population from the RSOM map initialization.

b. Presentation of serial data sequences to the map and prototyping of neurons representing individuals in the population.

c. Calculation of scores on the overall effectiveness of individuals.

d. Selection of parents to cross.

e. Calculating the local effectiveness of each parent and considering genes groups to cross.
f. Recombination of parents and creation of a child individual.

g. Mutation of this child.

h. Perform a loop from the selection until a new population.

i. Choosing a winner individual in the population appointed by the 'BMU' of the RSOM map.

j. Calculating the recognition rate of the gaining unit.

k. Check the stopping criterion of the algorithm.

VIII. Experimental Results of GA-RSOM Model

The evaluation of the developed model noted GA-RSOM in multi speakers' speech recognition and regardless of context was made on the TIMIT corpus. This database contains a total of 6300 sentences. Each group of 10 sentences is uttered by one of 630 speakers from 8 major dialect regions of the United States of America. The wording of the speech is sampled at a sampling frequency of 16 kHz in 16 bits. The obtained frames of speech are filtered by a first order filter whose transfer function is given by the following equation:

$$H(z) = 1 - a.z^{-1}, \quad 0.9 \le a \le 1.0 \tag{2}$$

We proceeded in our experiments, to the separation of the entire database in macros classes each containing a phonemes family. Each macro class is treated separately. The phonemes list for each macro-class is shown in the following table:

TABLE V. THE SEVEN CLASSES OF TIMIT PHONEMES

Macro class	phonemes
Vowels	/iy/, /ih/, /eh/, /ey/, /ae/, /aa/, /aw/, /ay/, /ah/, /ao/, /oy/, /ow/, /uh/, /uw/, /ux/, /er/, /ax/, /ix/, /axr/, /axh/
Semi-vowels	/l/, /r/, /w/, /y/, /hh/, /hv/, /el/
Nasals	/m/, /n/, /ng/, /em/, /en/, /eng/, /nx/
Stops	/b/, /d/, /g/, /p/, /t/, /k/, /dx/, /q/, /bcl/, /dcl/, /gcl/, /pcl/, /tcl/, /kcl/
Fricatives	/s/, /sh/, /z/, /zh/, /f/, /th/, /v/, /dh/
Affriquates	/jh/, /ch/
Others	/pau/, /epi/, /h#/

Starting from the consideration that most of the vowels used in the languages are spoken with the vocal cords vibration, they are called sound. However, the deaf vowels are pronounced without the vocal cords vibration, such as the Whisper [12; 13].

Thus, the vowels are different to consonants, because the latter are characterized by obstruction to airflow accompanied by noises such as whistling, clapping, etc.. In addition, the vowels are manifested by clear sounds and are more perceptive. For this, we have used this phonemes class, in our experiments to evaluate the GA-RSOM model we have developed. We have compared it to recognition rates average for other models, such as the SOM, and the GA-SOM for static data. These different values are mentioned in the two following tables:

TABLE VI VOWELS RECOGNITION RATES IN TIMIT TRAINING DATABASE

Phonemes	SOM	GA-SOM	GA-RSOM
‘aa’	72.60	96.17	96.83
‘ae’	76.37	95.51	96.02
‘ah’	70.50	62.00	59.17
‘ao’	77.10	80.50	83.11
‘aw’	83.75	71.82	76.99
‘ax’	72.60	44.30	45.16
‘axr’	74.15	82.58	83.00
‘ay’	73.83	52.89	54.10
‘eh’	68.41	84.85	86.88
‘er’	82.71	84.00	81.45
‘ey’	78.10	77.47	80.34
‘ih’	63.91	80.43	80.50
‘ix’	54.23	87.29	89.66
‘iy’	75.48	93.17	95.23
‘ow’	72.86	80.00	81.19
‘uh’	85.28	63.16	65.72
‘uw’	90.40	78.36	77.94
‘ux’	83.25	80.64	82.38
Average rates	74.95	78.85	79.11
Allotted time	t = 14 h for GA-SOM t = 14 h 30 min for GA-RSOM t = 35 min for SOM		

TABLE VII VOWELS RECOGNITION RATES IN TIMIT TEST DATABASE

Phonemes	SOM	GA-SOM	GA-RSOM
'aa'	62.97	74.54	72.77
'ae'	70.46	89.18	91.09
'ah'	38.98	57.00	61.34
'ao'	70.04	84.74	84.12
'aw'	41.66	87.25	88.45
'ax'	54.01	61.00	66.13
'axr'	56.19	89.52	87.44
'ay'	49.59	72.64	73.69
'eh'	54.08	65.48	68.28
'er'	50.35	79.20	80.80
'ey'	51.28	83.96	85.67
'ih'	43.80	79.61	79.44
'ix'	60.41	79.20	79.00
'iy'	74.20	97.16	97.21
'ow'	46.78	80.00	82.73
'uh'	30.08	60.00	59.01
'uw'	47.54	73.07	74.78
'ux'	49.76	80.19	81.55
Average rates	52.89	76.35	77.64
Allotted time	t = 12 h 20 min for GA-SOM t = 12 h 50 min for GA-RSOM t = 25 min for SOM		

The allocated time is given in training and test case of adopted models.

A comparison of the experimental results of the different models mentioned above, tells us to:

The determined recognition rates from the test database are in all models cases lower than those provided by using the training database. This means that the self consistency rate of our developed model as for the other models considered for comparison, is still high compared to their generalization levels.

In addition, an evolutionary SOM model GA-SOM performs better in a global search space and taking into account the advantage of a diverse population favored by the SOM map property. However, this model is beneficial only for the static data recognition.

In improvement of the GA-SOM model, previously seen, we have proposed an evolutionary recurrent SOM model in order to take into account the data temporal aspect presented in phonemic recognition. Experimental results show that this model has an important position in terms of recognition of even dynamic data. However, the improvements gap is usually small compared to the results of the GA-SOM model for static data; this is illustrated by the two following figures which give the different gaits of obtained recognition rates. The latter allows us to think of other improvements required to our model to further tighten the gap and to benefit our dynamic recognition model.

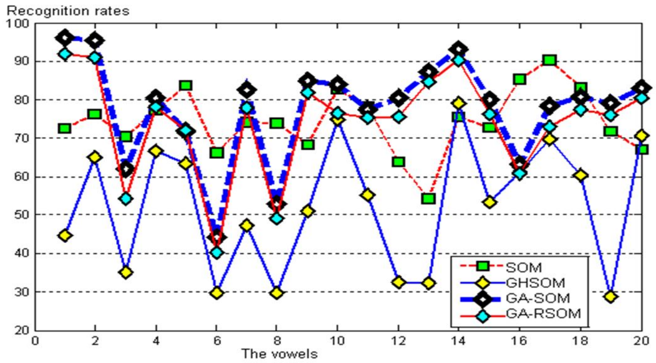

Fig. 4: Recognition rates for SOM, GHSOM, GA-SOM and GA-RSOM models in training

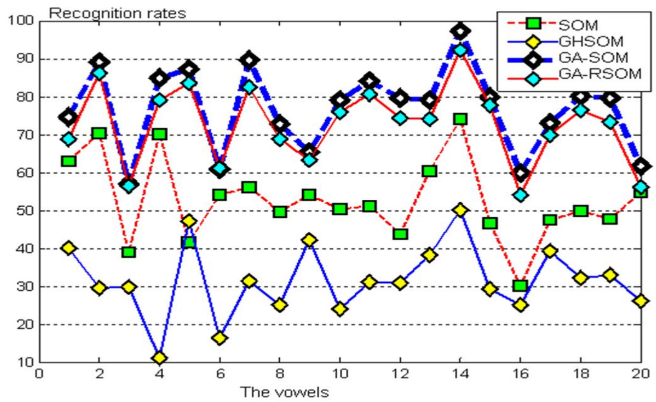

Fig. 5: Recognition rates for SOM, GHSOM, GA-SOM and GA-RSOM models in test

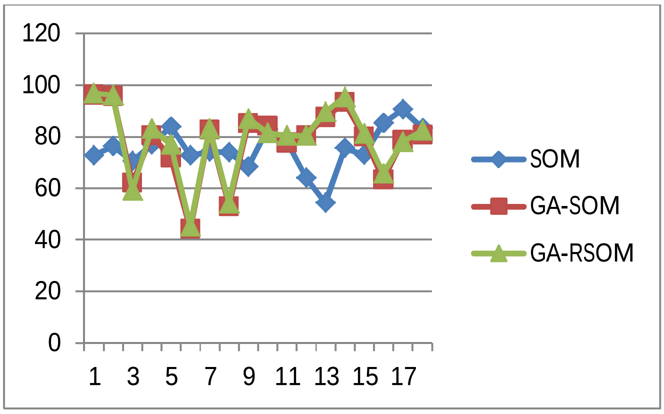

Fig. 6: Recognition rates for SOM, GA-SOM and GA-RSOM models in training

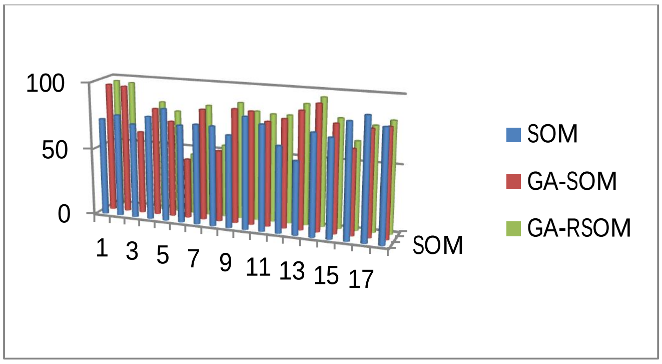

Fig.7: Histogram of recognition rates for SOM, GA-SOM and GA-RSOM, training case

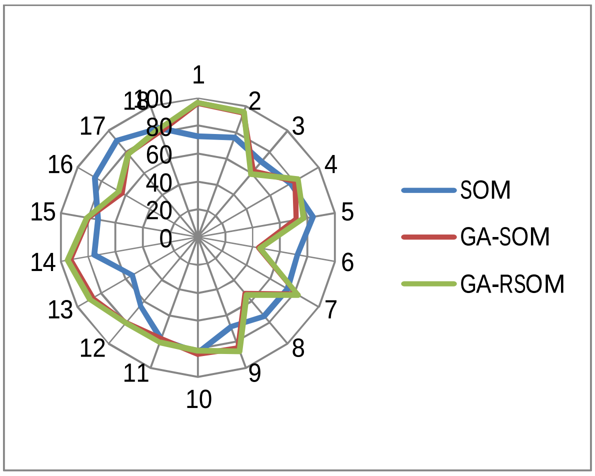

Fig. 8: Recognition rates for SOM , GA-SOM and GA-RSOM models in training

IX. CONCLUSION

We never stop to learn and discover in life. Knowledge is helpful only if it is truly shared.

Several models of speech recognition, based on the application of Kohonen self organizing map SOM, were used during the last decade in an attempt to improve the recognition rate. The results appear remained in local optimal with respect to the algorithm of followed model. In other circumstances, most of these models are used in handling static data. To overcome these limitations, we developed an evolutionary and recurrent self-organizing model aimed to optimize the results in a larger search space and take into account the temporal variability of speech.

This model has been developed is called GA-RSOM. Promising results are obtained by comparing its average recognition rates to other models, such as the SOM, the GHSOM, and the GA-SOM for static data.

This task is performed by means of adaptive operators manipulating extracted data by a self-organizing map SOM laying individuals of previous generations. The evaluation of this approach shows that the model GA-SOM is a tool to address the issue of early convergence in GA, which may have a solution for other problems. Thus, the GA-SOM model with recurrence GA-RSOM offers slightly improved results compared to the static model GA-SOM, this is explained by the property of phoneme stationarity to recognize which forces the GA-RSOM model to behave as being a static model GA-SOM.

REFERENCES

[1] Julien RACHEDI : 'Reconnaissance et classification de phonèmes', Laboratoire IRCAM, Paris, Aout 2005.

[2] Marie Cottrell et al : 'Cartes auto-organisées pour l'analyse exploratoire de données et la visualisation', Université Paris 1, 2003.

[3] Gaussier E, et al: 'A hierarchical model for clustering and categorising documents', Advances in Information Retrieval, Proceedings of the 24th BCS-IRSG European Colloquium on IR Research ECIR-02, Glasgow. Lecture Notes in Computer Science 2291, p.229-247, Springer, 2002.

[4] Kohonen, T.: "*Self-organizing maps*", Springer, Berlin, 1995.

[5] Thomas voegtlin 'thèse': "Réseaux de neurones et autoréférence", université lumière- lyon II, 2002.

[6] Mihail Crucianu : ' Algorithmes d'évolution pour les réseaux de neurones', Rapport de recherche 187, Février 1997, Révisé Avril 1997, Laboratoire d'Informatique Ecole d'Ingénieurs en Informatique pour l'Industrie 64 avenue Jean Portalis 37200 Tours, France.

[7] Shimon Whiteson, Peter Stone: "Evolutionary Function Approximation for Reinforcement Learning", to appear in Journal of Machine Learning Research, 2006. Department of Computer Sciences.University of Texas at Austin.

[8] J.H. Holland : "Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to biology, control and artificial intelligence". MIT Press, ISBN 0-262-58111-6. 1998.

[9] Mohamed Salah Salhi, Najet Arous, Noureddine Ellouze: 'Principal temporal extension of SOM', journal IJSIP Koréa, laboratoire LSTS-ENIT_Tunis, septembre 2009.

[10] Mohamed Salah Salhi, Najet Arous, Noureddine Ellouze: 'A suitable model of evolutionary SOM for phonemes recognition', journal IRECOS Napoly-Italy, (Indexé COMPENDEX – Elsevier–Copernicus). Laboratoire LSTS- ENIT_Tunis, septembre 2009.

[11] Andreea Lazara, Gordon Pipaa, Jochen Triescha : 'Fading memory and time series prediction in recurrent networks with different forms of plasticity', Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Max-von-Laue-Str. 1, 60438 Frankfurt am Main, GermanyNeural Networks 20 (2007) 312–322, Elsevier, 2007 Special Issue.

[12] S. Mallat : '*Une exploration des signaux en ondelettes*', Editions de l'Ecole Polytechnique, 2000.

[13] Julien BRICHE : 'Adaptation d'un algorithme génétique pour la reconstruction de réseaux de régulation génétique : COGARE', thèse de l'Université du Sud Toulon, 2009.