Arabic Vowels Recognition by Modular Arithmetic and Wavelets using Neural Network

Emad F. Khalaf, Khaled Daqrouq, Ali Morfeq

Electrical & Comp. Eng. Dept, King Abdulaziz Univ., Jeddah, Saudi Arabia

Abstract: Recently, the speech recognition is very attractive for researchers because of the very significant related applications. For this reason, the novel research has been of very importance in the academic community. The aim of this work is to find out a new and appropriate feature extraction method for Arabic language recognition. In the present study, wavelet packet transform (WPT) with modular arithmetic and neural network were investigated for Arabic vowels recognition. The number of repeating the remainder was carried out for a speech signal**.** 266 coefficients are given to probabilistic neural network (PNN) for classification. The claimed results showed that the proposed method can make an effectual analysis with classification rates may reach 97%**.** Four published methods were studied for comparison. The proposed modular wavelet packet and neural networks (MWNN) expert system could obtain the best recognition rate.

[Emad F. Khalaf, Khaled Daqrouq, Ali Morfeq. **Arabic Vowels Recognition by Modular Arithmetic and Wavelets using Neural Network.** *Life Sci J* 2014;11(3):33-41]. (ISSN:1097-8135). http://www.lifesciencesite.com. 6

Keywords: Speech; Arabic vowel, Wavelet; Modular arithmetic; Neural network.

1. Introduction

The recent rising increase of loudening activity in mobile communication domain draw new opportunities and shed some lights for applications of speech recognition including words and sentences. Text to speech or vice versa is incredibly critical issues in many computer applications. Where, English language has achieved vast success of the major part of interest. However, Arabic language speech recognition has been less attractive than English; because of its many nature difficulties, in term of, several alphabets forms and various dialects.

The key study of investigating a speech recognition of Arabic language tackling with the morphological structure may be found in (Kirschhoff, 2003; Datta et al., 2005) or utilizing the phonetic features in order to recognize the famed Arabic phonemes (pharyngeal, geminate and emphatic consonants) (Debyeche et al., 2006;Selouani and Caelen, 1999) and interpretation their further inference in a biger vocabulary recognition speech design. This allocates and encourages interested researchers of Arabic language with several dialect at different countries. The achievement in term of implementation of recognition systems dedicated to spoken separated words, or continuous talk are not widely conducted. Only some examples have been investigated, improved in many research papers to establish a new trend of research of Arabic spoken with several dialects. Derivative proposal, denoted by the concurrent generalized regression neural networks (GRNN), was considered by Shoaib et al. (2004). This was implemented for precise Arabic phonemes recognition with the purpose of automate the intensity and formants-based feature extraction. Alotaibi, (2005) used multiple layer neural network (MLN), to investigated a separated word speech recognition. The accurateness of the recognition rate in speaker-independent system was 94.5% and in speaker-dependent system was 99.5%. A group of Arabic speech recognition different systems was studied in Amrouche, A., et al., (2009).

The Fuzzy C-Means approach has been given to the traditional ANN/HMM speech recognizer by perceptual linear prediction (PLP) or RASTA-PLP features vectors. The Word Error Rate (WER) is above 14.4%. With the same way, a method utilizing data fusion had a WER of 0.8%. Nevertheless, this technique was examined only on one personal corpus and the authors showed that the accomplished improvement required the use of three neural networks running in parallel. In Bourouba et al. (2006), an additional hybrid method was proposed, where the K nearest neighbour (KNN) and the Support Vector Machine (SVM) were given to the ANN in the conventional hybrid system. On the other hand, the recognition rate, for KNN/HMM and for SVM/HMM did not go beyond 92.72% and 90.62%, respectively.

An innovative Algorithm to recognize separate spoken Arabic digits was studied in Saeed and Nammous, (2005b). In feature extraction stage, the algorithm of minimal eigenvalues of Toeplitz matrices with other methods of speech processing, were implemented. The recognition rate taken in the conducted experiments was almost excellent, and exceeded 98%, for many examples. A hybrid approach has been applied to Arabic digits recognition by (Lazli and Sellami, 2003).

Neural networks attract many researchers to extract and recognize Arabic language properties fo example: emphasis, gemination and vowel lengthening. (Alotaibi, 2005). This was studied using ANN and other techniques (Selouani and Douglas O., 2001), where many systems and configurations were considered including time delay neural networks (TDNNs). (Salam M., et al., 2001) proposes 10 Malay digits ANNs identification. Saeed and Nammous (2005a) introduced a heuristic method of Arabic digit classification, by the Probabilistic Neural Network (PNN). A neural network method, with utilizing of a nonparametric activation function could present good results to enhance the quality of speech recognition systems, mostly in the case of Arabic language.

The advantages of the GRNN speech recognizer over the HMM in calm environment and the multilayer perceptron MLP, was demonstrated in Amrouche and Rouvaen, (2003). Also, the noisy environments degrade the recognition performance considerably. Strength to noise is then indispensable for professional using recognition systems mainly in mobile networks context; (Karray,. L. and Martin A, 2003). From this conception variety of studies have been conducted (Savoji, 1989; Mauuary and Monne, 1993; Junqua, et al., 1994; Mokbel et al., 1997). To decrease or remove the noise effects in the speech before adding to a recognizer different pre-processing techniques have been developed. In (Berouti et al., 1979; Mokbel et al., 1997) eenhancement process such as spectral subtraction, eliminates surrounding noise. In (Mokbel et al., 1995; Hermansky et al., 1993) the transmission effects equalization techniques, for instant cepstral normalization and adaptive filtering are proposed. Modular arithmetic with wavelet transform was proposed for gender and speaker recognition by (Khalaf et al., 2011a; Khalaf et al.,2011b).

This paper presents a new combination of wavelet transform with modular arithmetic and probabilistic neural network. The objective of such conjunction is to construct an Arabic vowels classifier, with high performance. The paper presentation will be presented as follows: In Section 2 we present the short introduction to Arabic language. The proposed method is described in Section 3 and Section 4. The experimental results and discussion in Section 5, followed in Section 6 by conclusion.

2. Arabic Language

There are 22 Arabic countries with around 350 millions Arabic speakers living in it or distributed all over the world. For this reason, Arabic language considered one of the most important and widely spoken languages in the world. Arabic is Semitic language that is characterized by the existence of particular consonants like pharyngeal, glottal and emphatic consonants. As well as Arabic language has some phonetics characteristics that are built, around pattern roots (CVCVCV, CVCCVC, etc.) (Zitouni I. and Sarikaya R. ,2009).

The 28 letters can be used in a set of 90 additional combinations, shapes, and vowels (Tayli and Al-Salamah, 1990). The 28 letters enclose consonants and long vowels such as ى and أ (both pronounced as/a:/), ي (pronounced as/i:/), and و (pronounced as/u:/). The short vowels and some other phonetic pronouncing like consonant doubling (shadda) are not introduced using letters directly, but by diacritics. The diacritics are short strokes, where each can be located above or below the consonant. Complete set of Arabic diacritics are represented in Table 1. Arabic discretization is interpreted by three groups: short vowels, doubled case endings form, and syllabification marks. First set, short vowels are written as symbols above or below the letter in the word using diacritics. In some cases, are used all together in text without diacritics. We have short vowels: fatha: it pronounce as /a/ sound and is an slanted dash above the letter, damma: it pronounce as /u/ letter sound and has form of a comma above the letter and kasra: it pronounce as /i/ letter sound and is an oblique dash under the letter as tabulated in Table 1 (Zitouni I. and Sarikaya R. ,2009).

Table 1. Diacritics above or below consonant letter

Short Vowel Name (Diacritics)	Diacritics above or below letter 'ب' (sounds B)	Pronunciation
Fatha	بَ	/ba/
Damma	بُ	/bu/
Kasra	بِ	/bi/
Tanween Alfath	بً	/ban/
Tanween Aldam	بٌ	/bun/
Tanween Alkasr	بٍ	/bin/
Sokun	بْ	/b/

Therefore, it is essential to realize that, what we usually denote to "Arabic" is not single linguistic variety; rather, it is a set of separate dialects and communities. Classical Arabic is an older and literary figure of the language, exemplified by the type of Arabic used in the Quran, the holly book for Islam. Modern Standard Arabic (MSA) is a version of Classical Arabic based on a modern vocabulary. MSA is a formal standard popular to all Arabic-speaking countries and communities. It is the language utilized in the newspapers, radio and TV, in official speeches, in courtrooms, and in any kind of formal communication.

Though, it is not utilized for everyday speech, informal communication, which is classically applied in one of the particular dialects. The dialects of Arabic may roughly speaking be divided into two sets:

Western Arabic, which consists of the dialects spoken in Morocco, Algeria, Tunisia, and Libya, and Eastern Arabic, which may be further divided into Egyptian, Levantine, and Gulf Arabic countries. These different dialects differ significantly from each other and from Modern Standard Arabic. Differences influence all levels of language, i.e. pronunciation, phonology, vocabulary, morphology, and syntax. Table 2 lists examples of the differences between Egyptian Arabic Dialect (EAD), Jordanian Arabic Dialect (JAD), Palestinian Arabic Dialect (PAD) and Modern Standard Arabic. EAD is that dialect which is most broadly understood through-out the Arabic-speaking world, because of a great number of TV programs which are made in Egypt and exported to other Arabic rigions. Native speakers from different dialect regions are for the most part capable of communicating with each other, especially if they have had some preceding exposure to the other speaker's dialect. However, widely contradictory dialects, such as Moroccan Arabic, may hinder communication to the extent that user (speaker) adopt Modern Standard Arabic as a lingua franca.

Table 2: Three examples of four different Arabic dialects

Gloss	MSA	EAD	JAD	PAD
'Three' ثـــلاث	thā-lā-thāh	tā-lā-tāh	thā-lā-thěh	tā-lā-těh
'Eight' ثمانيـــة	thā-mā-nê-yah	tā-mā-n-yah	thā-mā-n-yeh	tā-mā-n-yeh
'Two' أثنيـــن	ith-nān	te-nān	?ith-nen	?it-nān

Several issues of Arabic language, such as the phonology and the syntax, do not have difficulties for automatic speech recognition. Standard language-independent techniques for acoustic and pronunciation modelling, such as context-dependent phones, can easily be carried out to present of the acoustic-phonetic properties of Arabic. The most puzzling difficulties in increasing performance speech recognition systems to Arabic language are the predominance of non-discretised text, the huge dialectal change, and the morphological complexity.

The primarily difficult of the dialectal variability, is by reason of a present absence of training data for spoken Arabic; while, MSA data could freely be acquired from several media sources.

In assumption, morphological complexity is approved to present solemn problems for speech recognition. A high scale of affixation, derivation etc. donates to the explosion of unlike word forms, making it difficult if not impossible to robustly estimate language model probabilities. Affluent morphology also leads to elevated out-of-vocabulary rates and larger search spaces through decoding, therefore, make the recognition process slow (Kirschhoff, 2003).

Modular arithmetic

For a positive integer n, two integers a and b are assumed to be congruent modulo n , written:

$$a \equiv b \pmod{n}, \tag{1}$$

if their difference $a-b$ is an integer multiple of n . The number n is called the modulus of the congruence.

In computer science discipline, it is the remainder operator that is usually indicated by either "%" (e.g. in C, Java, Javascript, Perl and Python) or "mod" (e.g. in BASIC, SQL, Haskell and Matlab). These operators are commonly pronounced as "mod", however, it is specifically a remainder that is computed. The remainder operation can be represented using the floor function (Cormen al., 2001).

If $a \equiv b \pmod{n}$, where n > 0, then the remainder b is calculated:

$$b = a - \left\lfloor \frac{a}{n} \right\rfloor \times n \,, \tag{2}$$

where $\left\lfloor \frac{a}{n} \right\rfloor$ is the largest integer less than or equal to $\frac{a}{n}$, then $a \equiv b \pmod{n}$ and, $0 \le b < n$

If instead a remainder b in the range $-n \le b < 0$ is required, then

$$b = a - \left\lfloor \frac{a}{n} \right\rfloor \times n - n. \tag{3}$$

3. Modular Arithmetic Wavelet Packet Feature Extraction Method

To decompose the speech signal into WPT, we start from the general form of the equivalent low pass of discrete time speech signal

$$u(t) = \sum_m X_m p(t - mT), \tag{4}$$

where X_m is a sequence of discrete speech values, which are obtained by data acquisition block; the signal $p(t)$ is a pulse, whose figure constitutes an important signal design problem when there is a bandwidth restriction on the channel; and T is the sampling time. Considering that $\phi(t-mT)$ is a scaling function of a wavelet packet, i.e., $\phi \in W^0_{2^N}$, then a finite set of orthogonal subspaces can be built as (Daubechies, 1988; Lei et al., 2005)

$$W^0_{2^N} = \underset{(1,n)\in\rho N}{\oplus} W^0_{2^l}, \tag{5}$$

where $W^0_{2^N} \subset L^2(R), \rho N = \{(l,n)\}$ is a dyadic interval that forms a disjoint covering of $[0,2^N]$, $W^n_{2^l}$, denotes the closed linear span of process $\sqrt{2^l}\psi_n(2^l t-m), m\in Z$, and $\{\psi_n(t)\}_{n\in \mathrm{N}}$ is called the wavelet packet calculated by the scaling function ϕ. According to (5), the speech signal model in (4) is customized as

$$u(t)=\sum_m \sum_{(l,n)\in\rho N} X_m \sqrt{2^l}\psi_n(2^l t-m). \tag{6}$$

The speech signal model in (6) is the basic form of wavelet packet transform, which is used in signal decomposition. The signal is carried by orthogonal functions, which shape a wavelet packet composition in $W^0_{2^N}$ space. We may use the inverse discrete wavelet packet transform (IDWPT) procedure and discrete wavelet packet transforms (DWPT) procedure. For the IDWPT, we have

$$\phi^n_l(i)=\sum_{k\in Z} h(i-2k)\phi^{2n}_{l+1}(k)+\sum_{k\in Z} g(i-2k)\phi^{2n+1}_{l+1}(k), \tag{7}$$

and for DWPT,

$$\phi^{2n}_{l+1}(i)=\sum_{k\in Z} h(k-2i)\phi^n_l(k) \tag{8}$$

$$\phi^{2n+1}_{l+1}(i)=\sum_{k\in Z} g(k-2i)\phi^n_l(k), \tag{9}$$

where $\phi^n_{l+1} \in W^n_{2^{l+1}}$ and $\phi^n_l \in W^n_{2^l}$. These two processes can be carried out recursively during the binary tree structure, with $0(N\log N)$ computational complexity. Using (7), (8), and (9), the coefficients of the linear combination can be shown to be the reversed versions of the decomposition sequences $h[k]$ and $g[k]$ (with appropriate zero padding), respectively.

Continuously, we can reconstruct $\phi^1_0(i)$ via the terminal functions of an arbitrary tree-structured decomposition

$$\phi^1_0(i)=\sum_{l\in L, n\in C_l} \sum_{k\in Z} f_{ln}(i-2^l k)\phi^n_l(k), \tag{10}$$

where L is the set of levels having the terminals of a given tree; C_l is the set of indices of the terminals at the lth level and $f_{ln}[i]$ is the equivalent sequence built from the combination of $h[k]$, $g[k]$ and decimation, which leads from the root to the (l,n)th terminal, i.e.,

$$\phi^n_l(i)=\sum_{k\in Z} f_{ln}(k-2^l i)\phi^1_0(k). \tag{11}$$

For a certain tree structure, the function ϕ^n_l in (11) is called the constituent terminal function of ϕ^1_0. In this work, the tree consists of three stages, and therefore has $G_{High} = 2^7$ high pass nodes (with original signal node) and $G_{Low} = 2^7$ low pass nodes with original signal. More generally, for a q-stage tree, there are $G=(G_{High}+\mathrm{G_{Low}})=2^{q+1}-2$ nodes.

In our work we propose the modular arithmetic to decrease the number of features in each WP node; For a speech signal in a WP node $\{u(t)\}=\{u(t_1), u(t_2), ..., u(t_M)\}$, where M is the length of $u(t)$, the remainder b is written as $\mathrm{mod}(n)_{u(t)}=0,1,2,...,n-1,$ for $n>2$. In this paper, the number of repeating the remainder $\mathrm{mod}(n)_{u(t)}$ is utilized as a wavelet speech signal feature vector: $WR(n)=r(0),r(1),r(2),...,r(n-1)$. R is the number of repeating the remainder $\mathrm{mod}(n)$ applied for one node of WP.

4. Classification

We create a probabilistic neural network algorithm for classification problem (Fig. 1):

$$Net = PNN(P,T,SPREAD),$$

where P is $4\times 2^{q+1}24$ matrix of 24 input vowel feature vectors for net training, of 2^{q+1} (minus 2, repeated original node) WP nodes number and modular arithmetic for number 4 (as example);

$$P = \begin{bmatrix} WR_{11} & WR_{12}, \ldots, WR_{124} \\ WR_{21} & WR_{22}, \ldots, WR_{224} \\ . & . \quad\quad . \\ . & . \quad\quad . \\ . & . \quad\quad . \\ WR_{4\times 2^{q+1}1} & WR_{4\times 2^{q+1}2}, \ldots, WR_{4\times 2^{q+1}24} \end{bmatrix}, \tag{12}$$

T is the target class vector

$$T=[1,2,3, \ldots,24], \tag{13}$$

and SPREAD is spread of radial basis functions. We employ a SPREAD value of 1 because that is a typical distance between the input vectors. If SPREAD is near zero the network acts as a nearest neighbour classifier. As SPREAD becomes larger the designed, network will take into account several nearby design vectors.

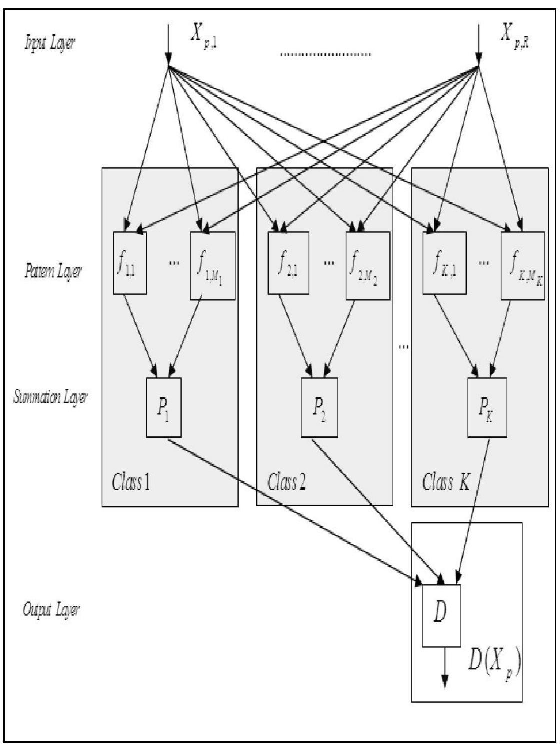


Fig. 1. Structure of the original probabilistic neural network

5. Results and Discussion

In this study, speech signals samples were recorded via PC-sound card, with a sampling frequency of 16 kHz. The Arabic vowels were recorded by 20 donors (speakers): 5 females and 15 males. The recording process was provided in normal university office environment.

These Arabic vowels are distinguishable in term of frequency domain and energy distribution. Figure 2 presents the flow chart of the suggested algorithm. In this paper we are willing to use a new classification method based on modular arithmetic (R, which is the number of repeating the remainder $\mathrm{mod}(n)$ applied for each node of WP). Our investigation of speaker-independent Arabic vowels classifier system performance is performed via several experiments depending on vowel type. To decide of which n is the most suitable for our testing signals. The comparison was performed based on the recognition sensitivity (RS) measure defined as follows:

$$RS = \rho_{XX} - \rho_{XY} \tag{14}$$

where ρ_{XX} is the correlation coefficient calculated for same two vowels of different signals and ρ_{XY} is the correlation coefficient calculated for different two vowels. The recognition sensitivity results calculated for 75 different vowels show that the methods for mod (13) and mod (19) are comparable but the mod (13) provides slightly best results. The results are illustrated by figure 3. So that mod (13) will be used in the next experiments.

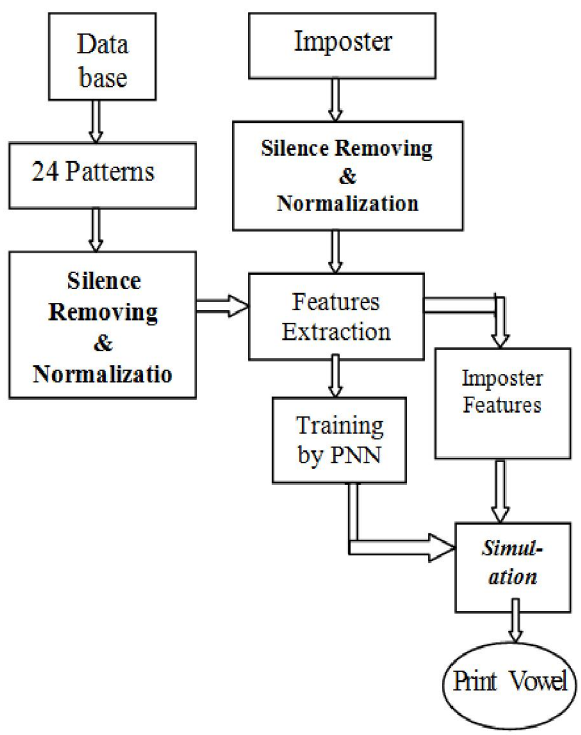


Fig. 2. Flow chart for proposed system

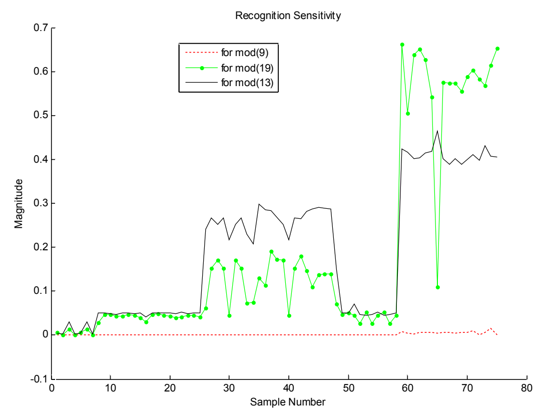
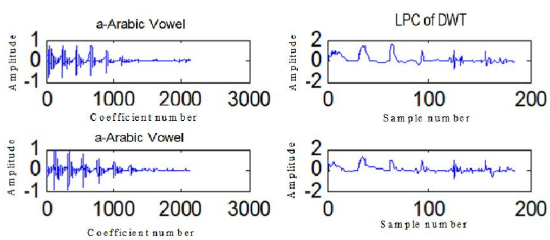
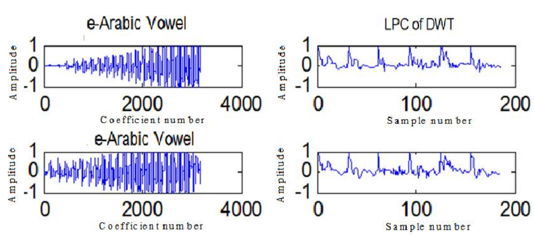


Fig. 3. The recognition sensitivity that is based on correlation coefficient. Three mod were used: mod(9), mod(13) and mod(19).

By means of figure 4, figure 5 and figure 6, we show the possibility of classification by some common functions. Figure 4 illustrates the linear prediction coding obtained from DWT, figure 5 illustrates the power spectrum density and Figure 6 illustrates the spectrogram. Features for Arabic vowels in figure 6 are presented in figure 7.

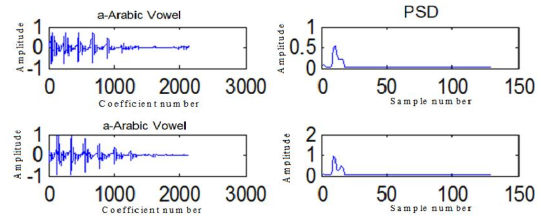


a.

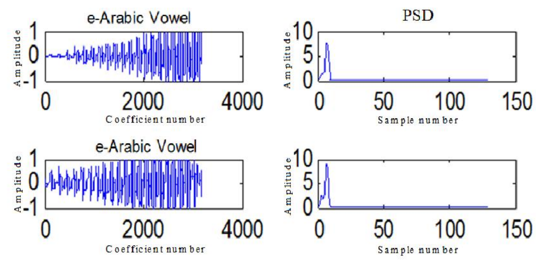


b.

Fig. 4. LPC with DWT features, taken for "a" and "e" Arabic vowels. a. Presents LPC with DWT for two different a-vowels. b. Presents LPC with DWT for two different e-vowels.

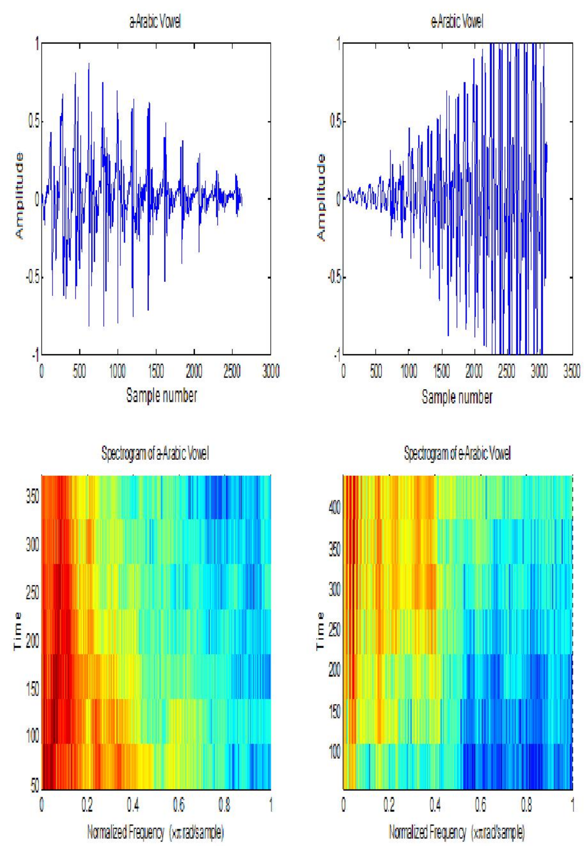


a.

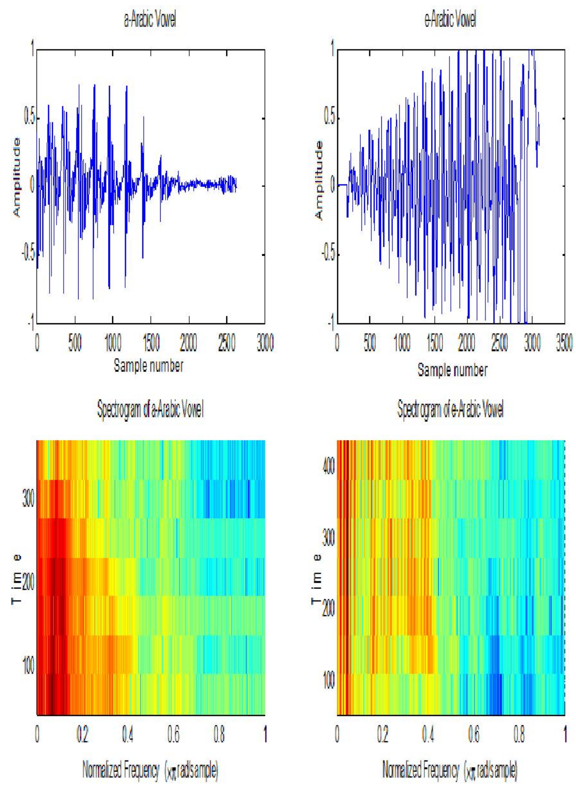


b.

Fig. 5. PSD features for "a" and "e" Arabic vowels. a. Presents PSD for two different a-vowels. b. Presents PSD for two different e-vowels is presented.



a.



b.

Fig. 6. Arabic vowels illustrated by spectrogram a. Vowels “a” and “e” of a speaker 1 with spectrogram b. Vowels “a” and “e” of a speaker 2 with spectrogram. The spectrograms of the same vowels are similar even for different speakers

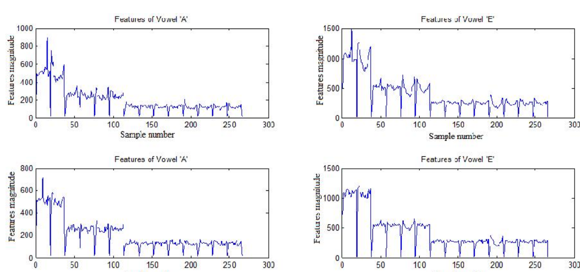


Fig. 7. Features for Arabic vowels presented in Fig. 6 by means of proposed method (MWNN). The feature vectors of the same vowels are similar

Experimental-1

We experimented 180 signals long Arabic vowel أ (pronounced as/a:/) signals of 20 speaker and 358 signals long Arabic vowelي (pronounced as/e:/) signals. The results indicated that 97.31% were classified correctly for Arabic vowel أ. For Arabic vowel ي, 84.36% of the signals were classified correctly. Table 3 shows the results of Recognition Rates (RR).

Table 3: The recognition rate results for long vowels

Long Vowels	Number of Signals	Accepted Signals	Rejected Signals	Recognition Rate [%]
Long A أ	186	181	5	97.31
Long E ي	358	302	56	84.36
			Avr. Recog. Rate	**90.83**

Experimental-2

In this experiment the recognition rates for long vowels linked with other letter such ل (pronounced as/l/) and ر (pronounced as/r/) are investigated. Table 4, reported the recognition rates. The results indicated 82.92% average recognition rate Table 4.

Table 4: The recognition rate results for long vowels connected with other letters

Long Vowels	Number of Signals	Recognized Signals	Not Recognized Signals	Recognition Rate [%]
Le لي	300	240	60	80
La لا	300	245	55	81.67
Re ري	300	261	39	87
Ra را	300	249	51	83
			Avr. Recognition Rate	**82.92**

Experimental-3

This experiment studies, short Arabic vowels: fatha: represents the /a/ letter sound and kasra: represents the /i/ sound (pronounced as/e:/) for each vowel 300 signals of 20 speaker results. The recognition rates of above mentioned two short vowels connected with other letter such ل (pronounced as/l/) and ر (pronounced as/r/). The average recognition rate was 81.83% (table 5).

A comparative study of the proposed feature extraction method with other feature extraction methods was performed. Mel-frequency cepstral coefficient (MFCC) (Hachkar et al., 2011), wavelet packet energy index distribution method (WPID) (Wu & Lin, 2009), genetic wavelet packet neural network (GWPNN) (Engin, 2007) and formants and Shannon entropy in conjunction with WP at level seven (SHWPF) (Daqrouq, 2011) are employed for comparison. For all these methods PNN classifier is utilized. The results were conducted for the whole recorded database (1728 signals). The best recognition rate selection obtained was 85.76 for MWNN (table 6).

Table 5**:** The recognition rate results for short vowels connected with other letters

Short Vowels	**Number of Signals**	**Recognized Signals**	**Not Recognized Signals**	**Recognition Rate [%]**
Le لِ	300	230	70	76.67
La لَ	300	247	53	82.33
Re رِ	300	263	37	87.67
Ra رَ	300	242	58	80.67
			Avr. Recognition Rate	**81.83**

Table 6: The recognition rate results for comparison

Rec. Method	**Number of Signals**	**Recognition Rate [%]**
MWNN	1728	85.76
MFCC	1728	84.54
WPID	1728	84.85
GWPNN	1728	83.34
SHWPF	1728	85.13

6. Conclusion

Probabilistic neural network based Arabic vowels recognition system with modular arithmetic was proposed in this paper. This system was developed using a wavelet packet feature extraction method. In this work, effective feature extraction method for Arabic vowels system is developed, which is able to present the Arabic vowels in reasonable short feature vector. Taking in consideration, the computational complexity is very crucial issue for neural network. The experimental results on a subset of recorded database showed that feature extraction method is appropriate for Arabic language recognition system. Our investigation of speaker-independent Arabic vowels classifier system performance is performed via several experiments depending on vowel type. The declared results show that the proposed method can make an effectual analysis with identification rates may reach 97%**.**

Acknowledgment

This paper was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University. The authors, therefore, acknowledge with thanks DSR technical and financial support.

References

1. Alotaibi Y. A., Investigating spoken Arabic digits in speech recognition setting, Information Sciences 173 (2005) 115–139.
2. Amrouche A., M. Debyeche A. Taleb-Ahmed J. M. Rouvaen and M.C.E. Yagoub. An efficient speech recognition system in adverse conditions using the nonparametric regression. Engineering Applications of Artificial Intelligence, 85 -94. 10.1016/j.engappai.2009.09.006.
3. Amrouche A., Rouvaen J. M., 2003. Arabic isolated word recognition using general regression neural network. Proceedings of the 46th IEEE Midwest Symposium on Circuits and Systems (MSCS), Volume 2, December 30-30, 2003, Cairo, pp: 689–692.
4. Berouti M., Schwartz, R., Makhoul J., 1979. Enhancement of speech corrupted by acoustic noise. In: International Conference on Acoustics, Speech, and Signal Processing, pp. 208–211.
5. Bourouba H., Djemili R., Bedda M., Snani, C., (2006). New hybrid system (supervised classifier/HMM) for isolated Arabic speech recognition. Proceedings of the 2nd IEEE International Conference on Information and Communication Technologies (ICTTA'06), Volume 1, April 24-28, 2006, Damascus, pp: 1264–1269.
6. Cormen T. H., Leiserson C. E., Rivest R. L., and Stein C., *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 31.3: Modular arithmetic, pp. 862–868.
7. Daqrouq K. (2011), Wavelet entropy and neural network for text-independent speaker identification, Engineering Applications of Artificial Intelligence 24 (2011) 796–802.
8. Datta, S., Al Zabibi, M., Farook, O., 2005. Exploitation of morphological in large vocabulary Arabic speech recognition. International Journal of Computer Processing of Oriental Language 18 (4), 291–302.
9. Daubechies I.(1988), Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*,41, 909–996.
10. Debyeche M., Haton J. P., Houacine A., (2006). A new vector quantization approach for discrete HMM speech recognition system. International Scientific Journal of Computing 5 (1), 72–78.
11. Engin A., (2007), A new optimum feature extraction and classification method for speaker recognition: GWPNN, *Expert Systems with Applications* 32, 485–498.
12. Hachkar Z., Mounir B., Farchi A. (2011), El Abbadi J., Comparison of MFCC and PLP Parameterization in pattern recognition of Arabic Alphabet Speech,

Canadian Journal on Artificial Intelligence, Machine Learning & Pattern Recognition Vol. 2, No. 3.
13. Hermansky H., Morgan N., Hirsch H. G., 1993. Recognition of speech in additive and convolutional noise based on RASTA spectral processing. In: International Conference on Acoustics, Speech, and Signal Processing, pp. 83–86.
14. Junqua J. C., Mak B., Reaves B., 1994. A robust algorithm for word boundary detection in the presence of noise. IEEE Trans. Speech Audio Process. 2 (3), 406–412.
15. Karray L. and Martin A., Towards improving speech detection robustness for speech recognition in adverse conditions, Speech Communication 40 (2003) 261–276.
16. Khalaf E.F., K. Daqrouq and M. Sherif, 2011a. Wavelet packet and percent of energy distribution with neural networks based gender identification system", Journal of Applied Sciences, 11:2940-2946.
17. Khalaf E., Daqrouq K., Sherif M., 2011b. "Modular Arithmetic and Wavelets for speaker Verification", Journal of Applied Sciences 11: 2782-2790.
18. Kirschhoff K., Novel approach to Arabic speech recognition. Final Report from the JHU Summer School Workshop, (2003). Proceedings of the International Conference on ASSP (ICASSP'03), pp. 344–347.
19. Lazli L., Sellami M., 2003. Connectionist probability estimation in HMM Arabic speech recognition using fuzzy logic. Lectures Notes in LNCS 2734, 379–388.
20. Lei Z., Jiandong L., Jing L., and Guanghui Z (2005). A Nove Wavelet Packet Division Multiplexing Based on Maximum Likelihood Algorithm and Optimum Pilot Symbol Assisted Modulation for Rayleigh Fading Channels, *CIRCUITS SYSTEMS SIGNAL PROCESSING* , VOL. 24, NO. 3, PP. 287–302.
21. Mauuary L., Monn_e J., 1993. Speech/non-speech detection for voice response systems. In: European Conference on Speech Communication and Technology, Berlin, September 1993,pp. 1097–1100.
22. Mokbel C., Jouvet D., Monn_e J., 1995. Blind equalization using adaptive filtering for improving speech recognition over telephone. In: European Conference on Speech Communication and Technology, pp. 141–1990.
23. Mokbel C., Mauuary L., Karray L., Jouvet, D., Monn_e, J., Simonin, J., Bartkova, K., 1997. Towards improving ASR robustness for PSN and GSM telephone applications. Speech Communication 23 (1–2), 141–159.
24. Saeed K., Nammous M., (2005), A New Step in Arabic Speech Identification: Spoken Digit Recognition b.
25. Saeed K., Nammous M., 2005. Heuristic method of Arabic speech recognition. Proceedings of the IEEE International Conference on Digital Signal Processing and its Applications (IEEE DSPA'05), 528–530a.
26. Salam M. , Mohamad D. , Salleh S., Neural Network speaker dependent isolated malay speech recognition system: handcrafted vs. genetic algorithm, in: International Symposium on Signal Processing and its Application (ISSPA), Kuala Lumpur, Malaysia, August (2001), pp. 731–734.
27. Savoji M.H., 1989. A robust algorithm for accurate endpointing of speech signals. Speech Communication 8, 45–60.
28. Selouani S.A., Douglas O., Hybrid architectures for complex phonetic features classification: a unified approach, in: International Symposium on Signal Processing and its Applications (ASSPA), Kuala Lumpur, Malaysia, August 2001, pp. 719–722.
29. Selouani S.A., Caelen J., (1999). Recognition of Arabic phonetic features using neural networks and knowledge-based system: a comparative study. International Journal of Artificial Intelligence Tools (IJAIT) 8 (1), 73–103.
30. Shoaib M., Awais M., Masud S., Shamail S., Akhbar J., (2004). Application of concurrent generalized regression neural networks for Arabic speech recognition. Proceedings of the IASTED International Conference on Neural Networks and Computational Intelligence, (NCI 2004), 206–210.
31. Tayli M. & Al-Salamah A.I. (1990). Building bilingual microcomputer systems. Communications of the ACM (May), 495-504.
32. Wu, J.-D. & Lin B.-F. (2009), Speaker identification using discrete wavelet packet transform technique with irregular decomposition Expert Systems with Applications 363136–3143.
33. Zitouni I. and Sarikaya R., Arabic diacritic restoration approach based on maximum entropy models Computer Speech and Language, 23 (2009) 257–276.

1/27/2014