Auto electronic recognition of the Arabic letters sound

Omar Ibrahim Alsaif[1], Kifaa Hadi Thanoon[1], Asmaa Hadi Al_Bayati[2]
[1]Computer Systems Department, Mosul Technical Institute, Northern Technical University, Mosul, Iraq
[2]Software Engineering Department, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

Article Info

ABSTRACT

In this research Arabic speech sounds have been studied and investigated, so as to find the distinctive features of each articulated sound. Therefore, certain Arabic sound which share certain approximate distinctive significant features have been chosen for study the ability of distinguishing among them through abstracting characteristic features for them. The signals of speech for the sounds have been recorded through the microphone which represented in a binary matrix. This procedure was implemented so as prepare these signals for processing operation through which two features for the co-occurrence matrix (contrast, energy) have been counted. The values of these features were studied and compared from one person to another to discover the certain speech sounds properties sharing certain common distinguishing features approximate in their articulation one another. The results analysis for this study gave the ability of the dependence to these features for distinguish the sound of speaker, in addition to the high ability which provided to distinguish among the arabic letters, where no connect between both co-occurrence matrix elements and the features of signaling of any arabic letters.

Corresponding Author:

Omar Ibrahim Alsaif
Computer Systems Department, Mosul Technical Institute, Northern Technical University
Mosul, Iraq
Email: omar.alsaif@ntu.edu.iq

1. INTRODUCTION

The Arabic language is considered as mother of Semitic languages, as it is basically a balanced language spoken before it was inherited in writing [1]. An important concern has taken to the Arabic letters in its spoken form. The phoneme is the natural means through which all other natural languages are transmitted. The sounds in any language are influenced by each other during the pronunciation process in order to agree in the output or adjective with the neighboring sounds, which leads to the change of some of the exits or their characteristics [2], [3].

Acoustics have many applications in our lives. It is involved in everything related to linguistic sounds and due to the great technical expansion, that we are seeing in this era and the want to use spoken language in communication between machine and man on the one hand and between people on the other hand, the practical phonetics is expanding, and we expect that it will witness a greater expansion in the coming years. expansion of the evolution of the transfer of sound waves, storage and control machines and devices one of the important and hopeful applications of distinguishing Arabic sounds is the development of an automated dictation method so that the phoneme is analyzed and converted into correct writing. Of course, ridding the spoken voice of its impurities and adding what was not spoken completely or correctly to writing is a very important matter [2], [4].

2.   LITERATURE REEVIEW

The "dictionary" which was arranged on the basis of phoneme, as well as Ibn Jinni through his books "Characteristics", "The Secret of the Synthesis of Expression", Ibn Al-Jazari in "Al-Nashar"and Sibawayh in the Book and others. These works are presented to describe the Arab voices in terms of their exits and characteristics. They are presented in terms of change and due to the juxtaposition, modalities of performance, and so on [5]-[7].

The linguistic lesson in the modern era has received a lot of examination, study and analysis, with new curricula that generally differ from what was prevalent in ancient linguistic studies. Modern linguistics deals with language through the study of four levels: the phonemic level, the morphological level, the grammatical, and the semantic [5]. Most of the research and authorship in the grammatical, morphological or semantic aspects, while the least of them deals with the phonemic aspect of the language. This topic need more research and study, as many of issues are still Mysterious and unknown. Desperately need to analyze and interpret it by adopting new methods in order to find commonalities between especially computer software, modern technology, and the Arabic language in writing and speech, this requires more attention to study these phonemes, in order to serve this significant aspect of language study.

Undoubtedly, many earlier studies have dealt with this topic. We mention in the old studies AL Khalil description of sounds in the "Al Ain dictionary, recent studies Kamal Bishr in "Phonetics", Muhammad Ali al-Khuli in "Arab Voices," Abdel-Qader Abdul-Jalil, and Ibrahim Anis in his two different books, which bear one title, "Linguistic Voices," and Saad Maslouh in "The Study of Hearing and Speech." and Muhammad Salih Al-Dali in "The Human Voice", However, most of these studies were not specific in such topic specifically, but were exposed to mention only some chapters refer to all whole [5]. Therefore, an attempt was made as much as possible to make the work focused in using some of the characteristics of the contrast matrix to indicate the Arabic letters spoken in a digital form. Processing by computer, and then study the values of those characteristics and the extent of their variation from one person to another and their proximity to each other within the same group of exits to which they belong.

3.   BASIC CONCEPTS IN PHONETICX

The letter is part of the language analysis, Sound is part of the speech analysis, and the study of phoneme is a necessary as an introduction to the study of the phonemic system [8]. The science that studies sounds of a language in physical aspect without considering s functions is defined by phonology, which is concerned with the study of the sounds spoken in a language, its analysis and classified, including the manner of pronunciation, transmission and perception [9]. Most of the students in the field of phonology divided into three branches: phonological phonology, physical or acoustic phonology, and acoustic phonology, with all its characteristics and field [10], [11], need to guide students in these directions towards the possibility of pushing them towards modern technology.

3.1. Sound components

Sound is produced as a result of ripples in the air and it's only transmitted through a material medium, whether it is air or a solid substance. The verbal sounds are produced in the human voice system through the air leaving the lungs. The lungs act as a blower for the air that as it passes through the throat, nose and mouth to generate the sound. The air is subjected to narrowing and modification in its path, producing different letter sounds [12]. It should be noted that some researchers point out that the common letter exits today in some Arab countries, such as (الضاد ، الظاء ، الجيم) and others. The correct Arabic pronunciation, as described by Arabic language scholars in the second and third centuries AH. The computer may be an important means in correcting this pronunciation and returning it to its predecessor, and even to standardizing the correct pronunciation among all Arabs. Recent studies divide sounds into sound parts made of different types, each of which has a different length that ranges between certain limits [13], [14].

3.2. Primary processing of the audio signal

Many parameters have primary affects to the noise of the acoustic signal through recording operation. The surrounding medium, the high frequencies within the computer, the microphone, the sound card, and the great difference in the sounds of the acoustic signal for each letter representing major characteristics that the sound influence in it. At this stage, speech acts as purified from noise depending on multi periods, while the noise of the environment medium is recorded alone without an audio signal. The shift rate, amplitude rate and the zero pass rate are calculated and recorded for all Arabic letters separately. The previous mathematical criteria are calculated for each letter, and it takes place after this cut of the letter into frames of a fixed size. Finally calculating the amplitude and zero pass rates for each frame separately are produced. The resulting values in calculating these two noise ratios must be compared with the amplitude and

zero pass rates for all frames. If the values are close, there is no word in the frame and it is deleted, and preserved on tires with spread rates and zero passes that are greater than both stretch rates and zero pass rates [8].

4. GRAY LEVEL COXISTENCE MATRIX (GLCM)

Haralick first proposed a gray-scale coexistence matrix, using (24) statistical features. It is a common statistical method to calculate the features of the image, relates to an account of the occurrence for pixel pairs within a certain distance and specific direction. Elements matrix correspond to proportional Gray-value pixel pairs occur. The size of the matrix depends on the maximum intensity of the gray level in the image [15], [16].

Figure 1 is a good example for calculating (GLCM), where matrix represents the original while matrix represent the final value of (GLCM). In GLCM matrix, top and left row of the columns with green pixels' cells, shows (0,1,2,3) value for each pair in matrix [a]. In matrix (a), pixels with Intensity of "1" occurs with pixels "0" in a pair (1,0) with zero orientation (horizontal pixel pair) and one space (adjacent) Pairs), only twice as indicated by a green circle around them the shape. The GLCM value of the pixel pair (1,0) is displayed with an arrow in the GLCM matrix and, therefore, in the GLCM, in position (1,0) the number "2" is occurred. Likewise, other elements of the GLCM are calculated [17]-[19].
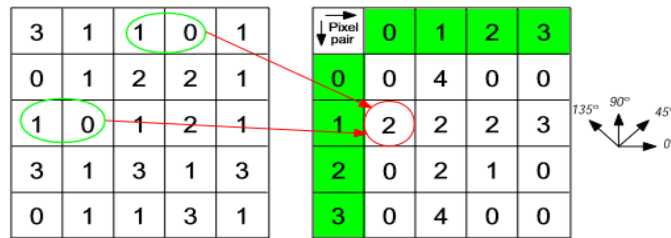

Figure 1. Gray level coexistence matrix calculation example

After generating (GLCM) matrix, you can extract many statistics forms using different formulas. These statistics provide information about the texture of an image. GLCM have good characteristics to provides information texture feature image. The second-order statistical texture feature computation representing the relationship between two pixels for mathematical computation. Features of reference and neighbor pixels are (energy, contrast, correlation, and homogeneity). In this paper, we focus on the properties of energy, and contrast as expressed below [20]-[22].

Contrast feature: returns a degree of contrast between the pixel density and its neighbor in the entire image, and the following equation shows how to calculate contrast [23]-[25].

$$\sum_{i,j} |i - j|^2 \, p(i.j) \tag{1}$$

Energy feature: returns the total elements in the g-l-c-m square. The energy value is calculated by:

$$\sum_{i,j} p(i - j)^2 \tag{2}$$

where $P(i, j)$ gives the statistical probability values for changes between gray levels i and j.

5. THE PROPOSED ALGORITHM

This research deal with a range of sounds for the Arabic letters by lesson and analysis, in order to find distinctive characteristics of each letter sound. Explain the minor parameters of the differences between them, and explain these differences, for the (28) Arabic sounds of twenty-five people were recorded, and each letter was stored in a single file. Figure 2 represent a block diagram has been applied:
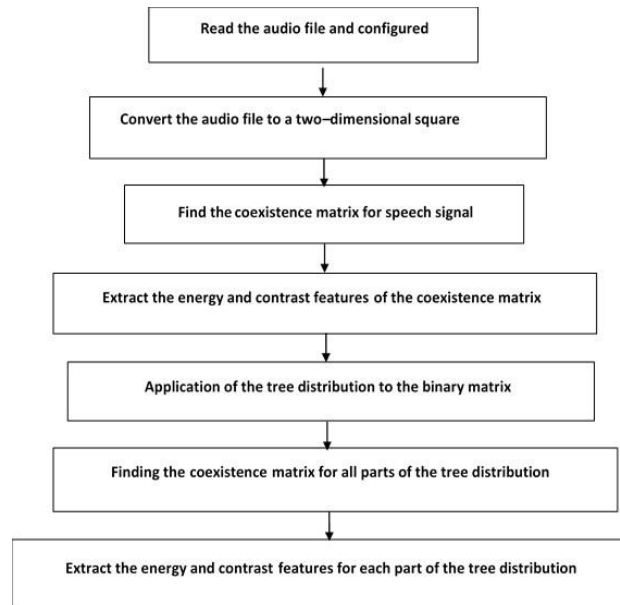
772   □   ISSN: 2502-4752

Figure 2. Represents a block diagram for the proposed algorithm

6.   RESULTS DISCUSSION

The relationship between energy and contrast attributes of a spoken voice have been studied by recording and speaking all the Arabic letters for (25) people. This study illustrates two features of energy and contrast for persons who uttered all the letters of the Arabic language. It was found that despite the great convergence in the exits of the groups of letters but the difference is so obvious for traits that taken to each person, so it can be depended as a good distinguishing for persons.

Appendix (A), shows the strength of adopting the descending sequence for each adjective of Arabic letters spoken by the persons (who were tested on), where the sequence of letters can be depended as a characteristic of the speakers. The figures in Appendix (B), showed that there is no similarity in the two features adopted in the sequence of letters regionally for those characteristics of the coexistence matrix to speech signal. The results in Table 1 and 2 which plotted in Appendix (C), shows that the two features of the energy and contrast of the Arabic letters, in alphabetical order, provide a characteristic report of the person, and they have shown that these characteristics accurately distinguish the speaker.

The results obtained in Appendix (A) are supported in good vision by data in Appendices (B) and (C), in addition to that Appendix (C) accurately describes the emergence of a clear contrast between the two speakers when pronouncing the same letter.

Table 1. The contrast feature of spoken Arabic letters for nine persons

ي	و	هـ	ن	… ….	ذ	د	خ	ح	ج	ث	ت	ب	أ	Per.	Feature
0.58	0.35	0.51	1.05	…	1.85	1.17	2.34	1.34	0.83	0.47	1.75	0.77	0.75	1	
23	48	09	35	….	65	8	13	27	58	24	06	86	33		
0.82	1.16	0.74	1.05	…	1.16	0.96	1.06	0.71	0.72	1.00	1.06	1.27	0.60	2	
71	88	25	38	….	8	73	51	29	1	14	97	79	39		
1.72	1.77	1.24	0.93	…	0.63	0.76	0.82	0.78	0.45	0.59	0.73	0.45	0.47	3	
51	1	07	26	….	81	05	63	95	02	78	91	01	38		
1.65	1.90	2.29	0.76	…	2.06	1.96	1.77	2.10	0.77	1.98	1.82	1.55	0.85	4	
19	8	42	1	….	29	56	91	94	11	45	48	42	45		
1.13	2.45	0.76	2.24	…	0.72	0.92	1.06	1.71	0.32	1.05	0.75	0.55	0.77	5	contrast
23	96	51	79	….	57	47	31	66	27	9	13	54	03		
2.00	3.04	2.89	1.05	…	2.42	2.90	3.46	2.23	0.96	3.69	2.66	3.04	1.60	6	
93	26	54	82	….	44	72	08	68	16	92	38	44	82		
0.44	0.49	1.15	0.32	…	0.61	0.42	0.35	0.36	0.68	0.34	0.32	0.62	0.33	7	
95	99	46	58	….	24	16	52	22	45	54	41	81	22		
3.05	5.28	2.97	2.47	…	2.17	2.35	2.81	3.39	0.61	2.00	3.03	2.03	1.63	8	
61	51	25	75	….	98	29	19	52	62	78	93	33	29		
1.80	2.69	1.18	0.59	…	1.59	1.04	1.24	1.13	0.69	1.44	2.66	1.80	0.42	9	
2	16	42	31	….	45	44	78	44	65	15	93	16	81		

Table 2. The energy feature of spoken Arabic letters for nine persons

| ي	و	هـ	ن	... .	ذ	د	خ	ح	ج	ث	ت	ب	أ	Per.	Feature
0.23 97	0.25 22	0.26 51	0.20 91	....	0.20 14	0.22 01	0.18 22	0.3	0.26 46	0.29 5	0.22 08	0.20 37	0.30 75	1	
0.26 8	0.21 11	0.27 42	0.21 89	....	0.21 4	0.25 05	0.23 73	0.26 87	0.25 54	0.28 19	0.31 9	0.26 57	0.29 34	2	
0.21 82	0.18 3	0.21 52	0.22 73	....	0.23 57	0.23	0.21 25	0.22 44	0.27 55	0.25 26	0.24 27	0.25 74	0.31 89	3	
0.18 02	0.16 14	0.15 2	0.25 06	....	0.18 43	0.15 95	0.20 95	0.16 26	0.19 99	0.24 58	0.23 09	0.21 86	0.24 21	4	
0.24	0.13 36	0.29 5	0.12 67	....	0.28 5	0.23 23	0.25 88	0.23 37	0.31 92	0.28 24	0.26 06	0.33 72	0.30 07	5	Energy
0.17 75	0.15 17	0.15 33	0.20 93	....	0.17 43	0.16 69	0.17 3	0.21 35	0.17 11	0.17 52	0.21 17	0.15 47	0.21 06	6	
0.31 33	0.27 05	0.21 85	0.31 91	....	0.22 21	0.25 69	0.28 39	0.35 48	0.23 69	0.29 77	0.36 07	0.26 13	0.35 37	7	
0.18 28	0.15 47	0.18 41	0.16 44	....	0.11 24	0.14 09	0.15 03	0.15 47	0.26 31	0.19 53	0.24 33	0.19 69	0.32 16	8	
0.15 12	0.15 36	0.19 92	0.24 37	....	0.14 64	0.16 06	0.21 41	0.21 2	0.26 35	0.20 9	0.18 55	0.17 97	0.35 1	9	

As can be observed in Appendix (D), finds the values of the two features at each of the quadrants of the quadrupole tree division. Which based on the idea of adopting the tree division in order to obtain the general average for the approved characteristics of the coexistence matrix. As the results shown the concentrated error in one of the quarters were distributed uniformly within the rest of the matrix. where are noticed that the general average of the matrix as a whole is close to the result of finding the average of the averages of the quarters that make up the coexistence matrix.

7. CONCLUSION

Through the above discussion, we conclude that the algorithm applied to all Arabic letters spoken by a group of people, based on founding the energy feature. It can be adopted in order to distinguish the letter in addition to its adoption in order to classify the letters' exits (i.e. work on distributing letters into classes based on the exits of the letter). While the contrast feature, which can be adopted as a distinctive feature of the speaker's voice, and the possibility of adopting two features together in order to obtain a more accurate distinction.
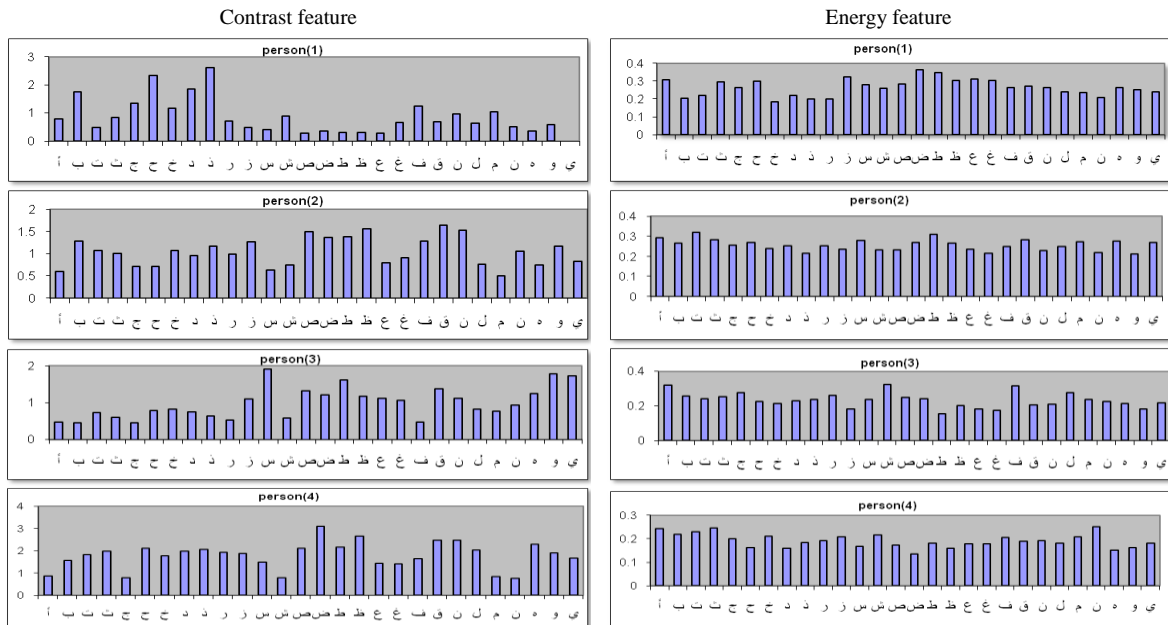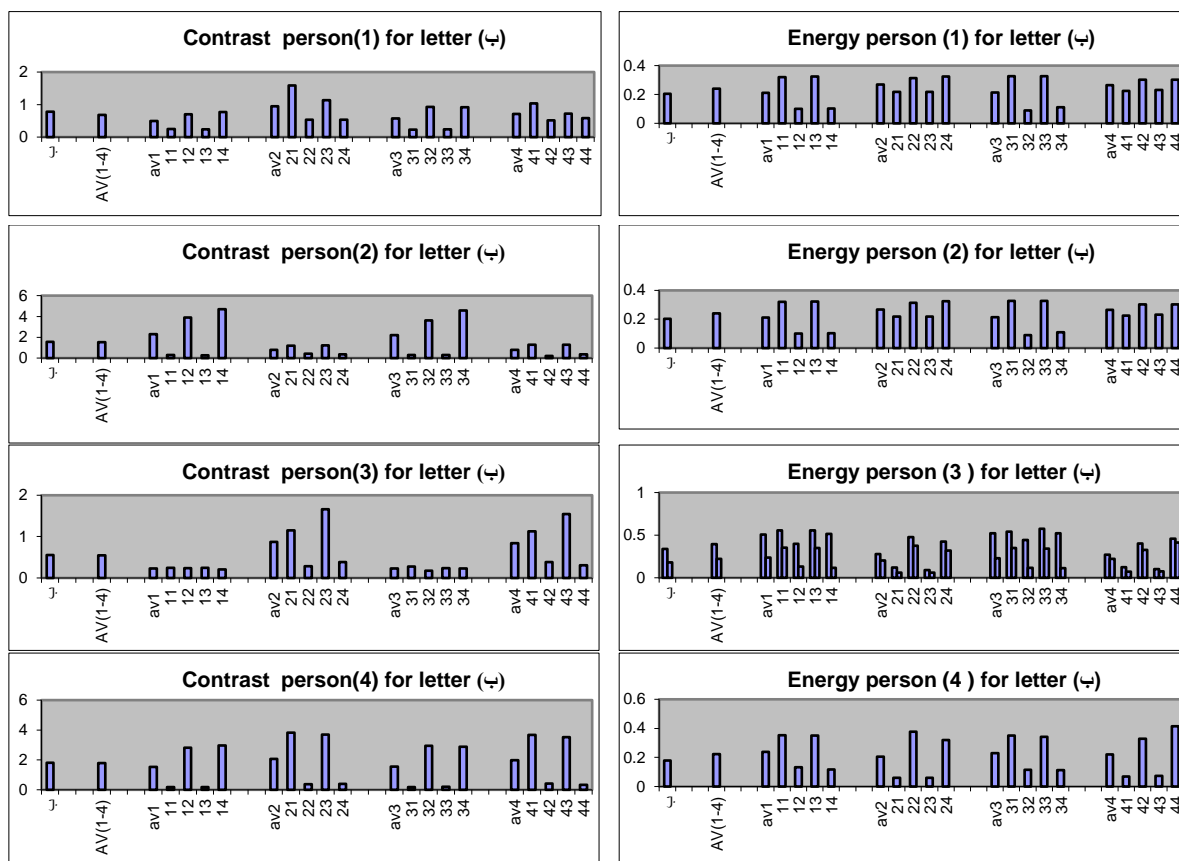
APPENDIX
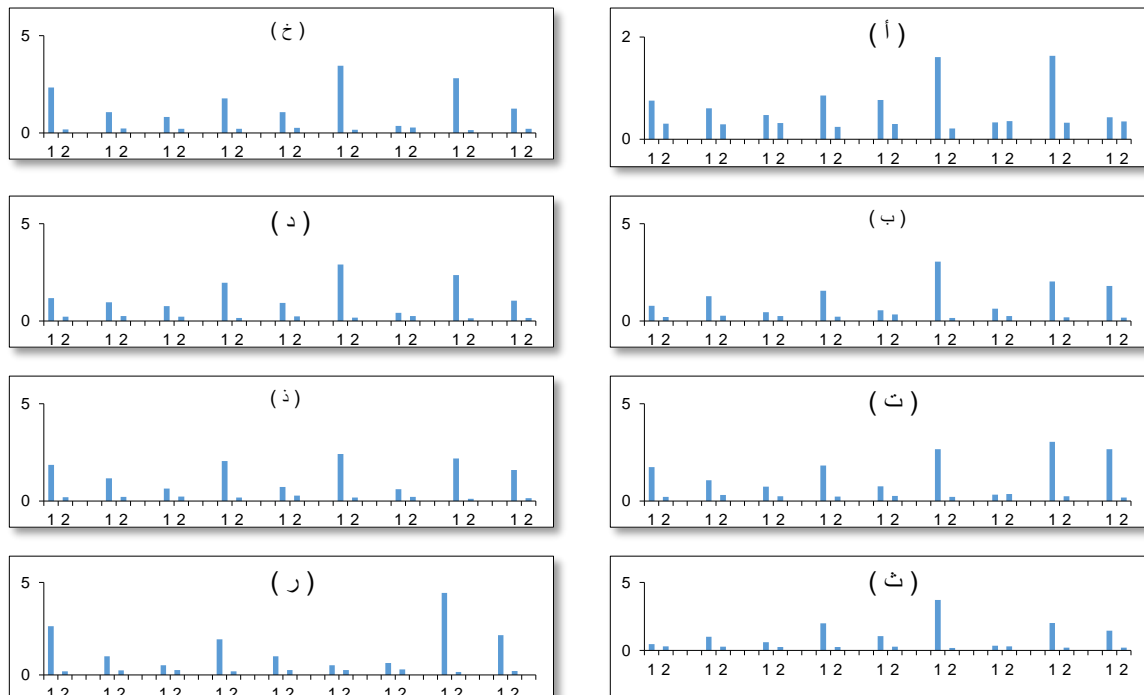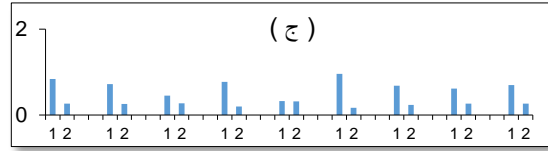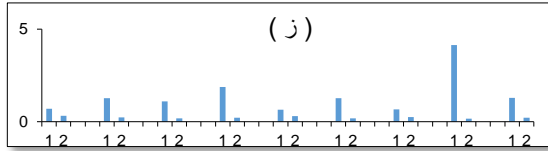
Appendix A. Energy and contrast features of persons who have pronounced all the letters of the Arabic language, where: X-axis (the contrast and the energy features). Y-axis, (value for each one)
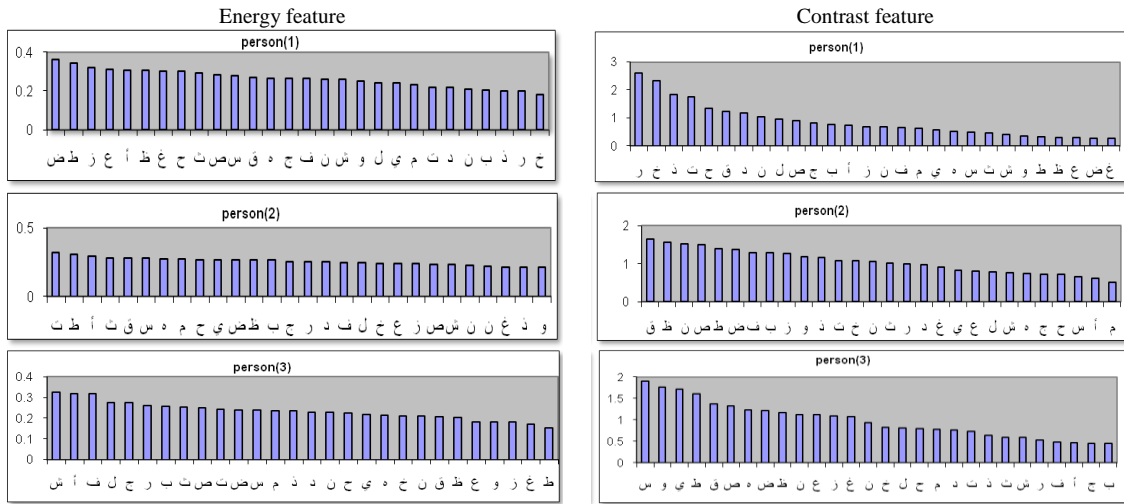
Appendix B. Adopting the descending sequence for each of the two features (energy, contrast) for Arabic letters spoken by the persons, X-axis represents the Arabic letter spoken and Y-axis Energy and contrast values

Appendix C. Features of energy and contrast. The Arabic letters of the region for persons, in alphabetical order, X-axis (Arabic letters of the region) and Y-axis (Contrast and Energy) values

Appendix D. Represents the tree representation of the energy and contrast features of a single letter and four persons and a modifier (the average of the attributes at each level), X-axis (the tree representation of the energy and contrast features) and Y-axis (The feature values)

REFERENCES

[1]   M. Alghamdi, *Arabic Phonetics and Phonology* (in Arabic), vol. 1. Altawbah, 2015.
[2]   M. Z. M. Khedher, "Arabic letter and computing (in Arabic)," Ammaan: Jordan, 2001. [Online]. Available: http://al-mishkat.com/files/arres/arabletter_computing.pdf.
[3]   B. M. Haddad, *Effectiveness of an educational program in the treatment of the problem of confusion between the Arabic letters of similar votes in the second-grade students* (in Arabic)*,* Ketabonline, 2008. [Online]. Available: https://ketabonline.com/ar/books/105014/read?part=1&page=5&index=2962552
[4]   A. A. H. Al-Arif, "Making use of computational linguistics in the service of arabic studies (in Arabic)," *The Journal of The Jordan Academy for Arabic*, vol. 7, no. 3, 2007. [Online]. Available: https://www.asjp.cerist.dz/en/downArticle/26/12/2/32177.
[5]   S. Shahrzad, M. Seyedali, and B. A. Lewis, "Grasshopper optimization algorithm: theory and application," *Advances in Engineering Software*, vol. 105, pp. 30–47, 2017, doi: 10.1016/j.advengsoft.2017.01.004.
[6]   S. A. Sini, "The Arabic language in the computer age, and in formatics, (an approach to the general structure of arabic systems) (in Arabic)," *Association of Arab Universities Journal for Arts*, vol. 1, no. 1, pp. 1–38, 2004.
[7]   A. A. Al-Sinjari Amir and A. M. Kazim, "In The bases of the phonetic method for the Arabic structure presentation and evaluation," *Journal of Kerbela niversity*, vol. 4, no. 8, pp. 1–23, 2012.
[8]   N. Hasan, "Arabic phonetics; Its developments, theory and utilization for teaching Arabic," *Al-Ta'rib, Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, vol. 6, no. 2, 2018. https://media.neliti.com/media/publications/281931--dc419c91.pdf.
[9]   S. 'A. Hanna, K. Z. H. Al-Din, and N. Greis, *Dictionary of modern linguistics English-Arabic*. Hardbound, 1997.
[10]   Al-Zabibi, M., "An Acoustic–Phonetic Approach in Automatic Arabic Speech Recognition," The British Library in Association with UMI. 1990.
[11]   K. H. Thanoon, S. Q. Hasan, and O. I. Alsaif, "Biometric information based on distribution of Arabic letters according to their outlet," *International Journal of Computing and Digital Systems*, vol. 9, no. 5, pp. 981–991, Sep. 2020, doi: 10.12785/ijcds/090518.
[12]   A. M. Al-Subaie and K. A. Al-Shehri, "Automatic spokesperson," 2005. [Online]. Available: https://slideplayer.com/slide/4967770/.
[13]   O. A. Al-Dakak, "Speech analysis and synthesis (in Arabic)," Syria Damascus: The Arabic Encyclopedia, 2005. [Online]. Available: https://phonetics-acoustics.blogspot.com/2015/11/blog-post.html (accessed Feb. 1, 2022).
[14]   S. Q. Alhashmi, K. H. Thanoon, and O. I. Alsaif, "A proposed face recognition based on hybrid algorithm for features extraction," in *2020 6th International Engineering Conference "Sustainable Technology and Development" (IEC)*, Feb. 2020, pp. 232–236, doi: 10.1109/IEC49899.2020.9122911.
[15]   M. Verma, B. Raman, and S. Murala, "Local extrema co-occurrence pattern for color and texture image retrieval," *Neurocomputing*, vol. 165, pp. 255–269, Oct. 2015, doi: 10.1016/j.neucom.2015.03.015.

[16] I. A. Saleh, A. H. Albayati, and K. H. Thanoon, "Measure the software quality based on grasshopper optimization algorithm," *International Journal of Computing and Digital Systems*, vol. 10, no. 1, pp. 955–961, Nov. 2021, doi: 10.12785/ijcds/100186.

[17] C. N. Rao, S. S. Sastry, K. Mallika, H. S. Tiong, and K. B. Mahalakshmi, "Co-occurrence matrix and its statistical features as an approach for identification of phase transitions of mesogens," *International Journal of Innovative Research in Science, Engineering and Technology (An ISO*, vol. 3297, no. 9, pp. 2319–8753, 2007, [Online]. Available: www.ijirset.com.

[18] H. S. Abdallah, W. Yassen, and K. Alsaif, "Feature extraction of images texture based on Co-occurrence Matrix," *Zanco Journal of Pure And Applied Sciences*, vol. 31, no. s2, pp. 60–69, May 2019, doi: 10.21271/ZJPAS.31.s2.8.

[19] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Canadian Journal of Remote Sensing*, vol. 28, no. 1, pp. 45–62, Jan. 2002, doi: 10.5589/m02-004.

[20] D. A. Clausi and B. Yue, "Comparing cooccurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 1, pp. 215-228, 2004, doi: 10.1109/TGRS.2003.817218.

[21] Z. Oleiwi, K. Thanoon, and K. Alsaif, "High frequency coefficient effect on image based on contourlet transformation," in *ICCISTA 2019 - IEEE International Conference on Computing and Information Science and Technology and their Applications 2019*, Mar. 2019, pp. 1–4, doi: 10.1109/ICCISTA.2019.8830649.

[22] A. H. Maray, O. I. Alsaif, and K. H. Thanoon, "Design and implementation of low-cost medical auditory system of distortion otoacoustic using microcontroller," *Journal of Engineering Science and Technology*, vol. 17, no. 2, pp. 1068–1077, 2022.

[23] I. A. Saleh, W. A. Alawsi, O. I. Alsaif, and K. Alsaif, "A Prediction of Grain Yield Based on Hybrid Intelligent Algorithm," Journal of Physics: Conference Series, vol. 1591, no. 1, p. 012027, Jul. 2020, doi: 10.1088/1742-6596/1591/1/012027.

[24] I. A. Saleh, O. I. Alsaif, and M. A. Yahya, "Optimal distributed decision in wireless sensor network using gray wolf optimization," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 4, p. 646, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp646-654.

[25] R. H. Ali and A. S. M. Khidhir, "Enhancement of daytime crescent image using wiener filter based de-blurring technique," in *2021 7th International Conference on Contemporary Information Technology and Mathematics (ICCITM)*, Aug. 2021, pp. 203–206, doi: 10.1109/ICCITM53167.2021.9677768.

BIOGRAPHIES OF AUTHORS

Dr. Omar Ibrahim Alsaif 🆔 🕸 SC Ⓟ he received his B.Sc. in electrical engineering from University of Mosul, Iraq, in 1992. The M.Sc. and Ph.D. degrees in electronics and microelectronic engineering from Mosul University, in 2005 and 2018 respectively. He is currently a lecturer in the Mosul Technical Institute/Northern Technical University in Mosul/ Iraq. His research interests include microelectronic, solid-state systems, renewable energy and nanotechnology devices. He can be contacted at email: omar.alsaif@ntu.edu.iq.

Kifaa Hadi Thanoon 🆔 🕸 SC Ⓟ she received her MSc. degree (in image processing from the University of Mosul, Iraq. Lecturer in 2019. She is lecturer at the Dept. of Computer Systems Techniques, Technical Institute/Mosul, Northern Technical University, IRAQ. Her interesting in Image and signal processing, neural networks. She can be contacted at email: kh.thanoon@ntu.edu.iq.

Asmaa Hadi Al_Bayati 🆔 🕸 SC Ⓟ born in Mosul/Iraq in 1977. She received her bachelor's degree in chemistry in Iraq, Mosul University In 2000, and in 2007 She received her bachelor's degree in software engineering sciences from Mosul University. Her MSc. degree (in software engineering) from the same university in 2013. Her interesting in Software engineering tools and information technology. She can be contacted at email: asmashade77@uomosul.edu.iq.