

Advance Natural Arabic Communication Recognition Uses a Speech Recognition Approach

Mohammed Basil Abdulkareem

Department of Engineering of Computer Technology,
Al-Maarif University College, 31001 Ramadi, Anbar, Iraq

Abstract: In the era of natural computing, automatic recognition of effective natural communication database for recognition is having major challenges. Human computer interface is always proved to be stimulating area in the speech and the language recognition domain. Speech is the important choice of man-machine communication. The objective of the speech recognition is to cultivate statistical methods and its use in real time automatic speech recognition systems. The statistical model of speech recognition plays an important role in the development of real time system. Arabic is a global language for Iraq country and has universal acceptance. This is an official language at many institutes in Iraq. This study addresses the implementation of an Advance Speech Recognition system (ASR) for an Arabic language towards identification of natural communication aspects. The statistical modeling and implementation of the system is done in this study. This research contributed the development, Arabic speech database for isolated (character, digit, word) and connected sentences which are used day to day activity in Arabic community. In this study, 10 subject databases were developed using the PRAAT Software in an office environment. The size of collecting database is consisting of 1200 samples of isolated words and 800 samples of connected sentences, the size of the database is 2000. The experimental analysis of Arabic natural communication is tested using the Mel Frequency Cepstral Coefficient (MFCC) feature extraction technique and dynamic time warping algorithm. The isolated natural recognition system gives 94% accuracy and continuous spoken natural communication proves the 87%. The Natural isolated recognition system gives a robust and dynamic performance than continuous speech recognition. This natural communication based emotional spoken database will be used for the emotive aware cloud computing and spoken interface domain.

Key words: Arabic language, speech, recognition, natural, Mel Frequency Cepstral Coefficient (MFCC), Dynamic Time Warping (DTW), classification

INTRODUCTION

Spoken language is that the most natural and economical variety of shift info between human. Speech recognition may be a technology that allows a knowledge processor to spot the words that an individual talks into a mike or phone. The spoken communication is the group of various operations that changes the sequences of information in language. It is the emerging trends in the signal processing and machine learning (Gaikwad *et al.*, 2010; Padmanabhan and Picheny, 2002). The goal of speech recognition is to develop methodology and real time automatic speech recognition systems. Automatic speech recognition has a motivating and helpful implementation in knowledgeable systems, a technology, whereby the PC act as a replacement for a personality's knowledgeable. Speech recognition systems facilitate users who in a technique or the opposite can't be ready to use the standard Input and Output (I/O) devices (Padmanabhan and Picheny, 2002). The person machine

interface mistreatment speech recognition has useful ways that modify the visually impaired and PC laymen to use the updated technologies (Raj and Susmitha, 2005). The operating diagram of the automated speech recognition is delineated in Fig. 1. The structure diagram illustrated the primary steps of speech recognition such as database collection, preprocessing, feature extraction and sorting or matching (Alghamdi *et al.*, 2007).

The communication among the human existence is generally done by vocally, thus, it's spontaneous for folks to expect speech interfaces with electronic computer (Shrishrimal *et al.*, 2012). A changed feature extraction technique is employed within the Linear Discriminate Analysis (LDA) for face and signature conjointly. Which is used two methodologies to achieve high performance as extract form and texture for image with the comparative with alternative work. Other researchers apply it on speech recognition yet at the feature extraction stage (Awang *et al.*, 2013, Talab *et al.*, 2013). Automatic speech recognition is the mechanism of obtaining hold of

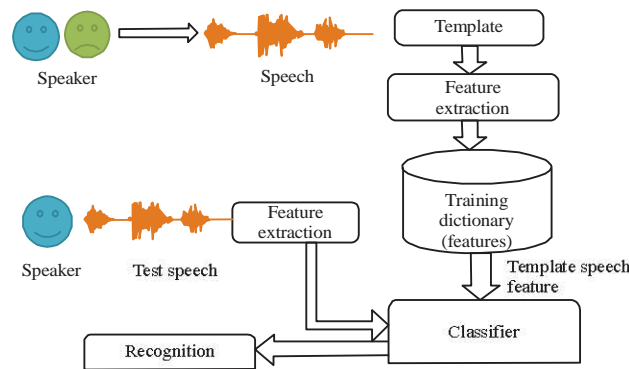


Fig. 1: Working of automatic speech recognition

AN vocalization of speech signal as AN input, captured by a mike and convert it into a text sequence as shut as potential to the spoken information (Aggarwal and Dave, 2011). The researchers try to evolve a theme which may record, interpret and perceive human language. The usage of speech for interacting with the pc could facilitate the developing nations because the language technologies are placed through within the e-Governance theme (Huang *et al.*, 2001).

It is well-known that the speech signal not solely takes the linguistic data (the message) however conjointly an excellent deal of information round the speaker himself: gender, age, social and regional origin, health and emotion and with a rather robust liableness, his individuality. For the popularity of speech recognition, the variability in speakers varies the performance of the system (Garvin and Ladefoged, 1963, Nolan, 1983).

The Arabic is the prescribed language of twenty-six countries and among the six official languages of the United Nations. It's spoken by over 300 million, that makes it one of the ten most widely-spoken languages within the universe (CIA., 2008). Despite that, the literature on machine-controlled Arabic speech recognition is restricted compared to alternative terminologies. This allows a shortage of conducting studies and a scarcity of depth in these surveys and public resources. The unavailability of speech corpus benchmark encourages to researcher for peruse research in this domain. Moreover, these corpora tend to be minor and not shared at intervals the community, preventing researchers from creating valid comparisons between the revealed arrangements. The work done of another language like English, Mandarin, Marathi and Hindi has reached an expanded degree. As Arabic speech is our communication language for day to day activity. This issue driven North American country is to require the Arabic speech recognition for natural communication understanding. The impact of this addition in academia and therefore, the

industrial world is the main motivation carry out this study. The target of this study is the implementation of AN Advance Speech Recognition system (ASR) for AN Arabic towards identification of natural communication aspects. During this investigation, connected analysis on the deserves of Arabic speech recognition was compactly given.

The Semitic incorporates a nature of the Semitic language of constant family of Hebrew and Aramaic (UNESCO., 2017). Arabic is presently the sixth most, generally, speech communication within the universe. The approximate variety of Arabic speakers is 250 million of that regarding 195 million square measure primary speakers and 55 million square measure secondary speakers. Arabic is a political candidate language in additional than twenty two states. Since, it's likewise the spoken communication of spiritual instruction in Islam, more speakers have a minimum of a passive data of the words. Arabic phonemes contain 2 distinctive classes, that square measure named tubular cavity and emphatic phonemes (Al-Zabibi, 1990; Alghamdi *et al.*, 2007). There square measure several dialects of a language like trendy customary Arabic, Egyptian Arabic, Gulf Arabic, Maghrebi Arabic, Levantine Arabic and plenty of others. Additional or less of those accents square measure, therefore, dissimilar from one another that speakers have a tough time understanding the opposite. Arabic is additionally turning an additional in style language to study within the Western world, even if Arabic synchronic linguistics is typically terribly tough to be told for native speakers of Indo-European languages (Gawali *et al.*, 2010; Karam and Campbell, 2007). The Arabic alphabet is employed in many languages, like Persian, Urdu and Malay (Huang *et al.*, 2001; Satori *et al.*, 2009). Analysis on the Semitic has principally targeting trendy customary Arabic that is employed throughout the media, courtrooms and educational establishments in Arab countries. Analysis

work has been rumored on developing ASR was in non-standard speech and conversational Arabic at intervals the 1997 federal agency benchmark analysis and additional recently on the popularity of colloquial and non-standard speech (Tomokiyo *et al.*, 2003). Image fusion is that the methodology of merging 2 pictures into one image. The obtained single amalgamate image exploitation numerous multi-modality medical pictures is increased anatomical, extremely fascinating spectral info compared to the raw single scanned image (Abdulkareem, 2018).

The analysis of speech recognition system development has made the extent layer. The study of speech recognition desires the efforts towards the expansion of regional language speech recognition. Evolution of Arabic speech recognition could be a multidiscipline effort which involves integration of Arabic phonetic, Arabic speech process techniques and natural speech process. The numbers of researcher's square measure addressing the Arabic speech recognition effort towards the isolated, connected and continuous speech modify application development (Elshafei, 1991; Alghmadi, 2003; Alghamdi, 2001; Alghamdi *et al.*, 2002).

The maid and dynamic tool for the Arabic speech recognition was developed by the scientist (Myers and Rabiner, 1981). The analysis towards the multidisciplinary effort with the combination of Arabic phonetic recognition techniques and natural speech process has been spoken by the scientist Elshafei (1991) Alghmadi (2003), Alghamdi *et al.* (2007), Elshafei *et al.* (2002), (2006). The Arabic speech recognition exploitation HMM primarily based Sphinx tool has been wiped out Carnegie altruist University's (CMU) and also the scientist rumored the 99.21% accuracy for the Arabic isolated speech info (Abushariah *et al.*, 2012; Hyassat and Zitar, 2006).

MATERIALS AND METHODS

The speech info plays a big performs within the development of automatic address recognition system (Gawali *et al.*, 2010). The speech info is volunteered by speakers resides from Iraq. The planning or the Arabic natural communication, corpus includes the isolated, connected and continuous sentences. All vocabularies are elect on the idea of day to day operating isolated words and sentences. The overall variety of speakers was ten, five females and 5 males. Every info is recorded victimisation 5 utterances level. The vocabulary size of the info consists of 1200 isolated words and 800 connected sentences. The complete size of gathering up speech info is 2000 samples. The actual technical specification of the info is given in Table 1.

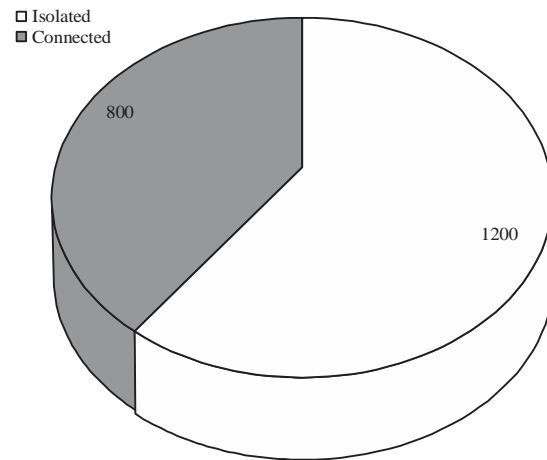


Fig. 2: Graphical description of vocabulary of collected Arabic speech corpus

Table 1: Detail technical specification of Arabic speech corpus

Parameters	Values
Speaker classification	Native Iraqi
Gender type	5 male, 5 female
Mother- tongue	Arabic
Accent	Iraqi
Age group	20-40
Region	Iraq
Environment	Office environment (closed room)

The graphical representation of the collected Arabic speech corpus for natural communication, understanding is presented in Fig. 2. Vocabulary of Arabic speech recognition is founded on the following scheme. The each sample of the Arabic speech word and sentences has been registered for 5 utterances level. The detail specification is presented in following Table 2.

Acquisition setup: To attain a high audio quality, the recording passed off within the traditional space while not screeching sound and result of echo. The frequency for all recordings was set to be 41500 rates at the area temperature and traditional humiddness. The speakers were asked to take a seat ahead of the mike with the space of concerning 8-12 m. The speech information was collected with the assistance of mike reattach and MATLAB Software system for the single channel recording. The pre-processing was finished the assistance of Praat Software system. The technical parameter thought of for recording is delineated in Table 3.

Segmentation of speech database: The recorded information is segmental and cropped using Praat tool. It is a package program for the analysis and reconstruction of acoustic speech signals. Praat is employed for corpus labelling and annotation (Alghamdi *et al.*, 2007). This allows labelling of designated segments and not essentially, all the sections in an exceedingly file. This

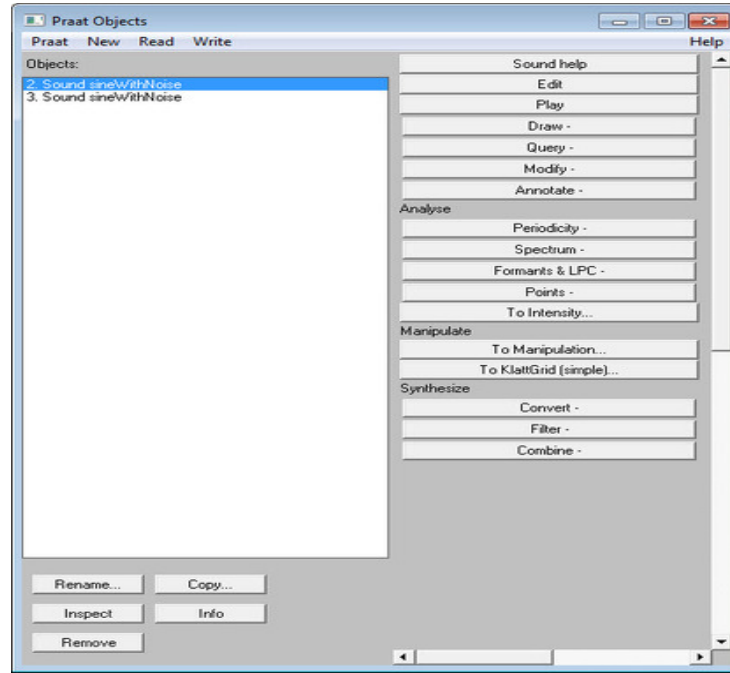


Fig. 3: The Praat object graphical view

Table 2: Characteristics of Arabic speech corpus

Type of speech corpus	No. of speakers	No. of sentences/word	Utterances	Size of corpus
Isolated	10	24	5	1200
Connected	10	16	5	800
Total size of corpus				2000

Table 3: Parameter used for the database creation

Parameters	Values
Sampling rate	41500
Speakers	Dependent
Condition of noise	Normal
Accent	Arabic
Pre-emphasis	1-0.9/(z-1)
Window	Hamming in nature, 25 milliseconds
Window base	20 millisecond

characteristic is particularly helpful for segmentation of diaphone like units in carrying. The graphical illustration of the study read of Praat is shown in Fig. 3. The speech segmentation and speech cropping is delineated in Fig. 4.

Methodology of Arabic speech recognition: For the planning and development of natural communication understanding feature extraction and have matching is plays a crucial role. For this analysis work we've got use Mel Frequency Cepstral Constant (MFCC) for feature extraction and Dynamic Time deformation (DTW) for matching the options. For the ASR system, the central frequency and detail of speech organ pulse isn't vital for recognizing the speech sample. Rather, the foremost necessary data could be a feature choice. In prescript to

extract feature for isolated word many ways and rule has been lined in literature. The MFCC is greatest and dynamic technique for feature extraction (Gawali *et al.*, 2010; Gaikwad *et al.*, 2011). The MFCC steps square measure sturdy. The Cepstral constant has the very helpful property that causes the variance of the distinction constant attended be unrelated (Jelinek *et al.*, 1975). The equation one describes a mathematical model MFCC:

$$c[n] = \sum_{n=0}^{N-1} \log \left(\sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} kn} \right) e^{j \frac{2\pi}{N} kn} \quad (1)$$

Where:

$c(n)$: Cepstral coefficient

$x(n)$: The input signal

The MFCC is the most popular feature extraction technique for ASR (Myers and Rabiner, 1981). The effectiveness of recognition system is directly relative to feature set selection. The human voice production system, the vocal tract is an important organ which reflects spectrum of language (Gawali *et al.*, 2010). The detail explanation of MFCC features extraction is described in Fig. 5.

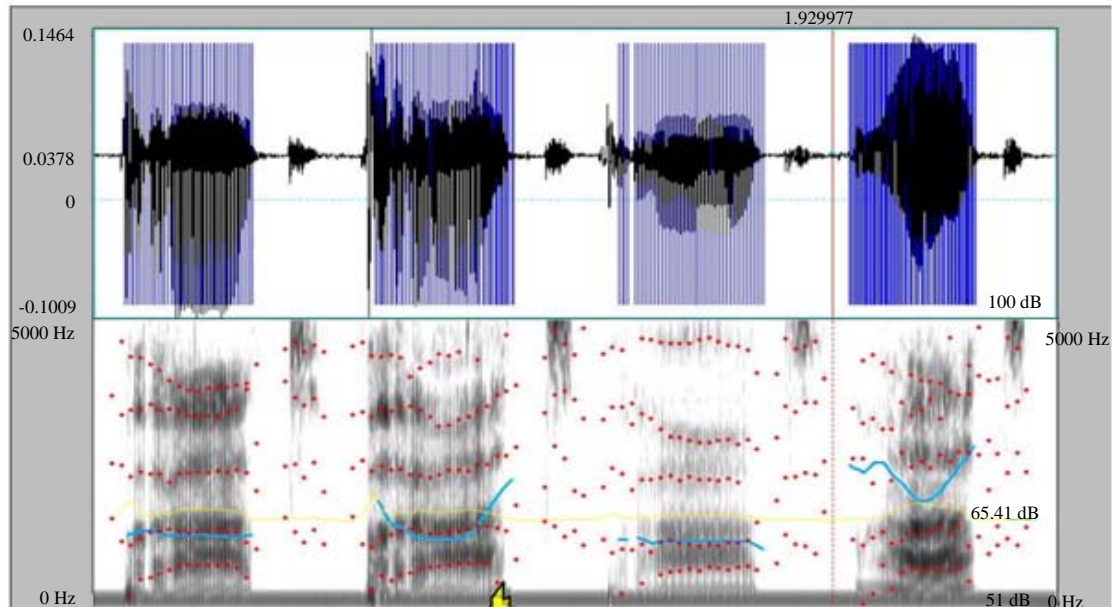


Fig. 4: Speech segmentation and cropping using the Praat tool

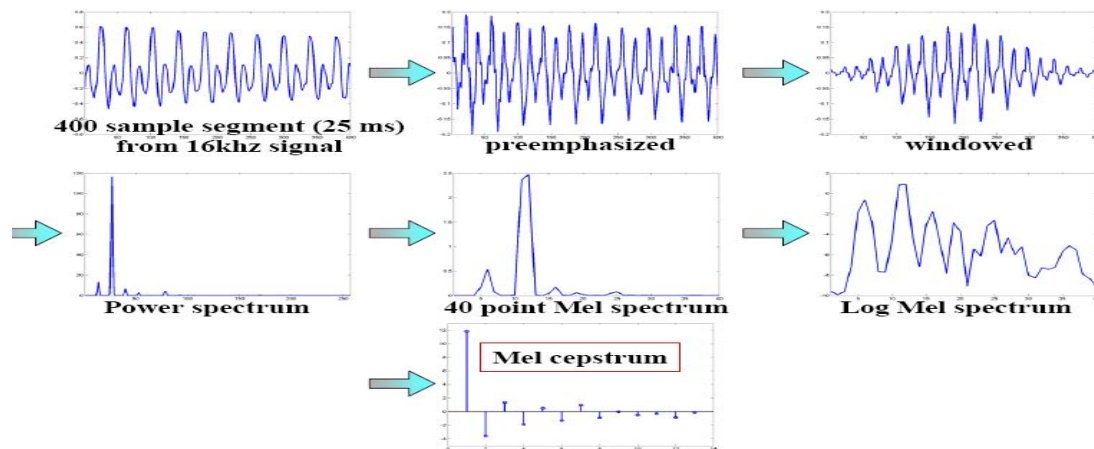


Fig. 5: Basic steps of MFCC feature extraction

Dynamic Time Warping (DTW) is the dominant and famous technique to find an optimal alignment between two given sequence under same condition. The sequence is warped in an un-similar fashion for matching the common rule. The Dynamic Time Warping (DTW) approach used in spoken language recognition for comparing the different speech pattern on the basis of distance measure (Gawali *et al.*, 2010). The variation of dynamic time warping in test and train template is described in Fig. 6. The dynamic time warping approach gives an efficient end product in a warped path as well as unwrapped path. The test and train pattern of connected speech corpus is shown in Fig. 7.

RESULTS AND DISCUSSION

The preprocessing of the dataset has been done using Matlab and Praat Software. Within the preprocessing, parameter of speech signal like pitch, energy contour, frequency contour, FFT water was taken. The voiced amount of speech signal was known and gave to any or all experiments used for preprocessing of speech signal. The detail FFT water Model and pic model of isolated sample word represented in Fig. 8 and 9 severally. The fundamental Frequency contour is shown in Fig. 10.

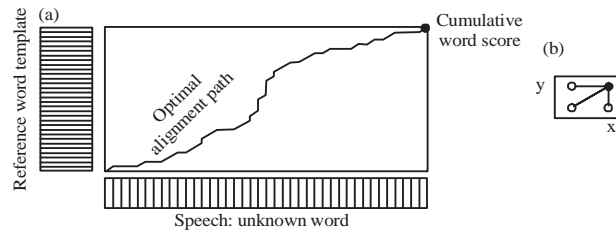


Fig. 6(a, b): Alignment path of dynamic time warping

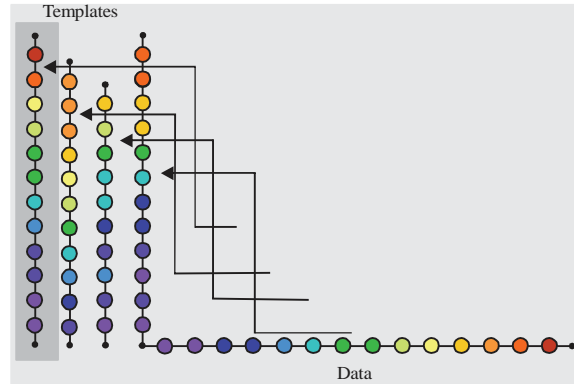


Fig. 7: Alignment of test and train pattern of connected speech corpus using DTW

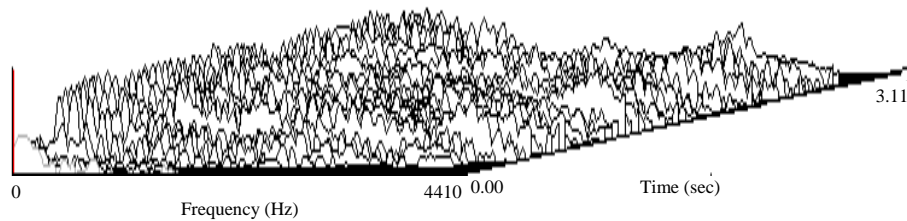


Fig. 8: The FFT waterfall model of sample isolated word

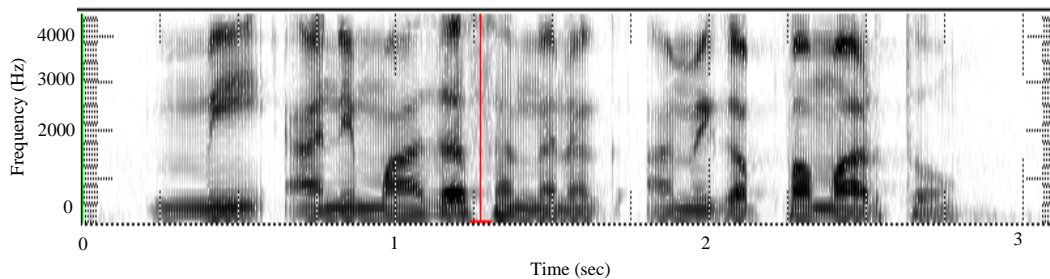


Fig. 9: Spectrogram variation of sample isolated word

The Arabic isolated word speech database sample test file is described in the Fig. 11. The connected speech sample of Arabic speech recognition test sample is shown in Fig. 12. In this study, 2000 sample have been tested of all created speech corpus but the results of only one test

sample graphical manner is shown here. The 12 MFCC feature are calculated using the MFCC approach. All extracted MFCC features were passed to the DTW approach. It matches the extracted features. The alignment of the feature pattern is an important

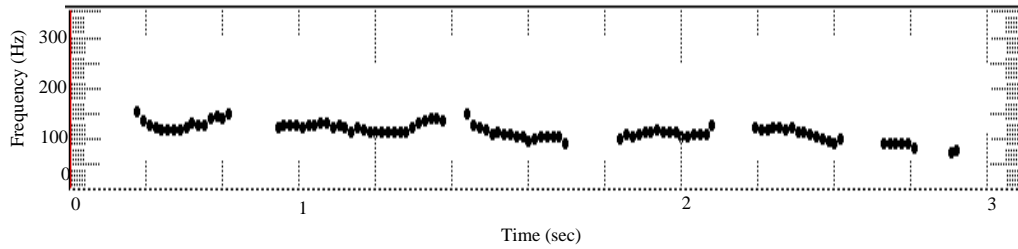


Fig. 10: Frequency contour of sample isolated word

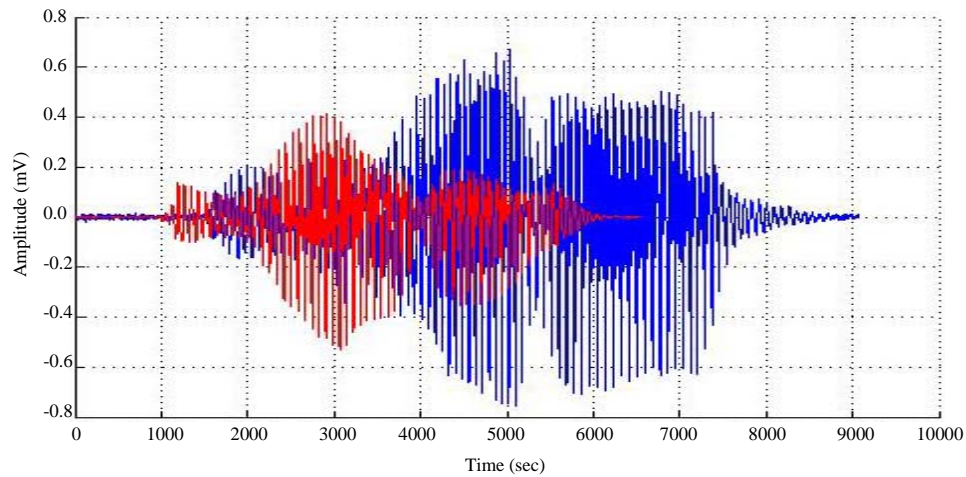


Fig. 11: Visualization of sample of isolated Arabic test speech sample (Original disaligned waves)

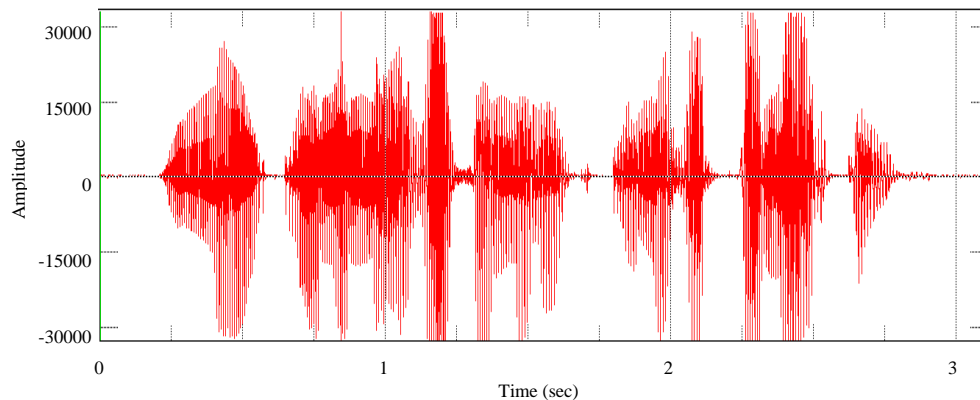


Fig. 12: Arabic speech corpus connected speech sample

factor in speech accent identification. The alignment of two test and train isolated sample speech using DTW approach is shown in Fig. 13. Figure 14 illustrate the extracted path by DTW for two test and train isolated sample speeches. The main contribution of this experiment is finding the unmatched path distance and implemented it for Arabic speech recognition. The extracted path of isolated test and train speech sample is

described in Fig. 15. The unmatched path distance for two different test and train isolated speech samples is shown in Fig. 16.

The unwrapped path from train and test of two isolated speech corpus using DTW is shown in Fig. 17. In the connected speech recognition applied as continuous counter-part of DTW. This study tried to solve the minimization problem faced due to recursion of DTW.

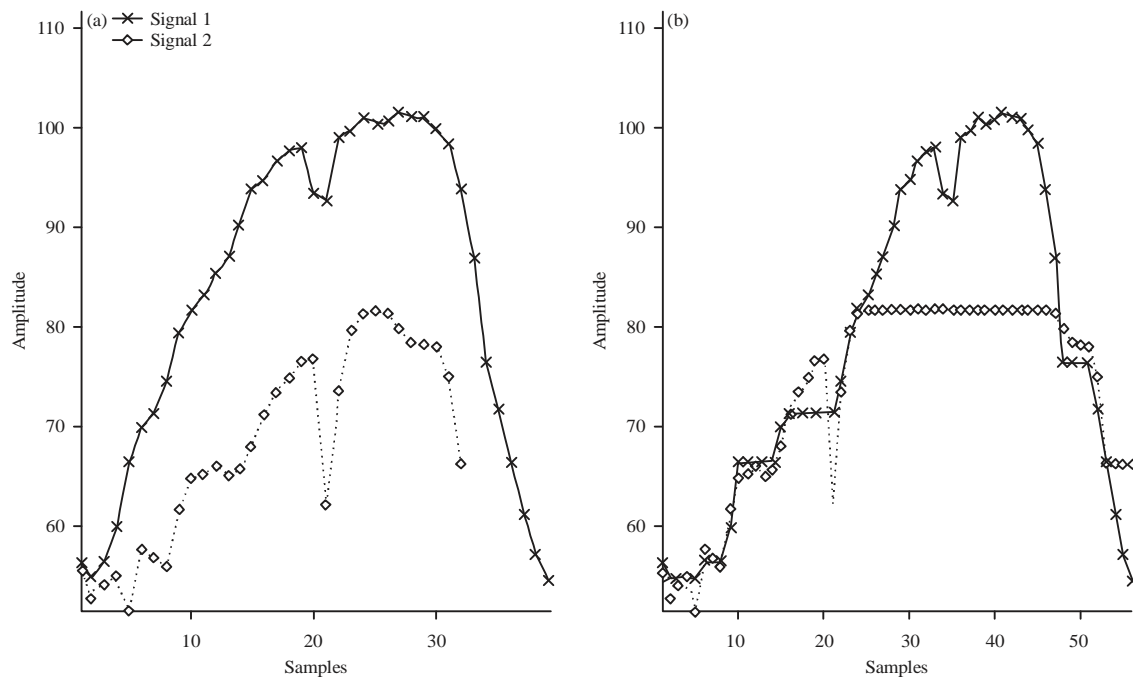


Fig. 13(a-b): Visualization of overlapping the different train and test sample of isolated speech sample using DTW (a) Original signals and (b) Warped signals

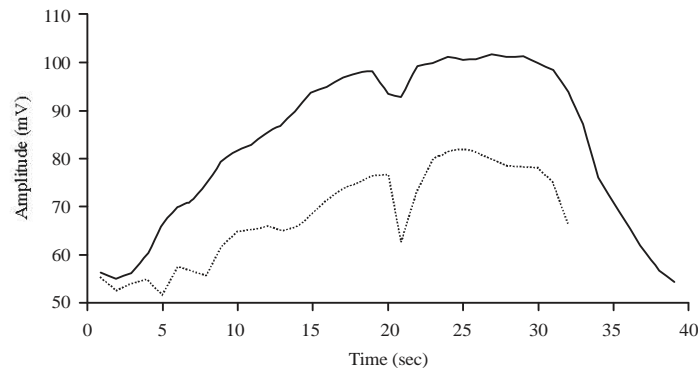


Fig. 14: Alignment and misalignment of different train and test sample of isolated speech corpus using DTW (Original disaligned waves)

Representation of overlapping for two connected train and test samples by DTW is shown in Fig. 18. The alignment of two train and test connected speech corpus samples is shown in Fig. 19. The representation of unwrapped path for two train and test connected speech corpus samples is represented in Fig. 20.

The unwrapped path of various take a look at and train connected speech sample exploitation DTW is delineate in Fig. 17. During this analysis, the performance of isolated word recognition system and connected speech recognition system has been done. The collected Arabic speech corpus has 5 male and 5 feminine speakers. This research tends to test the experiment on the premise of

male and feminine issue. The performance of the system is delineating in Table 4.

The graphical representation of isolated and connected speech corpus is shown in Fig. 21. The isolated word recognition system is tested using the male and female parameter. The graphical representation of the male and female parameter based isolated word recognition is shown in Fig. 22. The tested connected speech recognition using male and female parameter on the basis of male and female parameter is shown in Fig. 23.

The experiment outcome of the research shows the following observations. The recognition of isolated and

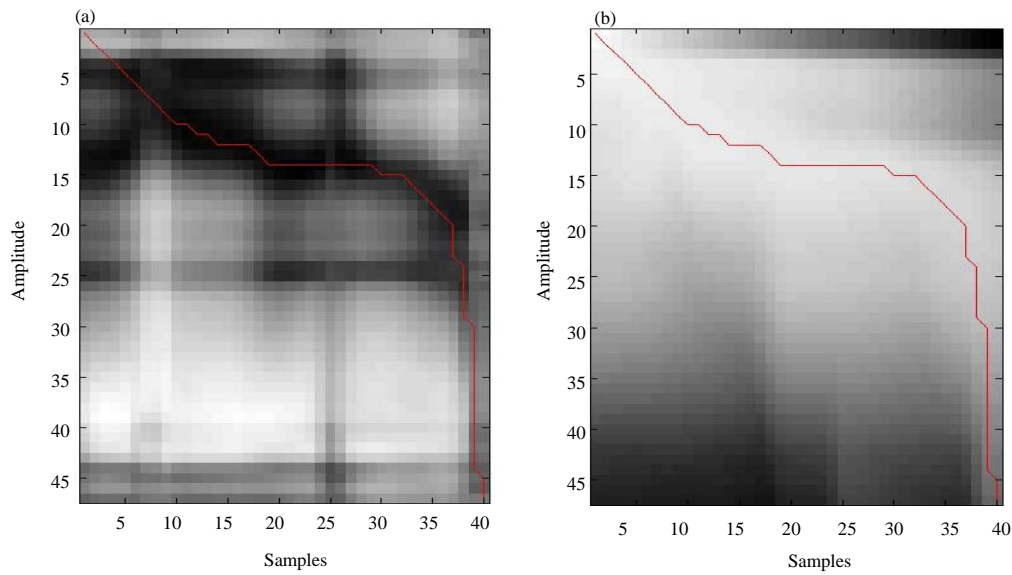


Fig. 15(a-b): Exacted path of different train and test sample of isolated speech corpus using DTW

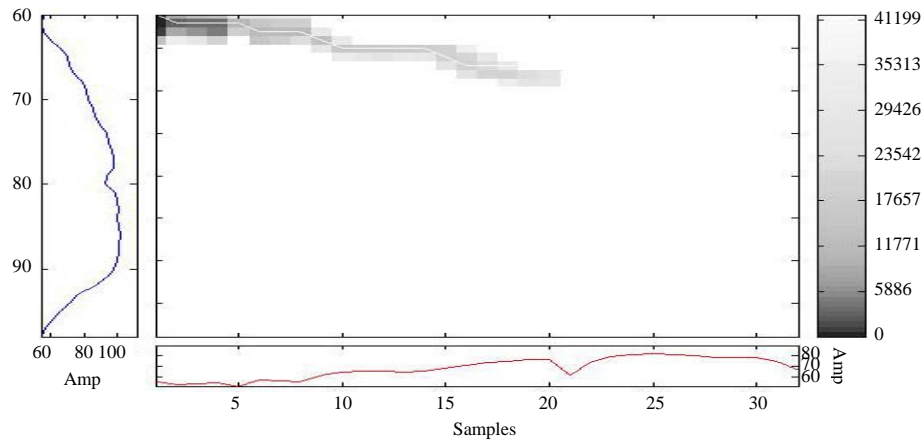


Fig. 16: Unwrapped path for train and test of isolated speech corpus using DTW

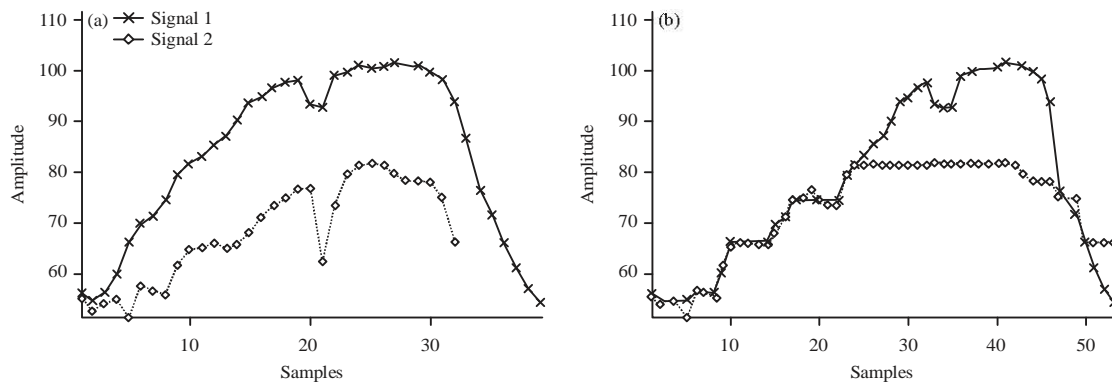


Fig. 17(a-b): Warped signals visualization of overlapping the train and test connected speech sample of Arabic speech corpus using DTW (a) Original signals and (b) Warped signal

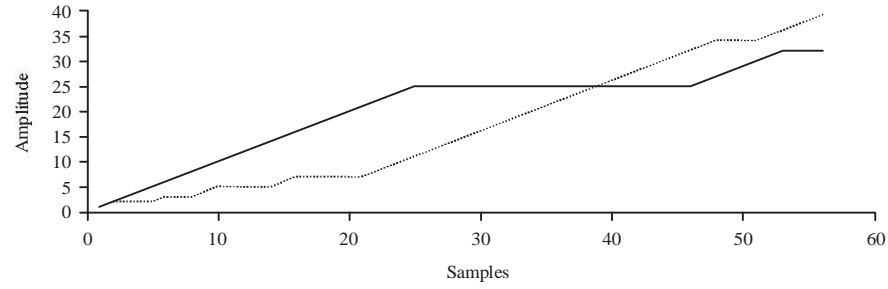


Fig. 18: Alignment and misalignment of different approach of train and test of connected speech sample using DTW

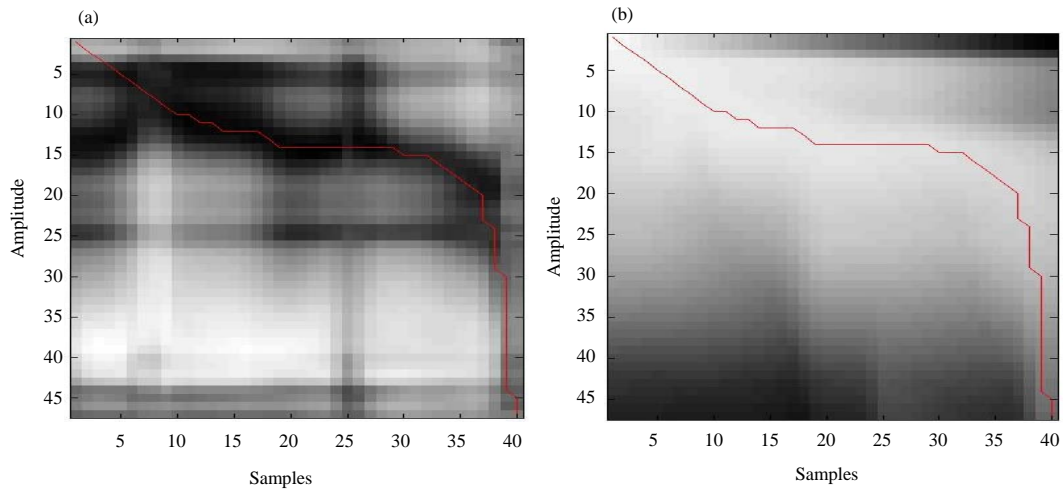


Fig. 19(a-b): Extracted path of two unequal train and test sample

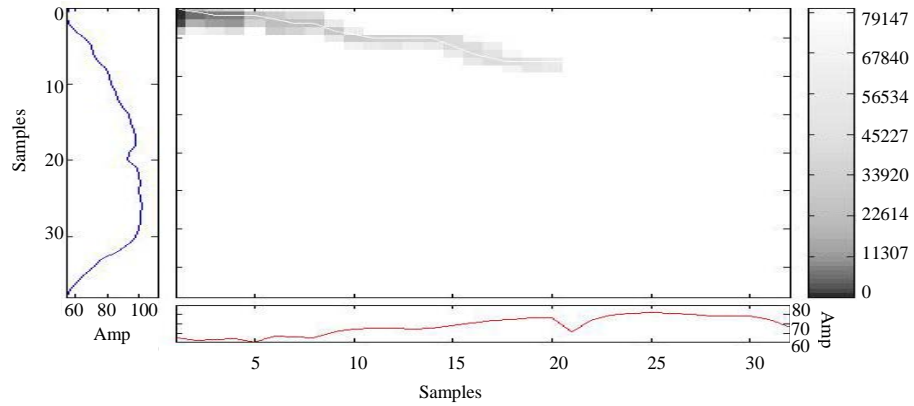


Fig. 20: Unwrapped path of different train and test connected speech sample using DTW

Table 4: Performance of Arabic natural communication system using MFCC and DTW

Type of recognition	Type	Accuracy (%)	Error rate	Average accuracy (%)
Isolated	Male	93	07	94
	Female	95	05	
Connected	Male	86	14	87
	Female	88	12	

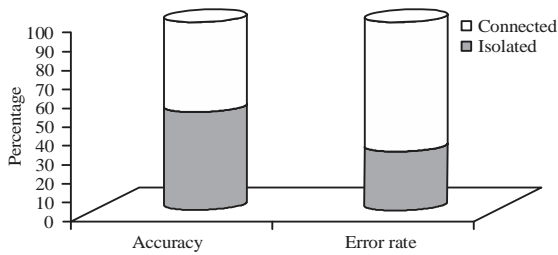


Fig. 21: Performance of isolated and connected speech corpus

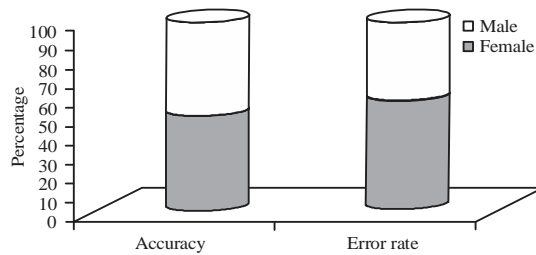


Fig. 22: Isolated word recognition using male and female parameter

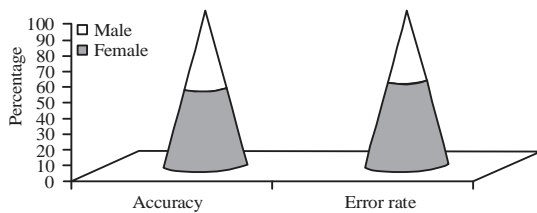


Fig. 23: Connected speech recognition using male and female parameter

connected natural Arabic speech communication recognition is implemented using unwrapped path and found to be effective.

- The experiment of isolated and connected system is conducted for gender group
- As the dominant variation is observed in frequency and pitch in female as compared to male subject, DTW gives better performance for male speaker in isolated and connected speech recognition system
- This study recommends that for advance natural Arabic communication understanding the MFCC in combination with DTW is preferable good choice

CONCLUSION

This study presents the MFCC and DTW approach towards advance natural communication system for

Arabic language. This research proposed the detail guideline for Arabic isolated and connected speech corpus development. The preprocessing of the collected corpus is done using the Praat and Matlab Software tool. The experiment tested the Arabic speech corpus using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) approach. The Arabic spoken corpus has been collected using 10 speakers in which 5 male and 5 female speakers. The overall size of collecting a speech corpus is 2000 sample. As the dominant variation is observed in frequency and pitch in female as compared to male subject, DTW gives better performance for male speaker as an isolated and connected speech recognition system. The performance towards isolated Arabic spoken database is 94 and 87% for the connected sentence corpus. The researcher recommended that the MFCC and DTW approach is the best choice for the Arabic natural communication understanding system.

REFERENCES

- Abdulkareem, M.B., 2018. Design and development of multimodal medical image fusion using discrete wavelet transform. Proceedings of the 2018 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT'18), April 20-21, 2018, IEEE, Coimbatore, India, pp: 1629-1633.
- Abushariah, M., R. Ainon, R. Zainuddin, M. Elshafei and O.O. Khalifa, 2012. Arabic speaker-independent continuous automatic speech recognition based on a phonetically rich and balanced speech corpus. Int. Arab J. Inform. Technol., 9: 84-93.
- Aggarwal, R.K. and M. Dave, 2011. Acoustic modeling problem for automatic speech recognition system: Conventional methods (Part I). Int. J. Speech Technol., 14: 297-309.
- Al-Zabibi, M., 1990. An acoustic-phonetic approach in automatic Arabic speech recognition. Ph.D. Thesis, Loughborough University, Loughborough, England.
- Alghamdi, M., 2001. Arabic phonetics. Al-Toubah Bookshop, Riyadh, Saudi Arabia.
- Alghamdi, M., M. Elshafei and H. Al-Muhtaseb, 2007. Arabic broadcast news transcription system. Int. J. Speech Technol., 10: 183-195.
- Alghamdi, M., M. Elshafei and H. Almuhtasib, 2002. Speech units for arabic text-to-speech. Proceeding of the 4th Workshop on Computer and Information Sciences, (WCIS'02), Dammam, pp: 199-212.

- Alghamdi, M., 2003. KACST Arabic phonetics database. Proceedings of the 15th International Congress on Phonetics Science (ICPhS'15), August 3-9, 2003, Barcelona, Spain, pp: 3109-3112.
- Awang, S., R. Yusof, M.F. Zamzuri and R. Arfa, 2013. Feature level fusion of face and signature using a modified feature selection technique. Proceedings of the 2013 International Conference on Signal-Image Technology and Internet-Based Systems, December 2-5, 2013, IEEE, Kyoto, Japan, pp: 706-713.
- CIA., 2008. World factbook. Central Intelligence Agency, Washington, DC., USA.
- Elshafei, M., 1991. Toward an Arabic text-to-speech system. *Arabian J. Sci. Eng.*, 16: 565-583.
- Elshafei, M., H. Al-Muhtaseb and M. Alghamdi, 2006. Statistical methods for automatic diacritization of Arabic text. Proceedings of the 18th National Computer Conference, March 26-29, 2006, Riyadh, Saudi Arabia, pp: 1-8.
- Elshafei, M., H. Almuhtasib and M. Alghamdi, 2002. Techniques for high quality text-to-speech. *Inf. Sci.*, 140: 255-267.
- Gaikwad, S.K., B. Gawali and P. Yannawar, 2011. Performance analysis of MFCC and DTW for isolated Arabic digit. *Int. J. Adv. Res. Comput. Sci.*, 2: 513-518.
- Gaikwad, S.K., B.W. Gawali and P. Yannawar, 2010. A review on speech recognition technique. *Int. J. Comput. Appl.*, 10: 16-24.
- Garvin, P.L. and P. Ladefoged, 1963. Speaker identification and message identification in speech recognition. *Phonetica*, 9: 193-199.
- Gawali, B.W., S. Gaikwad, P. Yannawar and S.C. Mehrotra, 2010. Marathi isolated word recognition system using mfcc and dtw features. *ACEEE. Int. J. Inf. Technol.*, 1: 21-24.
- Huang, C., E. Chang and T. Chen, 2001. Accent issues in Large Vocabulary Continuous Speech Recognition (LVCSR). MSR-TR-2001-69, Microsoft Corporation, Beijing, China. <https://www.microsoft.com/en-us/research/publication/accent-issues-in-large-vocabulary-continuous-speech-recognition-lvcsr/>
- Hyassat, H. and R.A. Zitar, 2006. Arabic speech recognition using SPHINX engine. *Intl. J. Speech Technol.*, 9: 133-150.
- Jelinek, F., L. Bahl and R. Mercer, 1975. Design of a linguistic statistical decoder for the recognition of continuous speech. *IEEE. Trans. Inf. Theory*, 21: 250-256.
- Karam, Z.N. and W.M. Campbell, 2007. A new kernel for SVMMLLR based speaker recognition. Proceedings of the 8th Annual International Conference on Speech Communication Association, August 27-31, 2007, Antwerp, Belgium, pp: 290-293.
- Myers, C. and L. Rabiner, 1981. A level building dynamic time warping algorithm for connected word recognition. *IEEE. Trans. Acoust. Speech Signal Process.*, 29: 284-297.
- Nolan, F., 1983. The Phonetic Bases of Speaker Recognition. Cambridge University Press, Cambridge, UK., ISBN: 9780521244862, Pages: 221.
- Padmanabhan, M. and M. Picheny, 2002. Large-vocabulary speech recognition algorithms. *Computer*, 35: 42-50.
- Raj, A.A. and R. Susmitha, 2005. A voice interface for the visually impaired. Proceedings of the 3rd International Conference on Sciences of Electronic, Technologies of Information and Telecommunications (SETIT'05), March 27-31, 2005, Tunisia, North Africa, pp: 1-5.
- Satori, H., H. Hiyassat, M. Harti and N. Chenfour, 2009. Investigation arabic speech recognition using CMU sphinx system. *Int. Arab J. Inform. Technol.*, 6: 186-190.
- Shrishrimal, P.P., R.R. Deshmukh and V.B. Waghmare, 2012. Indian language speech database: A review. *Int. J. Comput. Appl.*, 47: 17-21.
- Talab, M.A., S.N.H.S. Abdullah and M.H.A. Razalan, 2013. Edge direction matrixes-based local binary patterns descriptor for invariant pattern recognition. Proceedings of the International Conference of Soft Computing and Pattern Recognition, December 15-18, 2013, Hanoi, Vietnam, pp: 13-18.
- Tomokiyo, L.M., A.W. Black and K.A. Lenzo, 2003. Arabic in my hand: Small-footprint synthesis of Egyptian Arabic. Proceedings of the 8th European Conference on Speech Communication and Technology, September 1-4, 2003, Geneva, Switzerland, pp: 2049-2052.
- UNESCO., 2017. World Arabic language day. The United Nations Educational, Scientific and Cultural Organization, Paris, France.