Scopus

Documents

Alqadasi, A.M.A.[a b c] , Zeki, A.M.[b] , Sunar, M.S.[c d] , Salam, M.S.B.H.[c] , Abdulghafor, R.[b e] , Khaled, N.A.[f]

Improving Automatic Forced Alignment for Phoneme Segmentation in Quranic Recitation

(2024) *IEEE Access*, 12, pp. 229-244.

DOI: 10.1109/ACCESS.2023.3345843

[a] Taiz University, Al-Saeed Faculty of Engineering and Information Technology, Taiz, Yemen
[b] International Islamic University Malaysia, Faculty of Information and Communication Technology, Kuala Lumpur, 53100, Malaysia
[c] Universiti Teknologi Malaysia, Faculty of Computing, Johor Bahru, 81310, Malaysia
[d] Institute of Human Centered Engineering, Universiti Teknologi Malaysia, Media and Game Innovation Centre of Excellence, Johor Bahru, 81310, Malaysia
[e] Arab Open University Oman, Faculty of Computer Studies (FCS), Muscat, 130, Oman
[f] International Islamic University Malaysia, AbdulHamid A. AbuSulayman Kulliyyah of Islamic Revealed Knowledge and Human Sciences, Kuala Lumpur, 53100, Malaysia

Abstract
Segmentation plays a crucial role in speech processing applications, where high accuracy is essential. The quest for improved accuracy in automatic segmentation, particularly in the context of the Arabic language, has garnered substantial attention. However, the differences between Qur'an recitation and normal Arabic speech, especially with regard to intonation rules affecting the lengthening of long vowels, pose challenges in segmentation especially for Qur'an recitation. This research endeavors to address these challenges by delving into the domain of automatic segmentation for Qur'an recitation recognition. The proposed scheme employs a hidden Markov models (HMMs) forced alignment algorithm. To enhance the precision of segmentation, several refinements have been introduced, with a primary emphasis on the phonetic model of the Qur'an and Tajweed, particularly the intricate rules governing elongation. These enhancements encompass the adaptation of an acoustic model tailored for Qur'anic recitation as preprocessing and culminate in the development of an algorithm aimed at refining forced alignment based on the phonetic nuances of the Qur'an. These enhancements are seamlessly integrated as post-processing components for the classic HMM-based forced alignment. The research utilizes a comprehensive database featuring recordings from 100 renowned Qur'an reciters, encompassing the recitation of 21 Qur'anic verses (Ayat). Additionally, 30 reciters were asked to record the same verses, incorporating various recitation speed patterns. To facilitate the evaluation process, a Random sample of the Qur'anic database was manually segmented, comprised 21 Ayats, totaling 19,800 words, with 89 unique words (14 verses x 3 recitation levels: fast, slow and normal x 6 readers). The outcomes of this study manifest notable advancements in the alignment of long vowels within Qur'an recitation, all while maintaining the precise alignment of vowels and consonants. Objective comparisons between the proposed automatic methods and manual segmentation were conducted to ascertain the superior approach. The findings affirm that the classic forced alignment method produces satisfactory outcomes when employed on verses lacking long vowels. However, its performance diminishes when confronted with verses containing long vowels. Therefore, the test samples were categorized into three groups based on the presence of long vowels, resulting in a Correct Classification Rate (CCR) that ranged from 6% to 57%, contingent on whether the verse includes long vowels or not. The average CCR across all test samples was 23%. In contrast, the proposed algorithm significantly enhances audio segmentation. It achieved CCR values ranging from 16% to 70% within the same database categories, with an average CCR of 45% across all test samples. This marks a notable advancement of 22% in segmented speech accuracy, particularly within a 30 ms tolerance, for verses containing long vowels. © 2023 The Authors.

Author Keywords
Arabic phoneme segmentation; forced alignment; Phoneme alignment; phoneme duration; phoneme recognition; phoneme segmentation; recitation recognition; Tajweed recognition

Index Keywords
Hidden Markov models, Job analysis, Linguistics, Speech processing, Speech recognition; Arabic phoneme segmentation, Forced alignment, Hidden-Markov models, Manual, Phoneme alignments, Phoneme duration, Phoneme segmentation, Phonemes recognition, Recitation recognition, Tajweed recognition, Task analysis; Alignment

References

- Jiang, N., Liu, T.
  An improved speech segmentation and clustering algorithm based on SOM and K-means
  (2020) *Math. Problems Eng.*, 2020, pp. 1-19.
  Sep.

- Salam, M.-S., Mohamad, D.
  Insertion reduction in speech segmentation using neural network
  (2008) *Proc. Int. Symp. Inf. Technol.*, pp. 1-7.
  Aug.

- Kalamani, M., Valarmathy, S., Anitha, S., Mohan, R.
  Review of speech segmentation algorithms for speech recognition
  (2014) *Int. J. Adv. Res. Electron. Commun. Eng.*, 3 (11), pp. 1-3.

- Geetha, K., Chandra, E.
  Automatic phoneme segmentation of Tamil utterances
  (2015) *Proc. Int. Conf. Adv. Comput. Commun. Syst.*, pp. 1-4.
  Jan.

- He, S., Zhao, H.
  Automatic syllable segmentation algorithm of Chinese speech based on MF-DFA
  (2017) *Speech Commun.*, 92, pp. 42-51.
  Sep.

- Kamper, H., Jansen, A., Goldwater, S.
  Unsupervised word segmentation and lexicon discovery using acoustic word embeddings
  (2016) *IEEE/ACM Trans. Audio, Speech, Language Process.*, 24 (4), pp. 669-679.
  Apr.

- Gan, L., Zhang, Y.
  Investigating self-Attention network for Chinese word segmentation
  (2020) *IEEE/ACM Trans. Audio, Speech, Language Process.*, 28, pp. 2933-2941.

- Zhang, W., Pang, M., Du, R., Liu, Y.
  Automatic speech sentence segmentation from multi-paragraph databases
  (2010) *Proc. Int. Conf. Measuring Technol. Mechatronics Autom.*, 1, pp. 721-724.
  Mar.

- Rybach, D., Gollan, C., Schluter, R., Ney, H.
  Audio segmentation for speech recognition using segment features
  (2009) *Proc. Ieee Int. Conf. Acoust., Speech Signal Process.*, pp. 4197-4200.
  Apr.

- Awais, M., Masud, S., Shamail, S.
  Continuous Arabic speech segmentation using FFT spectrogram
  (2006) *Proc. Innov. Inf. Technol.*, pp. 1-6.
  Nov.

- Adami, A.G., Hermansky, H.
  Segmentation of speech for speaker and language recognition
  (2003) *Proc. 8th Eur. Conf. Speech Commun. Technol. (Eurospeech).*, pp. 1-4.
  Sep.

- Anh, T.T., Huu, M.N., Trong, K.N.
  A method for automatic Vietnamese speech segmentation
  (2019) *Int. J. Innov. Technol. Exploring Eng. (IJITEE).*, 8 (11).

- Saiegh-Haddad, E., Henkin-Roitfarb, R.
  The structure of Arabic language and orthography
  (2014) *Handbook Arabic Literacy.*, pp. 3-28.
  Cham, Switzerland: Springer

- Alashqar, A.M.
  A comparative study on Arabic POS tagging using Quran corpus

(2012) *Proc. 8th Int. Conf. Informat. Syst. (INFOS).*, pp. NLP29-NLP33.
May

- Rushdi, A., Swayd, A.R., Karul, K.
(2000) *Tajweed Rules in the Holy Quran*,
Syria-Damascus, Syria: Dar Al-Khair Islamic Books Publisher

- Czerepinski, K.C., Swayd, A.R.
(2006) *Tajweed Rules of the qur'An*,
Syria-Damascus, Syria: Dar Al-Khair Islamic Books Publisher

- Liew, S.-L.
A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms
(2022) *Sci. Data*, 9 (1), pp. 1-12.
Jun.

- Scharenborg, O., Wan, V., Ernestus, M.
Unsupervised speech segmentation: An analysis of the hypothesized phone boundaries
(2010) *J. Acoust. Soc. Amer.*, 127 (2), pp. 1084-1095.
Feb.

- Rahman, M.M., Khatun, F., Bhuiyan, M.A.-A.
Blocking black area method for speech segmentation
(2015) *Int. J. Adv. Res. Artif. Intell.*, 4 (2), pp. 1-6.

- Kaur, E.A., Singh, E.T.
Segmentation of continuous Punjabi speech signal into syllables
(2010) *Proc. World Congr. Eng. Comput. Sci.*, 1, pp. 20-22.

- Hossain, A., Nahid, N., Khan, N.N., Gomes, D.C., Mugab, S.M.
Automatic silence/unvoiced/voiced classification of Bangla velar phonemes: New approach
(2005) *Proc. 8th Iccit*,

- Sharma, N., Singh, P.
Automatic segmentation of wave file
(2010) *Int. J. Comput. Sci. Commun.*, 1 (2), pp. 267-270.

- Tolba, M., Nazmy, T., Abdelhamid, A., Gadallah, M.
A novel method for Arabic consonant/vowel segmentation using wavelet transform
(2005) *Int. J. Intell. Cooperat. Inf. Syst.*, 5 (1), pp. 353-364.

- Mporas, I., Ganchev, T., Fakotakis, N.
A hybrid architecture for automatic segmentation of speech waveforms
(2008) *Proc. Ieee Int. Conf. Acoust., Speech Signal Process.*, pp. 4457-4460.
Mar.

- Ziólko, B., Manandhar, S., Wilson, R.C., Ziólko, M.
Wavelet method of speech segmentation
(2006) *Proc. 14th Eur. Signal Process. Conf.*, pp. 1-5.
Sep.

- Siniscalchi, S.M., Schwarz, P., Lee, C.-H.
High-Accuracy phone recognition by combining high-performance lattice generation and knowledge based rescoring
(2007) *Proc. Ieee Int. Conf. Acoust., Speech Signal Process. (ICASSP).*,
Apr.

- Al-Qadasi, A.M.A., Sunar, M.S., Salam, M.S.H., Abdulghafor, R.
  Medd recognition for Al-Quran recitation using ANN-based model
  (2023) *Proc. Post Pandemic Challenges Embracing Soc.*, p. 1.

- Kamarauskas, J.
  Automatic segmetation of phonemes using artificial neural networks
  (2006) *Elektronika Ir Elektrotechnika*, 72 (8), pp. 39-42.

- Bansal, P., Pradhan, A., Goyal, A., Sharma, A., Arora, M.
  Speech synthesis-Automatic segmentation
  (2014) *Int. J. Comput. Appl.*, 98 (4), pp. 29-31.

- Dines, J., Sridharan, S., Moody, M.
  Automatic speech segmentation with HMM
  (2002) *Proc. 9th Austral. Conf. Speech Sci. Technol.*, pp. 544-549.

- Stolcke, A., Ryant, N., Mitra, V., Yuan, J., Wang, W., Liberman, M.
  Highly accurate phonetic segmentation using boundary correction models and system fusion
  (2014) *Proc. Ieee Int. Conf. Acoust., Speech Signal Process. (ICASSP).*, pp. 5552-5556.
  May

- Baby, A., Prakash, J.J., Murthy, H.A.
  A hybrid approach to neural networks based speech segmentation
  (2017) *Proc. Int. Symp. Frontiers Res. Speech Music*, pp. 15-16.
  Rourkela, India: National Institute of Technology (NIT).

- Yuan, J., Lai, W., Cieri, C., Liberman, M.
  Using forced alignment for phonetics research
  (2018) *Chinese Language Resources and Processing: Text, Speech and Language Technology.*,
  Cham, Switzerland: Springer

- Mahr, T.J., Berisha, V., Kawabata, K., Liss, J., Hustad, K.C.
  Performance of forced-Alignment algorithms on Children's speech
  (2021) *J. Speech, Lang., Hearing Res.*, 64 (6 S), pp. 2213-2222.
  Jun.

- Kim, H., Choi, H.-S.
  Towards trustworthy phoneme boundary detection with autoregressive model and improved evaluation metric
  (2023) *Proc. Ieee Int. Conf. Acoust., Speech Signal Process. (ICASSP).*, pp. 1-5.
  Jun.

- Wohlan, B., Pham, D.-S., Chan, K.Y., Ward, R.
  A text-independent forced alignment method for automatic phoneme segmentation
  (2022) *Ai 2022: Advances in Artificial Intelligence*, pp. 585-598.
  Cham, H. Aziz, D. Corrêa, and T. French, Eds. Cham, Switzerland: Springer

- Tamiru, R.M., Abate, S.T.
  Sentence-level automatic speech segmentation for Amharic
  (2022) *Proc. 6th Int. Congr. Inf. Commun. Technol. (ICICT).*, 2, pp. 477-485.
  London, U.K. Cham, Switzerland: Springer

- Leinonen, J., Partanen, N., Virpioja, S., Kurimo, M.
  Semiautomatic speech alignment for under-resourced languages
  (2022) *Proc. Workshop Resour. Technol. Indigenous, Endangered Lesser-Resourced Lang. Eurasia Within 13th Lang. Resour. Eval. Conf.*, pp. 17-21.
  [Online]

- Peng, W., Gao, Y., Lin, B., Zhang, J.
  A practical way to improve automatic phonetic segmentation performance
  (2021) *Proc. 12th Int. Symp. Chin. Spoken Lang. Process. (ISCSLP).*, pp. 1-5.
  Jan.

- Punnoose, A.K.
  A study on forced alignment error patterns in Kaldi
  (2022) *Proc. 8th Int. Conf. Signal Process. Commun. (ICSC).*, pp. 250-253.
  Dec.

- Backstrom, D.A., Tucker, B.V., Kelley, M.C.
  Forced-Alignment of the sung acoustic signal using deep neural nets
  (2019) *Can. Acoust.*, 47 (3), pp. 98-99.

- Milne, P.M.
  (2015) *Improving the Accuracy of Forced Alignment Through Model Selection and Dictionary Restriction*,
  M.S. Thesis, Dept. Linguistics, McGill Univ.

- Rahmatullah, G.M., Ruan, S.-J.
  Performance evaluation of Indonesian language forced alignment using Montreal forced aligner
  (2023) *Proc. 6th Int. Symp. Comput., Consum. Control (IS3C).*, pp. 163-166.
  Jun.

- Al-Manie, M.A.
  Arabic speech segmentation: Automatic verses manual method and zero crossing measurements
  (2010) *Indian J. Sci. Technol.*, 3 (12), pp. 1134-1138.
  Dec.

- Anwar, M.J., Awais, M., Masud, S., Shamail, S.
  Automatic Arabic speech segmentation system
  (2006) *Int. J. Inf. Technol.*, 12 (6), pp. 102-111.

- Nofal, M., Abdel-Raheem, E., El Henawy, H., Kader, N.S.A.
  Arabic automatic segmentation system and its application for Arabic speech recognition system
  (2003) *Proc. 46th Midwest Symp. Circuits Syst.*, pp. 697-700.
  Dec.

- Absa, A.H.A., Deriche, M., Elshafei-Ahmed, M., Elhadj, Y.M., Juang, B.-H.
  A hybrid unsupervised segmentation algorithm for Arabic speech using feature fusion and a genetic algorithm (July 2018)
  (2018) *Ieee Access*, 6, pp. 43157-43169.

- Karim, R.M., Suyanto
  Optimizing parameters of automatic speech segmentation into syllable units
  (2019) *Int. J. Intell. Syst. Appl.*, 11 (5), pp. 9-17.
  May

- Frihia, H., Bahi, H.
  HMM/SVM segmentation and labelling of Arabic speech for speech recognition applications
  (2017) *Int. J. Speech Technol.*, 20 (3), pp. 563-573.
  Sep.

- Ahcène, A., Aissa, A., Abdelkader, D., Khadidja, B., Ghania, D.
  Automatic segmentation of Arabic speech signals byHMMand ANN
  (2016) *Proc. Int. Conf. Electr. Sci. Technol. Maghreb (CISTEM).*, pp. 1-4.
  Oct.

- Javed, M., Baig, M.M.A., Qazi, S.A.
Unsupervised phonetic segmentation of classical Arabic speech using forward and inverse characteristics of the vocal tract
(2020) *Arabian J. Sci. Eng.*, 45 (3), pp. 1581-1597.
Mar.

- Alqadasi, A.M.A., Sunar, M.S., Turaev, S., Abdulghafor, R., Hj Salam, M.S., Alashbi, A.A.S., Salem, A.A., Ali, M.A.H.
Rule-based embedded HMMs phoneme classification to improve Qur'anic recitation recognition
(2022) *Electronics*, 12 (1), p. 176.
Dec.

- Yuan, J., Ryant, N., Liberman, M., Stolcke, A., Mitra, V., Wang, W.
Automatic phonetic segmentation using boundary models
(2013) *Proc. Interspeech*, pp. 2306-2310.
Aug., [Online]

- Mohammed, A., Sunar, M.S.
Toward A rich quranic Arabic speech corpus for Tajweed rules
(2019) *Proc. Universal Wellbeing (ICUW).*, p. 210.

- Alqadasi, A.M.A., Abdulghafor, R., Sunar, M.S., Salam, M.S.B.H.J.
Modern standard Arabic speech corpora: A systematic review
(2023) *Ieee Access*, 11, pp. 55771-55796.

- Al-Qadasi, A.M.A.
(2021) *Phoneme Duration Scheme for Tajweed Medd Rules Recognition in qur'An Recitation*,
Ph.D. dissertation, Dept. Comput. Sci., Universiti Teknologi Malaysia, Malasia, [Online]

- Wightman, C.W., Talkin, D.T.
The aligner: Text-To-speech alignment using Markov models
(1997) *Prog. Speech Synth.*, pp. 313-323.
Cham, Switzerland: Springer

- Flores Solórzano, S., Coto-Solano, R.
Comparison of two forced alignments systems for aligning Bribri speech
(2017) *Clei Electron. J.*, 20 (1), pp. 21-213.

- Young, S.
(2002) *The Htk Book*, 3, p. 175.
Cambridge, U.K.: Cambridge Univ. Press, Engineering Department

Correspondence Address
Alqadasi A.M.A.; Taiz University, Yemen; email: ammaralqadasi2@gmail.com

Publisher: Institute of Electrical and Electronics Engineers Inc.

ISSN: 21693536
Language of Original Document: English
Abbreviated Source Title: IEEE Access
2-s2.0-85181575250
Document Type: Article
Publication Stage: Final
Source: Scopus