Document details

Back to results | 1 of 1

Full Text | View at Publisher | Export | Download | Add to List | More...

Egyptian Informatics Journal

Volume 17, Issue 3, 1 November 2016, Pages 305-314

Building CMU Sphinx language model for the Holy Quran using simplified Arabic phonemes (Article)

El Amrani, M.Y.[ab], Rahman, M.M.H.[b], Wahiddin, M.R.[b], Shah, A.[b]

[a] Computer Science and Engineering Department, Jubail University College, Saudi Arabia

[b] Computer Science Department, Kulliah of Information Communication Technology, International Islamic University Malaysia, Malaysia

View references (23)

Abstract

This paper investigates the use of a **simplified** set of **Arabic phonemes** in an **Arabic** Speech Recognition system applied to **Holy Quran**. The **CMU Sphinx** 4 was used to train and evaluate a **language model** for the Hafs narration of the **Holy Quran**. The **building** of the **language model** was done using a **simplified** list of **Arabic phonemes** instead of the mainly used Romanized set in order to simplify the process of generating the **language model**. The experiments resulted in very low Word Error Rate (WER) reaching 1.5% while using a very small set of audio files during the training phase when using all the audio data for both the training and the testing phases. However, when using 90% and 80% of the training data, the WER obtained was respectively 50.0% and 55.7%. © 2016

Author keywords

Automatic speech recognition; **Holy Quran** recognition; Human voice

ISSN: 11108665 Source Type: Journal Original language: English

DOI: 10.1016/j.eij.2016.04.002 Document Type: Article

Publisher: Elsevier B.V.

View in search results format

References (23)

All Export | Print | E-mail | Create bibliography