

Myanmar Consonants Recognition System Based on Lip Movements Using Active Contour Model

Authors : T. Thein, S. Kalyar Myo

Abstract : Human uses visual information for understanding the speech contents in noisy conditions or in situations where the audio signal is not available. The primary advantage of visual information is that it is not affected by the acoustic noise and cross talk among speakers. Using visual information from the lip movements can improve the accuracy and robustness of automatic speech recognition. However, a major challenge with most automatic lip reading system is to find a robust and efficient method for extracting the linguistically relevant speech information from a lip image sequence. This is a difficult task due to variation caused by different speakers, illumination, camera setting and the inherent low luminance and chrominance contrast between lip and non-lip region. Several researchers have been developing methods to overcome these problems; the one is lip reading. Moreover, it is well known that visual information about speech through lip reading is very useful for human speech recognition system. Lip reading is the technique of a comprehensive understanding of underlying speech by processing on the movement of lips. Therefore, lip reading system is one of the different supportive technologies for hearing impaired or elderly people, and it is an active research area. The need for lip reading system is ever increasing for every language. This research aims to develop a visual teaching method system for the hearing impaired persons in Myanmar, how to pronounce words precisely by identifying the features of lip movement. The proposed research will work a lip reading system for Myanmar Consonants, one syllable consonants (ဂ (Nga) ဂ (Nya) ဂ (Ma) ဂ (La) ဂ (Wa) ဂ (Tha) ဂ (Ha) ဂ (Ah)) and two syllable consonants (ဂ (Ka Gyi) ဂ (Kha Gway) ဂ (Ga Nge) ဂ (Ga Gyi) ဂ (Sa Lone) ဂ (Sa Lain) ဂ (Za Gwe) ဂ (Da Dway) ဂ (Na Gyi) ဂ (Na Nge) ဂ (Pa Saug) ဂ (Ba Gone) ဂ (Ya Gaug) ဂ (La Gyi)). In the proposed system, there are three subsystems, the first one is the lip localization system, which localizes the lips in the digital inputs. The next one is the feature extraction system, which extracts features of lip movement suitable for visual speech recognition. And the final one is the classification system. In the proposed research, Two Dimensional Discrete Cosine Transform (2D-DCT) and Linear Discriminant Analysis (LDA) with Active Contour Model (ACM) will be used for lip movement features extraction. Support Vector Machine (SVM) classifier is used for finding class parameter and class number in training set and testing set. Then, experiments will be carried out for the recognition accuracy of Myanmar consonants using the only visual information on lip movements which are useful for visual speech of Myanmar languages. The result will show the effectiveness of the lip movement recognition for Myanmar Consonants. This system will help the hearing impaired persons to use as the language learning application. This system can also be useful for normal hearing persons in noisy environments or conditions where they can find out what was said by other people without hearing voice.

Keywords : feature extraction, lip reading, lip localization, Active Contour Model (ACM), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Two Dimensional Discrete Cosine Transform (2D-DCT)

Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development

Conference Location : Chicago, United States

Conference Dates : December 12-13, 2020