Abstract

Human uses visual information while trying to understand speech, especially in noisy conditions or in situations where the audio signal is not available. Lip reading is the technique of a comprehensive understanding the underlying speech by processing on the movement of lips. However, the recognition of lip motion is a difficult task since the region of interest (ROI) is nonlinear and noisy. In proposed method lip reading system we have used two stage feature extraction model which is precised, discriminative and computation efficient. The first stage 3D Discrete Wavelet Transform (3D-DWT) or 3D Discrete Cosine Transform (3D-DCT) is used and the second stage is Locality Sensitive Discriminant Analysis (LSDA) to trim down the feature dimensions. These features make a novel lip reading system with small feature vector size. In addition to the novel feature extraction technique, the performance of Naive Bayes and SVM classifier is compared. CUAVE database of 0 to 9 utterances in English is used for experimentation. Results of 3 dimension transform with LSDA are compared with 2 dimension transform with LSDA. Experimental results show that 3D-DWT+LSDA feature mining are compared with 3D-DWT with PCA or LDA. 3D-DWT+LSDA result is also compared with
3D-DCT + LSDA.

References

20. Deng Cai, X. He, K. Zhou, J.Han, H. Bao, “Locality discriminant analysis”, International

21. V. Kechman, “Learning and soft computing, support vector machines, Neural Networks
and Fuzzy logic models”, MIT Press Cambridge, 1-58, 2001


24. T. M. Mitchell,” Generative and Discriminative Classifiers: Naive Bayes and Logistic

database for multimodal human computer- interface research”, Proceedings of IEEE

Index Terms

Computer Science Artiﬁcial Intelligence

Keywords

LSDA, LDA, 3D-DWT, 3D-DCT, SVM, Naive Bayes, Lip reading.