



LIP FEATURE EXTRACTION BASED ON AUDIO-VISUAL CORRELATION (ThuAmOR11)

★ Author(s) :

Mehmet Emre Sargin

(Koc University, Turkey)

Engin Erzin

(Koc University, Turkey)

Yucel Yemez

(Koc University, Turkey)

A. Murat Tekalp

(Koc University, Turkey)

★ Abstract :

In this paper, the lip feature that has the highest correlation with audio features is investigated. Audio features are selected as Mel Frequency Cepstral Coefficients (MFCC) of the audio signal. Three different lip features are considered for the visual lip information, where these features are 2D DCT coefficients of the intensity based image and the optical flow vectors within the lip region, and the distances between pre-defined points on the lip contour which carries the lip shape information. In this study, we present two techniques based on class conditional probability analysis and canonical correlation analysis to estimate and compare the correlations between audio feature and each lip feature. The lip feature, which has the highest correlation to audio features, is identified among the above lip features. Isolation of lip features, which are highly correlated with audio signal, can be used for audio-visual speech recognition, audio-visual lip synchronization and estimation of lip shapes using audio signal for visual synthesis.