

Interactie: lip lezen

Keywords

Multi-modale mens-machine interactie, lip lezen

Probleemstelling

Automatic lip reading (also known as speech reading) is a relatively new area of the speech-related research. It comprises topics ranging from image and video processing to the recognition architectures and natural language processing. It is most often used in combination with audio based speech recognition in order to improve its robustness (or lack thereof) against noise. The development of an automatic lip reading system can be seen as an important intermediate step in the direction of multi-modal speech recognition.

Technisch probleem

The automatic lip reading system must firstly process the incoming video signal so that it is suitable for recognition. The vast amount of video data must be substantially reduced. The features relevant to the articulation need to be extracted from the signal. All this should happen in the initial stages of the processing and should be done in a real-time. Further, the recognizer itself must be designed and trained.

Soort oplossing

Our project in the lip reading started in 1998 with the aim of developing methods for the whole range of lip reading related problems. The research covered also several experiments with different recognizers and different application scenarios. We developed the artificial neural network (ANN) based voice activity detector that uses as input the visual modality only. We have also a working prototype of Hidden Markov Model (HMM) based lip reading system capable of recognizing sequences of spoken digits. Both systems are available as demonstrators.

Vergelijking met bestaande oplossingen

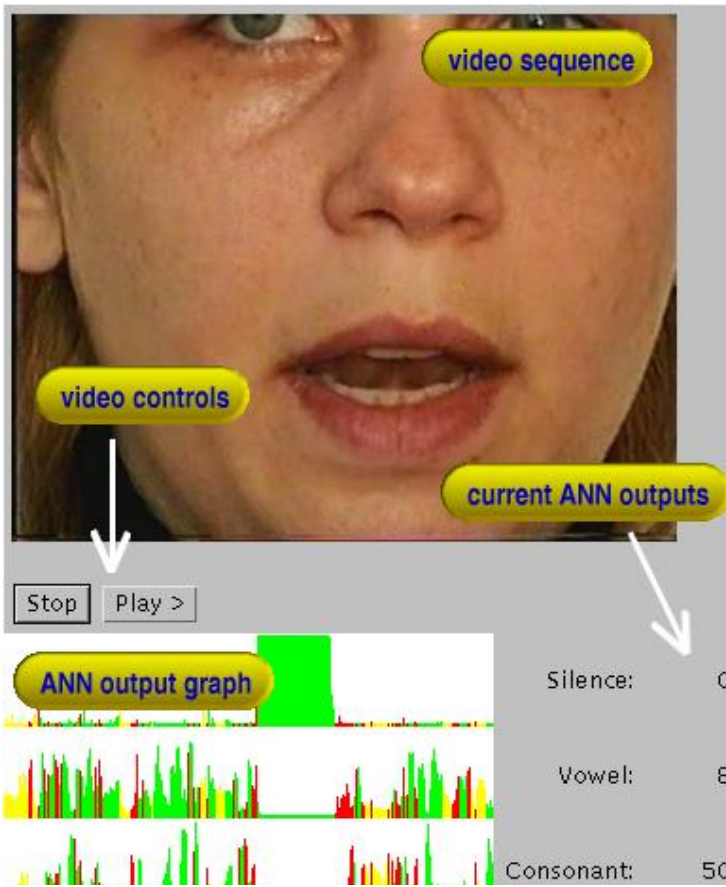
State of the art.

Potentiële afnemers

Multimedia information analysis, television, services to the deaf.

Contact

Technische Universiteit Delft
Afdeling Mediamatica
Tel. (015) 278 6052



Silence detection and vowel / consonant discrimination in video sequences.