ACOUSTICS2008/3565
Analysis and synthesis of non-verbal facial motion

Jonas Beskow and Björn Granström

KTH Speech, music and hearing, Lindstedtsvägen 24, 10044 Stockholm, Sweden

Until recently, most efforts in audio-visual speech synthesis have been concerned with verbal content. However, in human-human communication it is obvious that non-verbal signals plays an important role, such as when expressing emotions and attitudes. Interaction is also often regulated using facial cues, for example gaze, head and eyebrow movements. Some of these cues have a direct coupling to the speech signal, while other occur during both while speaking and listening. When applying interactive talking agents in man-machine systems, non-verbal signals may be very important in easing the flow of communication. In a series of experiments we have been exploring the function of non-verbal facial motion. These studies include an experiment on the interaction between expressive speech and prominence, as well as an attempt to synthesize emotions and attitudes in a talking head, using 3D motion capture data. Further we will report on a real-time experiment with human-human avatar-mediated conversation, where the subjects' turn-taking behavior is affected by facial motion in the avatars.