Adaptive Man-Machine Interfaces
MIT9904-15
Tomaso Poggio

Most significant progress so far

We have demonstrated the feasibility of a 3D text-to-visual-speech (TTVS) technique and we have significantly extended the TTVS technique of Miketalk by using a low-dimensional morphable model of the mouth.

Proposal for July 2001 to June 30, 2002

By June 2001 we will have technologies 1) to use 3D models of faces -- rather than face images -- and to output a 3D model of a speaking face and 2) to deal with coarticulation issues. We will also have feasibility demos.

In the year from July 2001 to June 2002 we propose to build prototype systems for a 3D TTVS and for a second-generation TTVS system, based on morphable models and HMMs. This will consolidate the work so far. We also plan to explore the following research issues:

1. Synthesize a realistic animation video of a speaking person from just one image of the person using 3D morphable models.

2. Incorporate higher-level communication mechanisms into our (2D and possibly 3D) talking facial model, such as various expressions (eyebrow   raises, head movements, and eye blinks).

3. Assess the realism of the talking face. We plan to perform several psychophysical tests to evaluate the realism of our system.
4. Extend our approach using morphable model from TTVS to TTS. We plan to first study morphing of audio sequences. The system will take as input 2 audio sequences, and produce as output intermediate audio sequences that approximate natural exemplars lying between the 2 input sequences. Audio morphing might have important applications in speech synthesis.

Collaboration

We had a visit by Drs Hagita, Sawaki, Murase in Aug 99
A visit by Dr. Matsuda in May 2000
A working stage at MIT by Dr. Minako Sawaki from February to April 2000
Another stage at CBCL by Dr. Minako Sawaki from August to September 2000