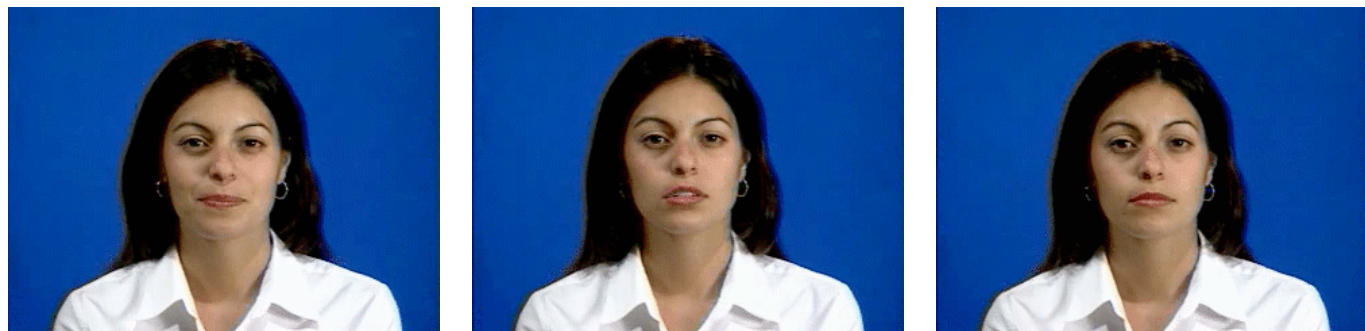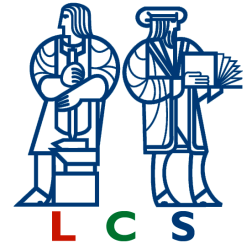# Project Overview

Three extensions of our text-to-visual-speech (TTVS) system:

- extend our morphing approach from video to audio
- extend the system to use morphing of 3D models of faces
- improve the performance of system

Ezzat, Geiger, Poggio, 2002

Progress Through December 2002

Audio Synthesis:

 Collected 15 minutes of spoken data

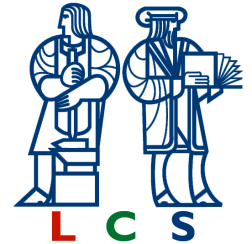 Implemented TD-PSOLA to warp duration & pitch

Original audio

Original audio,
Duration-normalized

Original audio,
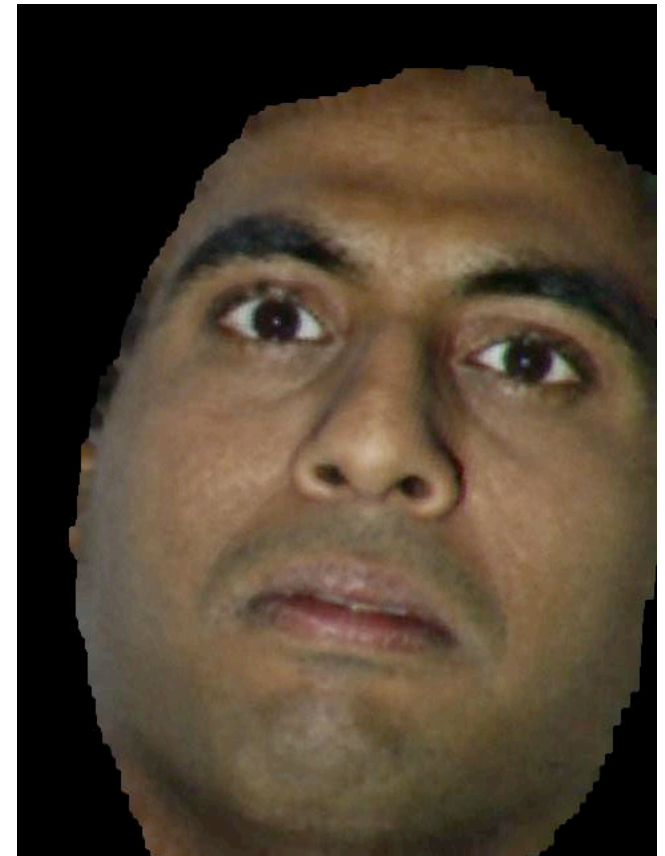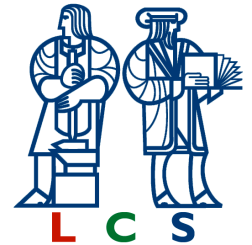Duration-normalized,
Pitch-normalized

Improving real-time performance:
- Acquired a fast 2 Ghz Linux machine
- Acquired Realnetworks streaming software to stream synthesized video

Extending our system from 2D to 3D:
- Acquired an Eyetronics structured light scanner for realtime 3D acquisition (example texture sample on the right)

Research Plan for the Next Six Months

We plan in the next six months to:

1) Look into spectral methods for audio warping & morphing,

2) Begin extending the system to use morphing of **3D models** of faces -- rather than face images using Eyetronics scanner

3) Improve the performance and intelligibility of the system