Carnival: a modular framework for automated facial animation

Michael Berger, Gregor Hofer, Hiroshi Shimodaira
Centre for Speech Technology Research
School of Informatics, University of Edinburgh
m.a.berger@sms.ed.ac.uk

Problem
Facial animation is difficult to do convincingly, particularly when synchronizing with speech. There are various ways to automate facial animation:

- Performance-driven animation
- Audio-driven animation
- Audio-visual text-to-speech synthesis (AVTTS)

The problem with these solutions is they bring together software and data formats from different fields—in particular speech technology and graphics technology—that are not well integrated.

- Conversion of facial dynamics into animation is cumbersome, slow and offline.
- Lack of live connection between speech and rendering pipelines. Difficult to backtrace animation problems, or see outcome of edits in the speech processing level.
- No standard control interface for different facial models, so adaptation process must be repeated in each case.

Solution
Software framework called “Carnival” which places speech and graphics components within a single object-oriented system.

- Fast and automatic end-to-end processing
- Real-time animation and linked display of time-varying representations for instantaneous feedback/feed-forward information
- Standardized object interfaces for easy integration of new components

The core of our solution is a platform independent C++ API.

Applications
• API may be used for fast prototyping of automated animation systems
• Suitable for performance-driven, audio-driven, or AVTTS applications
• Our implemented tool built on the API is suitable for in-house industrial or academic use

The Component class hierarchy of the Carnival API. Components are objects that can be loaded, edited and saved. They include both data objects and the processing modules out of which systems are built.

GUI window for a Visualizer, with real-time output and manual slider controls for the deformation parameters.