

MULTILINGUAL AVATARS

Cities and towns around the world have a huge wealth of historically significant and culturally important material in the form of buildings and events. We are creating a toolkit of applications specifically to enable this information to be recreated as virtual reality worlds for visitors to see for themselves how buildings once looked; to hear from virtual guides the history of significant buildings and events; to explore locations unavailable to them.

Modelling and rendering. A typical city scene might consist of a small (500 m²) region with important buildings and a much larger surrounding region of generic housing. By providing tools to rapidly create these generic buildings and automatically position them



Recreation of a street in Wolfenbützel



Wolfenbützel

along roads, a large proportion of the scene can be constructed in a relatively short time. A scene will also need important buildings (modelled using traditional modelling packages) and parametric trees to add realism and movement. Once the background has been created a number of people can be added around the scene. These



avatars can talk to the visitor and will be combined with a background crowd of city residents so that the model comes to life with an interactive information guide.

Speaking Avatars Speech is required to make interaction as natural and engaging as possible. The avatar must be able to respond to questions from a range of users with replies in a number of languages. Tools are required that allow new speaking avatars to be quickly developed. To this end, mouth shapes for all the speech sounds are created by mixing six base morphs and applying them to the reference face. Realistic visual speech can be generated using this limited set of morphs.

Technical Specifications The Wolfenbützel showcase application is being developed using the open-source scenegraph OpenSG with custom add-ons for rendering speedup techniques including subdivision surface work from Braunschweig University.

The virtual environment needs to be related to a knowledge base of information about the site being visited, and this knowledge has to be formulated into the appropriate natural (text) language response, with synchronized movement and gestures as the visitor



asks questions of the guide.

Speech can be sourced from any Text to Speech Synthesizer or recorded by an actor. Phoneme timings are obtained by using the HTK recognizer. These phoneme timings are then used to create a series of morphs which drive the facial animation. Both



modelling and rendering packages are designed for mid- to high-end consumer PCs; with support for multi-channel big screen theatres.

Partners This showcase is being realized by the following EPOCH partners. It will be demonstrated at the VAST 2004 conference.

- ▶ University of East Anglia, UK
- ▶ Technical University of Braunschweig, Germany
- ▶ University of Brighton, UK



Interested?

Are you interested in this showcase? Do you think that this approach can help you in creating effective Cultural Heritage presentation projects or can be integrated in new research projects? Please contact Dr. Andy Day (amd@cmp.uea.ac.uk) of University of East Anglia at +44 (0)1603 592604.

EPOCH is a Network of Excellence on Intelligent Cultural Heritage within the IST (Information Society Technologies) section of the Sixth Framework Programme of the European Commission. EPOCH showcases demonstrate innovative solutions and technological integration for target application areas in the Cultural Heritage domain. As they are created with real world content, they stimulate creative thinking about the use of the technologies in Cultural Heritage, and are used to validate new technological approaches with key stakeholders in the Cultural Heritage domain. For more details, visit the project web site:

www.epoch-net.org

EPOCH is funded by the European Commission under the Community's Sixth Framework Programme, contract no. 507382. However, this leaflet reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained herein.

