

Google Earth Multimodal

UIST 2007 Demo

Study Multimodal Interaction with an Existing Application

Motivation

This demo allows a user to navigate around the familiar Google Earth application using a combination of simple actions with the SHAKE Bluetooth sensor pack and speech recognition. The demo provides an example of a relatively complex pipeline constructed using the OIDE, including feedback, consisting of a total of 9 distinct components, and is built around an existing closed-source application. In addition, it takes advantage of some of the generic debugging and monitoring components available to OpenInterface applications.



Modalities

This demo uses speech recognition and several features of the SHAKE Bluetooth sensor pack, including capacitive sensing and vibrotactile feedback. OpenInterface also offers a selection of generic debugging and monitoring components, some of which are used in the demo. For example, an “oscilloscope” component is used to display realtime sensor data from the SHAKE together with terrain height data from the Google Earth application.



Using this demo

Hold the SHAKE device flat in your hand. Tilt the device left and right to scroll the map west and east. Tilt the device forward (away from you) to scroll north, and tilt it back (towards you) to scroll south. To stop moving, return the device to the original horizontal position.

To zoom in, first stop moving the device and hold it horizontally, then press your finger over the sunken square on the left of the top surface of the SHAKE. This should cause the Google Earth viewpoint to zoom in. Zooming out is done in the same fashion, using the sunken square on the right of the top surface of the SHAKE.

Finally, a list of city names is displayed on the screen. To navigate immediately to one of these cities, place the SHAKE in the “neutral” position, hold down the button on the right side of the device, speak the name of the city, then release the button. The Google Earth viewpoint should rapidly move to display the location of the selected city.

Component details

The three purple components form the main part of the application. From bottom to top, they consist of: an interface to the SHAKE device, a speech recognition component, and a component which provides an interface to the Google Earth API. The button on the SHAKE device is used as a “push-to-talk” trigger for the speech recognition component. Speaking the name of a city into the microphone then causes the Google Earth API component to move the current viewpoint to the location of that city.

The two blue components provide both vibrotactile and audio feedback to the user. The “landscape audio” component simulates a ball bouncing on a landscape, the height data being provided by the Google Earth API. Each time the ball impacts the landscape, audio feedback is produced, and an event is passed to the second component, which triggers a vibration through the SHAKE device.

