Comparing Single-modal and Multimodal Interaction in an Augmented Reality System

Zhimin Wang[1,2] ,Huangyue Yu[1] ,Haofei Wang[2] ,Zongji Wang[1] ,Feng Lu[1,2 *]

1 State Key Laboratory of VR Technology and Systems, School of CSE, Beihang University

2 Peng Cheng Laboratory, Shenzhen, China

➢ Motivation&Key Idea

- Challenge: 1) There are few works fusing three or more interaction modalities. 2) Current AR HMD need complex AR registration algorithms and cannot be easily adapted to the different needs.
- We propose a gaze-gesture-speech AR system (GGS-AR) and study the effects of different interaction techniques.

➢ Methods

Experimental setup

- The head-mounted parts of the system sense user's input from different communication channels.
- The modality fusion and display parts of the system are left to the remote end.

We investigated five representative modalities:

- two single-modal techniques (Gesture Only and Gaze Only)
- three multimodal techniques (Gesture+Speech, Gaze+Speech and Gaze+Gesture+Speech).

➢ Results

- We evaluated the system in the table lamps scenario and compared the performance of different interaction techniques.
- The experimental results show that the Gaze+Gesture+Speech is superior in terms of performance.