

Issues in Multimodal Telecommunications

Tsuyoshi MORIMOTO

Abstract

We have been investigating the applicability of a multimodal user interface to an interpreted telecommunications system. In that interface, not only speech but also other modalities such as keyboard, graphics or handwriting will be prepared. As a first trial of the study, a 'direction-instructing task' was chosen as an application field. The results of the experiments have demonstrated the effectiveness of multimodality, and suggest applicability to other kinds of tasks as well.

In this paper, first, we describe the results of the experiments. The major effectiveness of multimodality can be summarized as follows:

(1) The roles of information provider and information recipient become clearer in a multimodal environment. As a result, an information recipient is able to use fewer words to achieve a goal (to get necessary information). Another phenomenon observed is a decrease of far-deictic expressions (locative or demonstrative polynomials). Identifying their referent in the dialogue context is one of the most difficult problems in current natural language processing technologies. On the other hand, the number of near-range deictic expressions increases, but they are used to point to some shared visual objects shown on the CRT, so their identification would be less difficult.

(2) The experiment results of simulated interpreted conversations have shown the flexibility of human users. They tend to use the same vocabularies used by the system, and in cases in which the system cannot process some complex or lengthy utterance, they speak more clearly and slowly, and rephrase their utterances in syntactically simpler forms. In the latter cases, the multimodal environment can allow users to change input modality from speech to non-speech media so that the input is more easily accepted by the system.

Then, we will discuss the applicability of multimodality to wider tasks by analyzing the currently available interpreting telephone service using human interpreters.

References

- [1] Loken-Kim, K.H., Yato, F., Fais, L., Morimoto, T., Kurematsu, A. : "Linguistic and Paralinguistic Differences between Multimodal and Telephone-only Dialogues," Proc. of ICSLP'94,1994
- [2] Fais, L. : "Effects of Communicative Mode on Spontaneous English Speech," Technical Report of the Institute of Electronics, Information and Communication Engineers, NLC94-22,1994
- [3] Fais, L., Loken-Kim, K. H., Park.Y.D. : "Speakers' Responses to Requests for Repetition in a Multimedia Cooperative Dialogue," Proc. of Int. Conf. on Cooperative Multimodal Communication, Eindhoven, The Netherlands, 1995