TAICAR - The Collection and Annotation of an In-Car Speech Database Created in Taiwan

Hsien-Chang Wang*, Chung-Hsien Yang+, Jhing-Fa Wang+,

Chung-Hsien Wu** and Jen-Tzung Chien**

Abstract

This paper describes a project that aims to create a Mandarin speech database for the automobile setting (TAICAR). A group of researchers from several universities and research institutes in Taiwan have participated in the project. The goal is to generate a corpus for the development and testing of various speech-processing techniques. There are six recording sites in this project. Various words, sentences, and spontaneously queries uttered in the vehicular navigation setting have been collected in this project. A preliminary corpus of utterances from 192 speakers was created from utterances generated in different vehicles. The database contains more than 163,000 files, occupying 16.8 gigabytes of disk space.

Keywords: TAICAR, in-car speech, speech database, multi-channel recording, corpus collection and annotation

1. Introduction

1.1 In-car speech corpora review

Driver information systems are becoming increasingly complex as more and more functions are integrated into modern cars. Speech-enabled functions will enhance the safety and convenience of operating for future vehicles. To realize such functions, in-car speech processing techniques need to be built and tested first. Thus, it is necessary to collect an in-car speech database. Although many speech corpora [Tapisa *et al.* 1994], [Roach *et al.* 1996], [Kudo *et al.* 1994], [Bernstein *et al.* 1994] have been created to improve speech-processing effectiveness, few in-car speech databases have been reported.

---

* Department of Information Management, Chang Jung Christian University,
  396 Chang Jung Road, Sec.1, Kway Jen, Tainan, Taiwan, R.O.C.
  Tel: 886-6-2785123 ext. 2071; Fax: 886-6-2785657
  E-Mail: wangbb@mail.cju.edu.tw
+ Department of Electrical Engineering, National Cheng Kung University
** Department of Computer Science and Information Engineering, National Cheng Kung University