



中华人民共和国国家标准

GB/T 35312—2017

中文语音识别终端服务接口规范

Specification of programming interface for Chinese speech recognition
terminal service

2017-12-29 发布

2018-07-01 实施

中华人民共和国国家质量监督检验检疫总局 发布
中国国家标准化管理委员会

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 中文语音识别终端服务系统能力范围	3
4.1 概述	3
4.2 基本功能	3
4.3 扩展功能	3
5 输入数据描述	4
5.1 前端信号处理	4
5.2 语音唤醒	5
5.3 语音识别	6
6 输出数据描述	7
6.1 前端信号处理	7
6.2 语音唤醒	7
6.3 语音识别	7
7 服务接口描述	8
7.1 概述	8
7.2 前端信号处理服务	8
7.3 语音唤醒服务	9
7.4 语音识别服务	10
附录 A (规范性附录) 中文语音识别终端服务系统性能指标定义和测试基本要求	12
附录 B (资料性附录) 语音唤醒和语音识别结果消息示例	14
参考文献	16

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本标准起草单位:中国科学院声学研究所、中国电子技术标准化研究院、中国电信集团公司、中兴通讯股份有限公司、科大讯飞股份有限公司、上海计算机软件技术开发中心。

本标准主要起草人:付强、颜永红、董滨、国雁萌、王晓飞、黎塔、杨智慧、韩卫生、赵菁华、杨震、赵孙平、李洪亮、蔡立志。

中文语音识别终端服务接口规范

1 范围

本标准规定了在终端设备(如移动电话、平板电脑、无线音箱、车载导航仪和机器人等)上的中文语音识别系统所提供服务的功能集合,包括服务系统能力范围、输入/输出数据描述和服务接口的基本要求。

本标准适用于终端设备上的中文语音识别系统的设计和开发。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 13000 信息技术 通用多八位编码字符集(UCS)

GB 18030 信息技术 中文编码字符集

RFC 5646 语言识别标签(Tags for identifying languages)

3 术语和定义

下列术语和定义适用于本文件。

3.1

前端信号处理 frontend signal processing

在语音终端进行语音识别处理前,对输入声音信号进行增强、检测和规整处理的过程,其有助于提高语音识别终端服务系统在复杂环境中的性能。

3.2

目标语音 target speech

用户对语音识别终端所发出的有用语音,即希望语音识别终端服务系统识别的语音信号。

3.3

背景噪声 background noise

语音识别终端应用环境中的噪声。

3.4

干扰声音 interfering sound

语音识别终端应用环境中的声音,会引起语音识别的错误或虚警。

3.5

房间混响 room reverberation

室内声源发出的声波在各方向来回被房间壁面反射,而又逐渐衰减的现象。

3.6

声学回波 acoustic echo

由语音识别终端上的扬声器所发出的,又被该终端上的语音采集设备所采集到的声音信号。