

ICS 35.040
L 71



中华人民共和国国家标准

GB/T 21024—2007

中文语音合成系统通用技术规范

General specification for Chinese speech synthesis system

2007-06-29 发布

2007-11-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会

发布

目 次

前言	I
引言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 中文语音合成系统的功能分类	4
4.1 概述	4
4.2 从语言处理能力上分类	4
4.3 从所处区域或方言上分类	4
4.4 从合成的应用范围上分类	4
4.5 从韵律和音色调节能力上分类	4
4.6 从合成方法上分类	4
4.7 从应用平台上分类	5
5 中文语音合成系统数据交换格式	5
5.1 概述	5
5.2 元素与属性	6
5.3 输入要求	17
5.4 输出要求	17
5.5 音频数据格式要求	18
5.6 一致性	18
附录 A (规范性附录) CSSML V1.0 Schema 定义	19
参考文献	31

前 言

本标准的附录 A 是规范性附录。

本标准由中华人民共和国信息产业部提出。

本标准由全国信息技术标准化技术委员会归口。

本标准由安徽中科大讯飞信息科技有限公司、中国科学院自动化研究所、中国电子技术标准化研究所起草。

本标准主要起草人：尹波、严峻、吴晓如、陶建华、吴志刚。

引 言

本标准的发布机构提请注意如下事实,声明符合本标准时,可能涉及第 5 章中有关中文语音合成系统数据交换方法的相关专利。

本标准的发布机构对于专利的范围、有效性和验证资料不提出任何看法。

专利持有人已向本标准的发布机构保证,他愿意同任何申请人在合理和非歧视的条款和条件下,就使用授权许可证进行谈判。在这方面,该专利持有人的声明已在本标准的发布机构备案。有关资料可从以下地址获得:

安徽中科大讯飞信息科技有限公司

安徽省合肥市国家级高新技术产业开发区信息产业基地讯飞语音大厦 230088

请注意除上述已经识别出的专利外,本标准的某些内容有可能涉及专利。本标准的发布机构不承担识别这些专利的责任。

中文语音合成系统通用技术规范

1 范围

本标准规定了中文语音合成系统的术语定义、分类标准、数据交换格式标准和应用规范。
本标准适用于各种计算机、网络 and 智能设备配置的中文语音合成系统。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件,其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本标准,然而鼓励根据本标准达成协议的各方,经过研究决定是否使用这些文件的最新版本。凡是不注日期的引用文件,其最新版本适用于本标准。

GB 2312—1980 信息交换用汉字编码字符集 基本集

GB 13000.1—1993 信息技术 通用多八位编码字符集(UCS) 第一部分:体系结构与基本多文种平面(idt ISO/IEC 10646-1:1993)

GB 18030—2005 信息技术 中文编码字符集

RFC1766 语言识别标签

SSML 1.0(W3C) 语音合成标记语言

3 术语和定义

下列术语和定义适用于本标准。

3.1

语音合成 **speech synthesis**

通过机械的、电子的方法合成人类语言的过程。

注:该过程所产生的语音称为合成语音,和人的发音器官产生自然语音相区别,有时也叫人工语音(artificial speech)。

3.2

语音合成系统 **speech synthesis system**

由特定算法、模块和数据资源等构成、用来产生合成语音的软硬件系统。

注:由于目前语音合成最常用的实例是将智能设备中的文字信息转换为语音信号,以作为智能设备的语音输出形式,语音合成系统也常常称为文语转换系统(text-to-speech system),但作为语音合成系统的输入,实质上不限于文字,可以是概念、意图、思想等。在目前的技术水平下,文语转换是实现语音合成的主要方式,本技术规范中不另作特别说明时,文语转换将等同于语音合成的概念。

3.3

单语种语音合成系统 **mono-lingual speech synthesis system**

针对某一语种设计的,功能限于单一、特定语种的语音合成系统。

3.4

多语种语音合成系统 **multi-lingual speech synthesis system**

可以实现两种及两种以上语言语音合成功能的语音合成系统。

3.5

有限词汇语音合成系统 **limited vocabulary speech synthesis system**

为某一特定任务设计的、只能完成有限词汇量语音合成功能的语音合成系统。