

# 语音合成技术的现状及发展前景

沈晋苑

(北京广播学院应用语言学系00级, 北京 100024)

E-mail: jinyuansh@sohu.com

**摘要:** 语音合成技术(Text To Speech)是中文信息处理领域的一项前沿技术。

本文介绍了从采用参数合成方法到基音同步叠加, 语音合成的研究运用了基于语音数据库的语音合成方法和运用大量的统计和规则的方法, 使之较好地解决了三个大的处理环节: 特殊符号处理, 分词处理和拼接处理; 一定领域内实现了中英文混读的高自然度和高流畅度。而使用数据驱动模式生成自然语音的技术, 基于新的韵律结构分析和语音建模技术, 采用了分散式语音识别和语音合成技术, 使得系统的语音合成更流畅、更平滑、更自然。

语音合成技术在现阶段已经取得了长足的发展, 但是由于还处于技术研究开发的初级阶段, 因此还存在着许许多多急需解决的问题。基于语音数据库的语音合成方法有望进一步提高语音合成的自然度; 近年来提出的基于LMA(对数振幅近似)技术的语音合成器, Hybrid Harmonic/Stochastic模型, Sinusoidal模型等已被证实是一些新颖的能合成出高质量语音的参数合成方法, 为此应继续深入这方面的研究, 以期在参数合成技术上取得突破; 为了适应社会的需求, 扩大文语合成的应用场合, 既要求要提高语音合成的质量, 又要降低语音合成的复杂度; 多语种的文语合成以及多方言文语转换, 需求共用一种合成算法或语音合成器, 必须有新的思路。

本文从计算机模拟人的大脑神经活动系统的功能, 把人大脑活动中枢对声音信息的接收、加工分析、传递等功能与计算机应用结合起来, 提出基于认知解释的计算分析使语音合成语法规则具有可操作性的思考: 即把计算机语音的信号编码基于韵律的音素重组和连接, 运用电位信号的自然脉冲输送方法, 并采用一定的数据算法来实现语流的平滑度、流畅度和自然度。这也许是一种可实现的计算途径。自然语言是一种基于“声音—听觉识别”的信息系统, 人耳能清晰地加以分辨的声音元素(即音素, phone)不过五十个左右, 但它具有很大的能产性。我们在语音合成中借助语言信息的编码原理, 既可降低数据库的开销, 又可实现语言表现力的丰富。方言在很大程度上也是原型范畴, 是人们根据不同方言点之间在语音、词汇和语法等方面各种程度不同的相似性而概括出来的基于原型的语音相似性分类。我们可以根据边界模糊性、相似性与原型规则, 来达到语音合成的合理运用。最后, 本文就语音的无序性和规则性、瞬时性、记忆功能和遗忘筛选、基于声音的数据库链接阐述了自己的一些思考。

展望未来, 由于国际国内语音合成技术的不断取得突破, 我们不难想象未来的语言世界, 语音合成与语音识别、语音翻译的共同运用将使人类步入一个崭新的领域去生活。

**关键词:** 语音合成 文语转换 信号编码

**作者简介:** 沈晋苑(1982—), 男, 云南昭通人, 大学二年级学生, 现在北京广播学院应用语言学系学习。