

# 语音合成语料库管理系统的研究与设计

## The design and research of speech-synthesis management systems

艾克拜尔

新疆维吾尔自治区经济信息中心 邮编: 830002

赵世霞 蔡莲红

清华大学计算机科学系 邮编: 100084

**摘要:** 本文主要叙述了语料及其管理系统的研究与设计, 用最新的开发工具和现有软件达到系统的设计目标。

**关键词:** 语料, 语料库, 语音合成, 数据

### 1. 前言:

语音合成是计算机智能接口的一个重要分支, 而语料是进行语音识别、合成的起点。语料是指包括有可能用来发现和研究语言学 and 语音学知识的一切素材。随着计算机技术及相关电子技术的发展, 十多年来, 无论是在国际还是在国内, 在语音识别和合成技术方面有了很大的发展。目前, 不论是在计算机技术方面, 还是在语音学研究领域里, 对这一课题的研究投注了很大的人力物力。研究人员各自依据自己的方式方法展开对语料的样本采集和语料设计的研究, 给这一研究领域带来了很大不便, 也浪费了大量的人力和财力。因此, 建立一个能综合反映目前研究状况的、从不同出发点设计、并为不同研究目的服务的语料库势在必行, 它能给研究者提供丰富的研究资料。与此同时, 为满足研究者不同的需求, 应该建立一个语料丰富、能够统一管理、应用目的广泛、查询统计能力强、可视界面良好、快速、简捷、方便的语料库管理系统。这就是我们研究、设计<<语音合成语料库管理系统>>的目的。

### 2. 管理系统的研究

随着计算机多媒体技术的飞速发展以及基于窗口的操作系统 Windows 的问世与不断完善。我们可以利用 Windows 的强大功能把管理系统和原始语料数据有机地结合起来。这对我们所研究的语料库管理系统而言, 无论对语音识别还是对合成都具有非常重大的意义。语料库数据是丰富多采的。它不仅有我们常见的格式数据、文本数据, 而且还有与其相对应的非格式化的波形数据和声音数据。它们由不同音节长度的词组组成。因此, 建立语音合成语料库管理系统, 首先要确定我们建立这样一种管理系统的出发点和所遵循的原则:

#### 2.1. 出发点:

语音合成语料库管理系统不只是一个传统的单一数据格式的数据库管理系统。而且还

是一个面向对象的多媒体系统。它要具备外存管理数据可靠性和共享性、对象的继承性和计算的完备性等特点。要求满足图一所示的系统结构和图二所示的查询(访问)模式。我们知道,语料库中的四种数据(. wrd, .txt, .wav, .phn)虽然是相互独立的,但是我们可以通过 . wrd 库将其它三类库有机地联接在一起。即使用多媒体技术提供的数据动态链接法来管理数据(对象实体)。用容错技术和动态提示的方法来达到规范管理,预防系统出现的异常现象。

图 1. 系统结构:

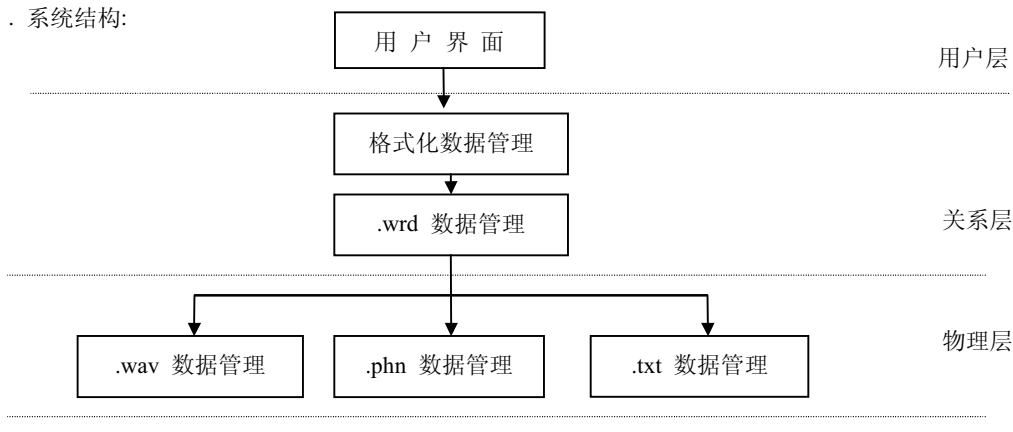
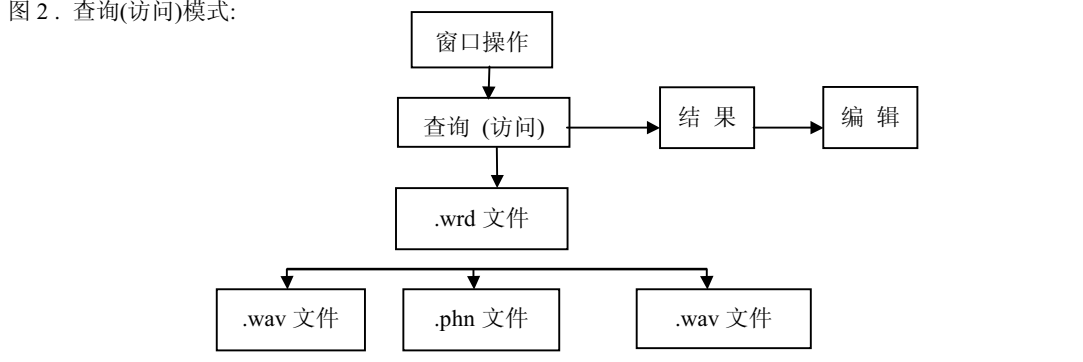


图 2. 查询(访问)模式:



## 2.2. 原则：我们建立这种系统的原则有如下几点

第一点,用优秀的开发工具:在数据库系统领域,除了大型数据库象 Sybase 这样的开发工具以外,还有应用非常普遍,绝大多数 PC 用户非常熟悉且容易掌握的开发工具 FoxPro。目前, Visual FoxPro 3.0 For Windows 由于其方便.灵活.快速.美观.可移植性好.功能强大.与传统开发工具 Xbase 兼容的特点,又有类似于 Windows 的可视界面,已经广泛应用于多媒体数据库管理系统。另外, FoxPro 还提供了 memo 和 general 两种字段,来存储位图和变长数据(声音,波形等)。因此, Visual.FoxPro 3.0 For Windows 作为支持语料库管理系统的开发平台是非常适用的。

第二点,用集成方法实现软件共享:在 Windows 系统里有很多现成的用于编辑图形和声音的软件,如: PaintBrush 和 SoundRecord。这些软件功能齐全,容易操作。假如,我们采用集成方法,用嵌入链接(OLE)技术把上述软件集成到系统里,不仅减少了很多重复性的工作,而且可达到事半功倍的目的。

第三点, 数据(库)可扩展, 增强系统寿命: 虽然系统所采用的语料数据是通过不断采集、不断筛选、不断更新得到的, 反映了目前语料数据建设的总体现状。但是, 在今后的语料研究工作中, 将对语料数据提出更多的要求, 即随着研究工作的不断深入和拓宽, 会对原始语料提出增减或更新的要求。这就使得系统具有对数据的扩展功能, 用户可根据自己的需求对语料数据进行增减、更新工作。这样, 既方便了用户, 又增加了系统寿命。

### 3. 系统设计

#### 3.1. 设计目标

语音合成语料库管理系统是一个多媒体数据库管理系统, 是基于 Visual Foxpro 3.0 For Windows 平台开发的, 并利用现有软件集成于该管理系统, 使之有效地达到系统目标, 该管理系统的最终设计目标是为用户提供一种简单易懂的可视界面, 用户可随意操作, 去完成工作, 达到目的。那么, 系统要实现的目标是:

- a. 良好的用户界面: 从用户角度设计界面。要求界面友好, 一目了然, 无师自通, 免除相关培训。
- b. 模块的独立性: 模块化程序设计可降低系统的复杂性, 减少系统开销, 使系统容易修改。因此, 把每一步操作作为一个独立的模块通过消息事件驱动来完成。这样, 既能达到(保证)模块间的独立性及其相互调用, 又能实现功能交叉, 如: 在检索操作中调用统计模块, 反之亦然。
- c. 用户(外部)数据库: 我们最终的产品是要存储在 CD 里的。那么, 用户如何处理编辑过的多媒体数据? 怎样为编辑过的数据开辟新的存储空间? 正是建立用户(外部)数据库的目的所在。

#### 3.2. CD 的组织及其内容

.wav 数据和 .phn 数据不同于一般数据, 其数据量是非常庞大的, 必须要用 CD 介质来存储语料数据。如何组织 CD 也是我们着重考虑的问题。组织好 CD 将会给管理系统带来许多方便, 提高对数据的检索查询速度更是不可估量。我们是这样组织 CD 的: 它由四部分组成, 第一部分是系统说明, 第二部分是可执行的管理系统, 第三部分是一些格式转换工具, 第四部分是数据。如下所示:

#### 3.3. 数据结构

##### 3.3.1. 波形和语音数据的数据结构:

我们把每一个与字词(.wod)数据相对应的波形和语音数据都定义成一个文件, 存取时都按文件存取。其结构如下:

File:= < File-ID > < data >

其中: “ File-ID ” 表示文件标识, 它是一串有效编码和文件名的组合, 编码表示文件(数据)的存储地址, 是按事先约定给出的, 每一个编码唯一指定一个数据文件。“ data ” 是指对应与字词数据的波形或语音数据。

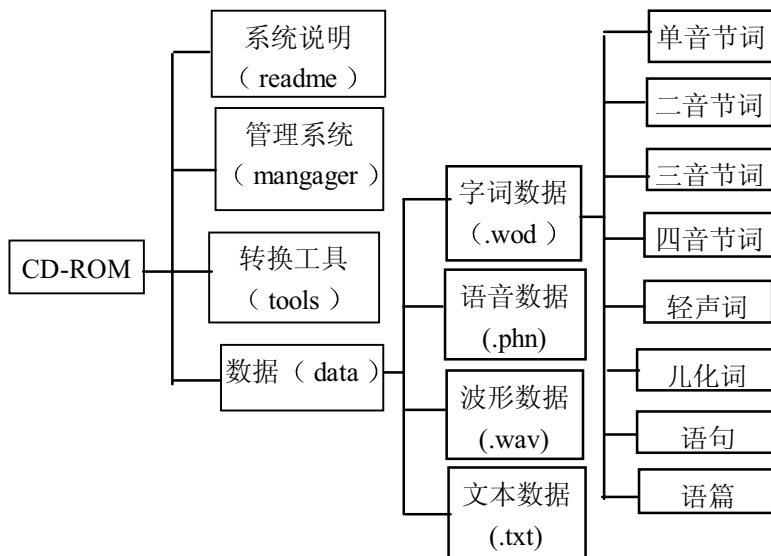


图3 CD的组织

### 3.3.2. 字词数据:

以二音节词为例, 其结构如下:

数据结构:= < Word > < Shdzh > < PY > < S1 > < Y1 > < D1 > < S2 > < Y2 > < D2 >  
< File-ID >

其中: “Word”表示字词, “Shdzh”表示声调搭配, “PY”表示字词的拼音加声调, “S1”表示第一个字的声母, “Y1”表示第一个字的韵母, “D1”表示第一个字的声调, “S2”表示第二个字的声母, “Y2”表示第二个字的韵母, “D1”表示第二个字的声调, “File-ID”同上述。

### 3.3.3. 文本数据:

根据需要通过转换工具 Tools 将文本数据转换成所要求的格式。

### 3.4. 系统功能及流程图:

首先, 系统要具备一般系统的功能, 例如: 查询, 浏览, 显示, 统计, 编辑, 输出等。除此之外, 还要具备剪切, 复制, 粘贴, 放音等用于多媒体数据的功能。实现上述功能需要用到一些特殊的技术, 特别是语音切分和链接嵌入 (OLE)。这些技术的合理应用将对实现系统功能起到关键作用, 也是判断系统优劣的重要标志。

系统流程图:

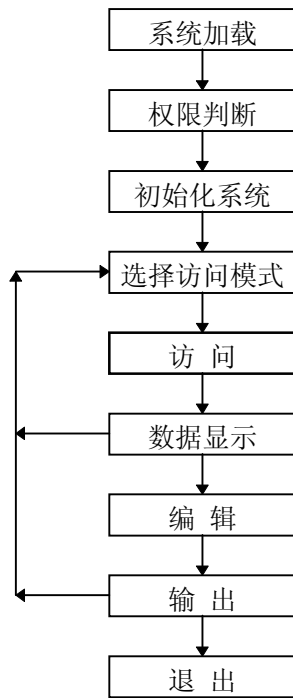


图 4 系统流程

#### 4. 结束语

这个任务是 863 项目的一个子任务。目前，已基本完成该管理系统的编程工作。我们将本着向用户负责的态度，遵从“以客为尊，用户至上”的原则，以应用为本，以需求为驱动力，规范技术支持标准，根据实际需要，不断改进系统，完善系统功能，使用户达到最大满意。使系统更具生命力。

#### 参考文献：

- [1] 朱瑞康等 CSCP 汉语文本—语音合成系统 小型微型计算机系统 1996 年 7 期
- [2] 肖健宇等 面向对象的多媒体数据库存储管理子系统的研究 计算机工程与科学 1996 年 7 期  
第 18 卷第 A3 期 62 页
- [3] 李国辉等 多媒体文献存储结构的研究 计算机工程与科学 1996 年 7 期第 18 卷 第 3 期 69 页
- [4] 曾晓勤 多媒体数据库：建模、存储和查询 计算机工程与应用 1996 年 1 期 1 页
- [5] 顾宁等 多媒体数据库管理系统 MEDB 的设计 多媒体世界 1994 年 7 期
- [6] 黄日日 面向对象数据库系统之我见 计算机科学 1995 年 Vol.22 No.1, 65 页
- [7] 胡晓峰等 多媒体系统原理与应用 北京：人民邮电出版社 1995 年 6 月
- [8] 陈诗嫒 FoxPro 2.6 For Windows 应用系统设计 北京：清华大学出版社 1996 年 3 月
- [9] 汪成为等 面向对象分析、设计与应用 北京：国防工业出版社 1992 年
- [10] 郑人杰等 实用软件工程 北京：清华大学出版社 1997 年
- [11] (美) Jeb Long 京京翻译组译 VISUAL FOXPRO 3 开发指南 北京：机械工业出版社

**ABSTRACT:** One of the 863 projects of China is considered in this paper. We mainly describe the design and research of phonetic-materials and their management system. We come to the designed aim by the recent development implements and software.

**KEY WORDS:** phonetic-materials, phonetic-materials base, phonetic synthetics, data.