

智能语音图书的设计与实现

彭崇梅^{*1} 徐雪连¹ 李全圣¹ 袁国顺¹ 陈正清²

(1. 中国科学院微电子中心 北京 100029; 2. 深圳市语音实业有限公司 深圳 518133)

【摘要】智能语音图书是以语音芯片FS20C01作为核心处理器的应用方案。该文介绍了智能语音图书的优点以及系统功能结构的组成,系统的硬件结构构成,并给出了具体学习、考查、帮助功能实现过程中卡片识别、按键识别、评判帮助算法及基于FS20C01语音芯片的语音合成算法的软件实现方案。

关键词 语音图书; 语音处理; 线性预测编码; 码本激励线性预测; 语音芯片
中图分类号 TP399

Design and Implement of Intelligent Voiced Book

Peng Chongmei¹ Xu Xuelian¹ Li Quansheng¹ Yuan Guoshun¹ Chen Zhengqing²

(1. Microelectronics Research and Development Center of the Chinese Academy of Science Beijing 100029;
2. Shenzhen Voicing Industrial Corp.Ltd Shenzhen 518133)

Abstract Intelligent voiced book is the application technique based on the core processor the voice chip FS20C01; This paper introduces the good qualities and system software structure constituting of the Intelligent voiced book; explains the system hardware constituting, narrates the arithmetic of the card identification, keystroke identification, judging help in the implemented procedure of the studying, examining and helping function; then introduces the software implemented scheme of the voice synthesis arithmetic based on the voiced chip FS20C01.

Key words voiced book; voice processing; linear predictive coding; code excited linear predictive; voice chip

语音智能图书是一种会发音的、图文并茂的智能型电子产品,它可以代替传统的印刷图书,也有别于电脑光盘等产品。相对于后两类产品,该图书具有以下优点:1)声、图、文、乐并茂,能丰富儿童的视觉、听觉、触觉感受,不会因电脑射线影响儿童的视力发育;2)采用ROM卡存储,内容丰富,涉及学科领域广泛,包括自然常识、学拼音识汉字、幼儿英语、智力游戏、地理常识、唐诗宋词等诸方面,图片印刷精美且存储量大;3)用人机对话的方式代替传统的面对面教授方式,能消除儿童学习心理障碍;4)操作简单方便,容易培养儿童学习兴趣。

1 语音智能图书的功能结构及硬件组成

语音图书系统可分为学习系统、考查系统和功能系统三个部分^[1]。系统的功能结构如图1所示。该系统以语音芯片FS20C01做为核心处理器,其硬件构成示意图如图2所示^[2]。

2002年3月4日收稿

* 女 24岁 硕士 助理研究员

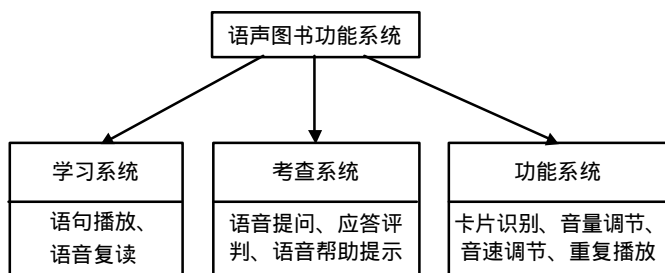


图1 语声智能图书的功能结构

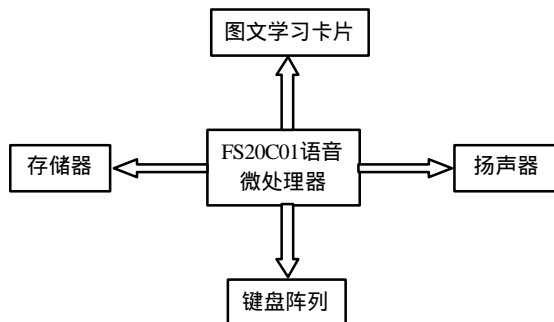


图2 语声智能图书系统硬件构成示意图

2 语声智能图书的实现技术

语声智能图书是基于FS20C01语音微处理器实现的电子产品系统。硬件界面生动活泼、声文并茂且操作方便；软件则采用FS20C01语音微处理器的指令集开发，程序从语句发音清晰度及程序空间存储量两方面综合考虑，对重要但级别不同的语句分别采用LPC和CELP两种发音技术；程序还大量使用编表查表的方法，实现按键的智能化，节省了程序空间。

2.1 卡片识别及按键发音的算法

语声智能图书能识别插入的不同类型的发音卡片，识别的卡片数量最大为 2^N （ N 为用来判别卡片序号的I/O口线数），待卡片正确插入后发音卡头语，根据卡片上标识的键位，播放按下键位语句的发音。同时还可以进入考查提问评判模式和帮助系统，并能根据需要调整音量大小。算法流程如图3所示。

2.2 提问应答评判及帮助系统的算法

提问考查是对学习效果的一种测试，对儿童来说在学习考查过程中不断采用表扬、鼓励、提示的方式能够增强孩子的学习自信心，提高孩子的学习热情。提问考查及应答评判系统是根据所学卡片的内容顺序提出相对应的问题，再根据回答的按键给予表扬、鼓励或帮助。算法流程如图4所示。

2.3 线性预测编码LPC及码本激励线性预测编码CELP的算法

该语声智能图书的核心是FS20C01语音微处理器，它主要采用LPC和CELP方式进行语音合成，下面简略介绍FS20C01语音微处理器中应用的这两种语音合成方式的算法。

2.3.1 线性预测编码LPC的算法

线性预测编码是一种低速率的语音参数表示方法，该方法的特点是用最少的传输、存储数据来表征语音的特征参数。线性预测的基本思想是根据最小均方误差的准则，用过去的若干语音数字信号来预测当前的语音信号，将估算出的语音参数送到语音合成模型后输出语音^[3]。合成器首先要判断帧类型，然后对接收的数值进行转换并在新旧帧之间平滑插值。时间寄存器Timer的量值用来决定何时调入新的一帧以及用作帧间插入的比例系数。

LPC的算法流程如下：

- 1) 设定语音合成算法为LPC方式。
- 2) 调入前两帧语音数据，设定时间计数器的门限值和比例参数。同时计算第一次帧间插入的参数，送往语音合成模型格型滤波器，然后使能LPC方式。
- 3) 判断是否有音量调节键，若是调用音量调节子程序，否则继续。
- 4) 判断是否有退出键按下，若是跳转步骤8)，否则继续。
- 5) 判断时间计数器是否达到门限值，若是调入新的一帧参数，否则继续。

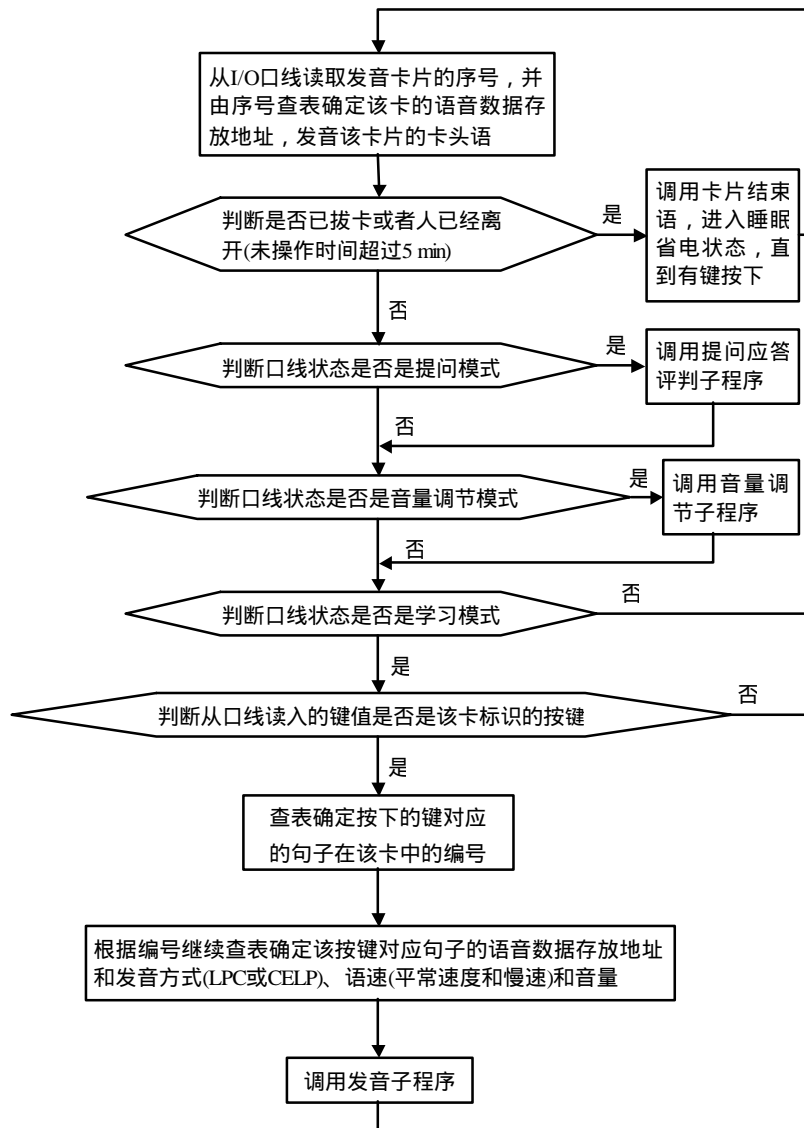


图3 卡片识别及按键发音的算法

6) 判断是否满足帧间插入的条件, 若是则进行帧间插入计算并送结果入语音合成模型格型滤波器, 否则继续。

7) 判断语音合成是否结束, 即LPC方式是否关闭, 若是跳转步骤8), 否则跳转步骤3)。

8) LPC合成算法结束, 返回主程序。

线性预测系数、增益、基音周期等参数值每个基音周期更新一次, 清音帧、浊音帧及静音帧间相互转换时因帧的性质不同而不进行插值更新。浊音和清音的激励信号是按基音周期从芯片内部ROM区取出并输入到音调寄存器, 与增益相乘后送到LPC滤波器^[4]。

2.3.2 码本激励线性预测编码CELP的算法

码本激励线性预测是用码本作为激励源的线性预测编码技术, 它把残差信号可能出现的各种样值的组合按一定的规则排列后存放在码本存储器中, 每一种样值的组合有一地址码。CELP算法是用从地址码本中搜索出的最佳码矢量乘以最佳增益来代替线性预测余量信号作为激励信号源。它采取分帧技术进行编码, 帧长一般为20~30 ms。CELP一般将每一语音帧分成2~5个子帧, 在每个子帧内搜索最佳的码矢量作为激励信号。

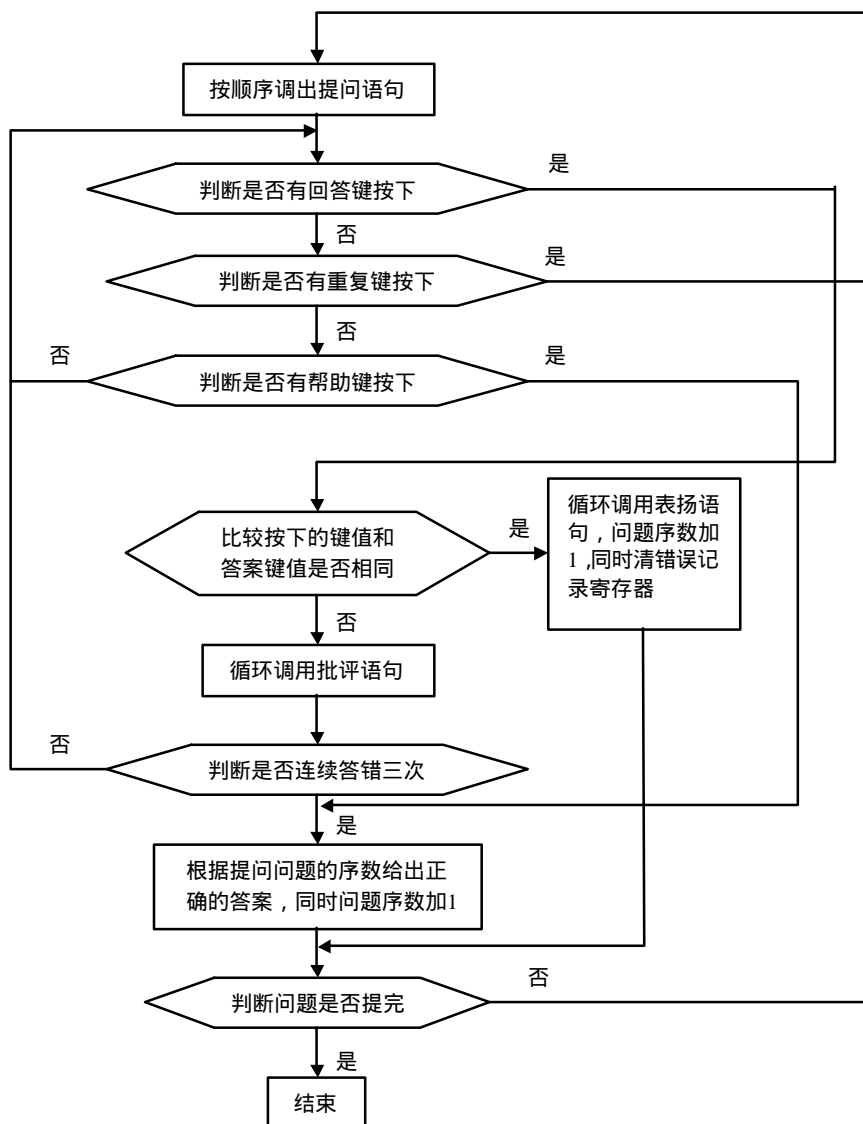


图4 提问应答评判及帮助系统的算法

CELP用一个自适应码本中的码字(码矢量)来逼近语音的基音信号,用一个固定的随机码本中的码字(码矢量)来逼近语音的随机激励信号。从两个码本搜索出来的最佳码矢量乘以各自的最佳增益后相加,其和即是CELP的激励信号源。将激励信号输入10阶LPC格型滤波器 $1/A(Z)$ 及低通滤波器后 $1/P(Z)$,得到合成语音信号 $s(n)$, $s(n)$ 与原始语音 $s(n)$ 的误差经感觉加权滤波器,得到感觉加权误差。CELP用感觉加权的均方误差最小作为搜索最佳码矢量及其幅度的度量准则^[5]。

在CELP算法中,格型滤波器的 K 参数及浊音帧的能量参数每帧传送一次。对于每个子帧,这些参数按新旧帧的数值进行平均加权,即若有3个子帧,则按:插值=旧值+(新值-旧值) $(i/3)$ 的模式进行插值, $i=1,2$ 。程序按照调入新帧、判断帧的属性、插入参数值一、插入参数值二,再调入新帧,再插入的流程循环直至结束。

在调入新帧及插值后,主程序不断地循环查询规定的采样点数是否已经到达,中断程序以8 kHz的间隔,在对采样点寄存器进行累加的同时,按照最佳码矢量标号,查出最佳码矢量乘以最佳增益后向激励寄存器输入激励码。当前子帧的自适应码本是由在此之前的激励矢量构成的140(EFH-64H)个元素系列。当码本搜索完成后,用当前子帧的60个激励码矢量值(由搜索到的最佳自适应码矢量和最佳随机码矢量经过最佳增益加权后得到)更新自适应码本,作为下一子帧的自适应码本。

CELP的算法流程如下:

- 1) 设定语音合成算法为LPC + PCM方式。
- 2) 判断是否达到规定的采样点数,若是进行CELP语音合成,具体步骤如图5所示,否则继续。
- 3) 判断是否有音量调节键,若是调用音量调节子程序,否则继续。
- 4) 判断是否有退出键按下,若是跳转步骤6),否则继续。
- 5) 判断CELP合成是否结束,若是跳转步骤6),否则跳转步骤2)。
- 6) LPC合成算法结束,返回主程序。

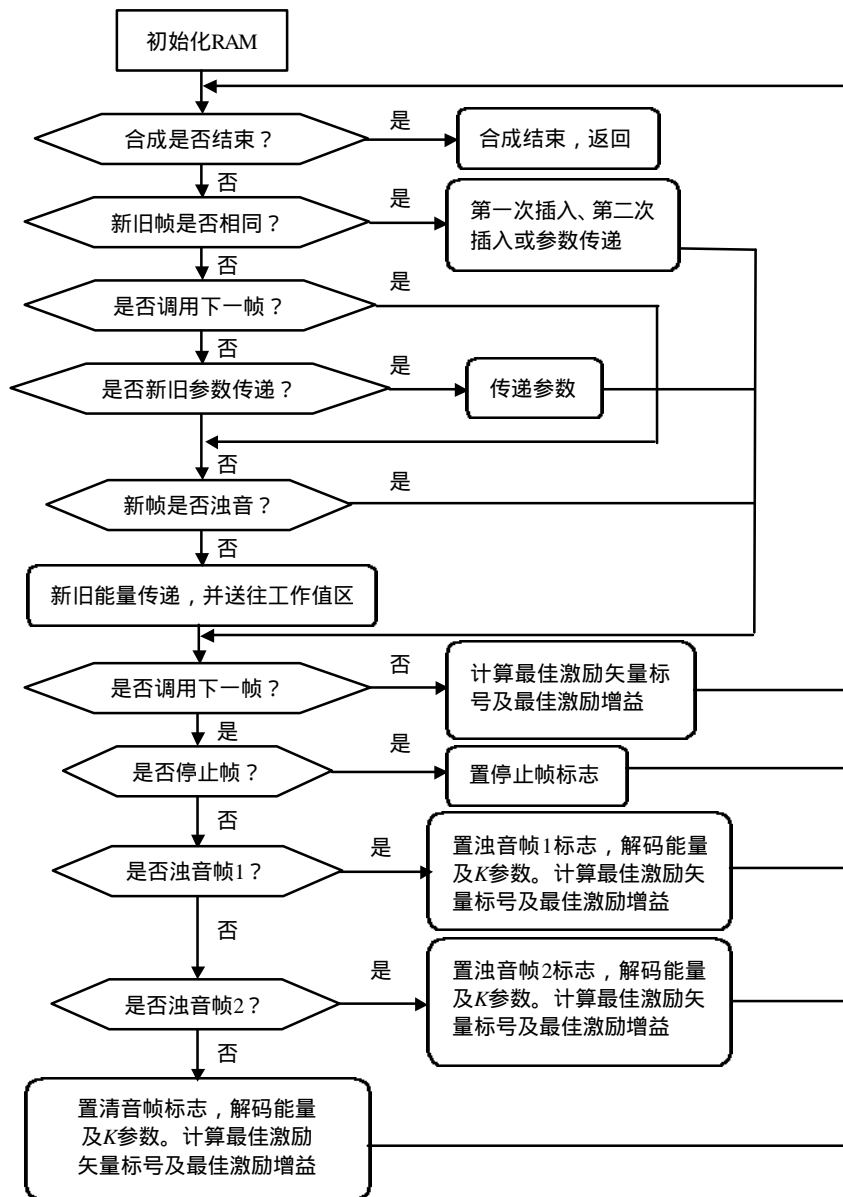


图5 CELP语音合成流程图

3 结束语

该语音智能图书系统基于语音芯片FS20C01所开发,发音清晰准确、内容丰富生动,画面活泼多样,它益智有趣,集教育、语音、音乐于一体,适合于儿童的综合智力开发。本系统设计实现的卡片按键识别、评判帮助和语音合成的算法实用性较强。

参 考 文 献

- 1 Zhang Jiang, Huang Diming, Liao Jianming. Design and Realization of General Examination System. Journal of University of Electronic Science and Technology of China, 2001, 30(2): 157-160 [张江, 黄迪明, 廖建明. 通用考试系统的设计与实现. 电子科技大学学报, 2001, 30(2): 157-160]
- 2 He Xinggao, Qin Zhiguang, Li Lingfei. The Design and Implement of Image File Reliable Transmission Software. Journal of University of Electronic Science and Technology of China, 2001, 30(6): 609-612 [何兴高, 秦志光, 李凌飞. 图像文件可靠传输软件的设计与实现. 电子科技大学学报, 2001, 30(6): 609-612]
- 3 [美]拉宾纳L R, 谢弗R W著. 语音信号数字处理. 朱雪龙译. 北京: 科学出版社, 1983
- 4 彭崇梅, 李全圣. 对线性预测语音解压缩方法的探讨. 电子产品世界, 2001, (4B): 11-13
- 5 杨行峻, 迟惠生. 语音信号数字处理. 北京: 电子工业出版社, 1995

· 科研成果介绍 ·

纳米晶稀土永磁材料应用研究

主研人员: 王豪才 杨仕清 蒋洪川 张万里 张文旭 彭斌 王瑞芳

纳米晶稀土永磁材料应用研究主要采用熔体快淬工艺, 在特种配方下, 其材料的主要指标达到: $(BH)_{\max}=101 \text{ kJ/m}^3$, $B_r=0.842 \text{ T}$, $H_{cJ}=777 \text{ kA/m}$, B_r 的温度系数优于 -0.1% ($20 \sim 80^\circ \text{C}$), $T_C > 600 \text{ K}$ 。利用厚膜烧结与复合涂层技术, 使磁体能耐氧化、酸碱腐蚀, 尤其能使磁体不发生氢脆, 可以应用于恶劣环境中, 具有重要的使用价值。利用化学镀方法在磁体表面镀一层Ni/P合金层, 成功地解决了稀土粘结磁体的防腐难题。根据所研究的微特步进电机的步进角度, 通过设计与制作充磁夹具, 实现了单根导线脉冲电流径向多极充磁, 使磁体用于微特电机中静态最大转矩为 $1.0 \sim 1.2 \text{ N} \cdot \text{cm}$ 。该项成果在交换耦合纳米磁体微磁学计算、熔体快淬低钕纳米晶交换耦合稀土永磁材料、纳米晶稀土粘结磁体在微特电机中应用的磁路分析、粘结磁体化学镀镍等方面有一定的创新。

军用高压智能功率控制电路

主研人员: 李肇基 方健 李鸿雁 张波 季晓宇 杨健 李学宁 李泽宏 庄圣贤 刘启宇 闫飞 杨舰

军用高压智能功率控制电路提出采用低能 He^+ 离子注入局域寿命控制技术来提高IGBT的关断速度。其局域性和热稳定性好, 且与集成电路工艺兼容; 研制出满足耐压要求的LDMOS器件; 研究了SPIC中高低压结隔离、高低电平转换和高压兼容以及电路保护技术。利用上述技术研制成功单相智能功率驱动集成电路, 同时还进行了高压SOI-LDMOS器件及全介质隔离技术的先期研究, 并在缓冲层的结构上取得了突破性进展。

软件无线电跳频模块技术

主研人员: 程郁凡 李少谦 雷霞 刘晓峰 王灿 余宏 张平 付斌

软件无线电跳频模块技术采用软件无线电的设计思想, 研究了可应用软件无线电台的软件跳频模块, 实现跳频电台跳频模块的全部功能。该成果具有良好的扩展性和灵活性, 改变相关参数软件后, 能兼容其他跳频电台。并且还可以应用于其他软件无线电的设备中。

· 科 卞 ·