

毕业论文

论文题目：基于STM32的智能语音播报系统的设计与实现

年 级：2015级

班 级：计算机科学与技术(嵌入式)一班

姓 名：毛思耀

学 号：20151104803

指导教师：李红霞

**目 录**

[1 引言 4](#_Toc533028381)

[2 系统整体方案分析设计 4](#_Toc533028382)

[2.1 技术需求分析 4](#_Toc533028383)

[2.2 具体功能实现分析 4](#_Toc533028384)

[3 硬件电路设计 5](#_Toc533028385)

[3.1 主控制器电路 5](#_Toc533028386)

[3.2 语音合成电路 5](#_Toc533028387)

[3.3 音频功放电路 5](#_Toc533028388)

[3.4 感光模块电路 6](#_Toc533028389)

[4 软件设计实现 6](#_Toc533028390)

[4.1 整体框架设计 6](#_Toc533028391)

[4.2 功能模块实现 6](#_Toc533028392)

[4.2.1应用软件的整体模块设计 6](#_Toc533028393)

[4.2.2 应用程序接口API的设计 6](#_Toc533028394)

[4.2.3 通信过程的设计 6](#_Toc533028395)

[5 主要实验流程 6](#_Toc533028396)

[6 总结与展望 6](#_Toc533028397)

[参考文献 6](#_Toc533028398)

基于STM32的智能语音播报系统的设计与实现

网络技术学院 2015级计算机科学与技术（嵌入式） 毛思耀 20151104803

指导教师 李红霞 老师

摘要 随着人工智能、大数据等前沿新兴技术的高速发展，带动了一大批与之配套的周边技术的兴起。正因此，TTS技术即语音合成技术近几年得到了迅速的发展，以语音合成、语音识别等为主的技术广泛运用于智能语音播报系统、智能语音控制系统等领域。基于此设计实现了 以STM32为核心，通过XF-S4240语音合成芯片将文本信息转换为语音信息的嵌入式智能语音播报系统。

关键词 STM32; 语音合成; 信号处理;

1 引言

在城市生活场合中，LED 显示屏和触摸屏已经不能满足复杂的人性化、自然化、方便化等服务。因此听觉交互成为了该领域一个非常热门的研究方向，而语音播报又是这里面一个很重要的分支。目前，语音播报系统广泛的应用于人们的日常生活，比如公交车的语音播报系统、银行的排队叫号系统等等。基于嵌入式 STM32 微控制器的语音播报系统和基于计算机的语音播报相比，具有能耗低、体积小、灵活性高、稳定性能优良和易携带等优点。通过语音合成芯片对文本解析成语音信息，可以达到语音效果清晰、准确、自然等效果，非常适用于当前流行的智能家居、交通服务等领域。

2 系统整体方案分析设计

## 2.1 技术需求分析

嵌入式语音播报系统的一个重要工作是语音的合成。语音合成，即 TTS ( Text to Speech) 技术，能将任意文字信息实时转化为标准流畅的自然语音并朗读出来，所要解决的主要问题就是如何将文字信息转化为可听的声音信息，即让机器像人一样开口说话。

语音合成的基本原理如图 1 所示。文本分析模块在文语转换系统中起着重要的作用，主要模拟人对自然语言的理解过程，使计算机对输入的文本能完全理解并给出后两部分所需的各种发音提示。韵律处理为合成语音规划出音段特征，如音高、音长和音强等，使合成语音能正确表达语意，听起来更加自然。这是语音合成中最重要的一个部分。要使得合成的语音符合通常说出的话语，最关键的是提取语言中的韵律参数。语音生成根据前两部分处理结果的要求输出语音，即合成语音。

## 2.2 具体功能实现分析

在本次设计过程中主要设计运用于两种场景，其一是大环境下的公共场所，例如火车站、公交车站、地铁站等人流量密集且需要长期固定播报语音的地方；其二是小环境下的私人场所，例如家里。系统设计以单片机为控制核心，配合语音芯片可以实现语音功能。当控制信号到来时，单片机接收到信号，即会给语音芯片发送控制信息，从而实现语音芯片相应的功能；采用单片机作为控制核心，对各个子模块控制，以实现需要的功能。语音控制芯片功能是将语音信号通过采样转化为数字，存储在IC的ROM中，再通过电路将ROM中的数字还原成语音信号。按键控制单元是通过按键发送信号给单片机，再通过单片机控制相应的模块，从而实现要求的功能。

3 硬件电路设计

嵌入式语音播报系统的硬件电路主要包括 3 个部分，即主控制器部分、语音合成部分和音频功放电路。上位机发送指令到主控制器部分，主控制器分析指令，通过 SPI 接口发送含文本信息的指令到语音合成板卡，语音合成板卡将文本信息转 换 为 语 音 信 息，送 往 音 频 功 放 电 路。期间，主控制器也会监测本系统的运行情况，向上位机返回当前系统的运行状态，并在紧急情况下进行语音提示。

## 3.1 主控制器电路

本系统选用STM32F103C8T6 作为主控制器。该芯片使用高性能的 ARM Cortex-M3 32 位的 RISC内核，工作频率最高可达 72 MHz，内置高速存储器( 64 KB 的闪存和 20 KB 的 SRAM) ，含有丰富的增强 I /O 端口和联接到两条APB总线的外设，完全满足本系统的需要。

## 3.2 语音合成电路

语音合成模块参照了科大讯飞的 XF-S4240 数据手册进行设计。XF-S4240语音芯片的主要特点是合成语音自然度高，控制接口简单方便，功能强大; 可以通过 USART、SPI接口和 I2C 总线 3 种方式接收待合成的文本 ，直接合成为语音输出; 采用双发音人，集成提示音效，支持软件调节语速、语调、音量。XF-S4240 中文语音合成板卡使用DC3.3V供电；采用 SPI 数据通信接口与主控制器连接，板卡在通讯当中设置为 Slave 身份，所需时钟信号由主控制器提供，即主控制器作为 SPI 通讯中的 Master 身份，传输的数据位数为 8 bits，每次合成的文本量最多达 1 K 字节; 由硬件 RDY 引脚指示板卡的工作状态，当引脚处于低电平时，表明板卡处于空闲状态，没有合成文本，当引脚处于高电平状态时，表明板卡处于合成文本的工作状态; RST为复位引脚，外接一个复位按键，板卡正常工作时，此引脚为高电平。

## 3.3 音频功放电路

音频功放电路，采用美国国家半导体的高效开关型低频功率放大器 LM4665MM，供电电压3. 3 V; 输出端省去了 D 类放大器特有的 LC 滤波器，直接接扬声器; 内部设有异或门控制电路提供关断控制功能，SDM 引脚接地电平，SD 引脚接高电平，放大器 处 于 工 作 状 态; GS 增 益 选 择 端 接 低 电平，集成电路增益为 12 d B。

## 3.4 感光模块电路

4 软件设计实现

## 4.1 整体框架设计

在整 个 软 件 系 统 中，语 音 播 报 任 务 App \_Task Speak 负责向 XF-S4240 发送含文本信息的语音合成命令。发送的命令和数据需要用“帧”的方式进行封装后传输，帧结构由帧头、待合成文本的长度、文本合成命令字、文本编码格式、待合成的文本组成。传输过程中，主控制器首先查询状态管脚 RDY输出工作状态，当引脚处于低电平时，表明板卡处于空闲状态，没有合成文本，可以发送文本信息; 当引脚处于高电平状态时，表明板卡处于合成文本的工作状态。传输时，SPI 使能选择信号 SSEL 保持低电平。程序中文本信息采用 GBK 编码。当 XF-S4240正在合成文本的时候，如果又接收到一帧有效的数据，板卡会立即停止当前正在合成的文本，转而合成所接收到的最新的文本。

## 4.2 功能模块实现

## 4.2.1应用软件的整体模块设计

流程如

## 4.2.2 应用程序接口API的设计

应用程序接口是用C语言设计的，

## 4.2.3 通信过程的设计

包括这两个等级都被不同的选择设备所支持。

5 主要实验流程

基于也必然将得到充分应用。

6 总结与展望

参考文献

[1] 谢希仁.《计算机网络》(第6版).北京:电子工业出版社，2013.

在本论文的写作过程中得到了李红霞老师的悉心指导，在此对老师表示衷心的感谢。