



在 STM32F101xx 和 STM32F103xx 微控制器上 使用 Speex 音频编解码器进行语音合成器演示

前言

本应用笔记说明了如何在 STM32F101xx 和 STM32F103xx 微控制器上利用编解码器 Speex 软件实现语音合成器应用。

Speex 是一款免费的音频编解码器，用于语音编码和解码。它能够对语音进行高度压缩同时保持良好音质。这让它成为那些使用留言回放或录音功能的应用的高性能解决方案，例如电话应答机、建筑和家庭安全系统、对讲机、智能家电、录音机或无绳电话等。

目录

1 Speex 编解码器概述 5

2 语音合成器应用 6

 2.1 语音合成器概述 6

 2.2 硬件描述 6

 2.3 Speex 编解码器实现 8

 2.3.1 Speex 编解码器设置 8

 2.3.2 Speex 编解码器优化 8

 2.3.3 Speex 编解码器要求 9

 2.4 语音合成器固件说明 9

 2.4.1 回放应用 10

 2.4.2 录音和播放应用 12

 2.4.3 环回应用 15

3 结论 19

4 修订历史 20



表格索引

表 1. Speex 实现要求 9

表 2. 文档修订历史 20

图片索引

图 1. STM32F101xx 和 STM32F103xx 语音合成器硬件说明 6

图 2. OLIMEX STM32F103-STK 音频输入 / 输出级 7

图 3. 回放应用流程图..... 11

图 4. 录音 & 播放框图 12

图 5. 主要的录音应用流程图 13

图 6. 主要的播放应用流程图 14

图 7. TIM3_IRQHandler 流程图..... 15

图 8. 实时环回原理 16

图 9. 环回流程图 17

图 10. TIM2_IRQHandler 流程图..... 18



1 Speex 编解码器概述

Speex 编解码器是一款开源的、取得专利并免版税的软件，用于语音压缩和解压缩。

Speex 基于 CELP（码激励线性预测），针对 2 至 44 kbps 比特率范围内的语音压缩而设计。

Speex 具有以下特点：

- 以相同的比特流进行窄带（8 kHz）、宽带（16 kHz）和超宽带（32 kHz）压缩
- 强度立体声编码
- 数据包丢失隐藏
- 可变比特率操作（VBR）
- 语音活动检测（VAD）
- 非连续传输（DTX）
- 定点端口
- 回音消除器
- 噪声抑制

Speex 具有很多其他编解码器所不具备的特点，例如可以进行强度立体声编码、在同一比特流中集成多种采样率以及具有 VBR 模式。

关于 Speex 编解码器的更多信息，请参考 Speex 网站：www.speex.org。

注：本应用笔记适用于 Speex 编解码器的 1.2rc1 版本。

2 语音合成器应用

2.1 语音合成器概述

语音合成器是一种语音处理应用，提供语音处理功能，如编码、解码、过滤和放大。

典型的语音合成器应用包括一个语音处理模块和音频输入 / 输出接口。

本应用笔记中，Speex 编解码器固件充当了语音处理模块，音频输入 / 输出接口则利用 STM32F101xx 和 STM32F103xx 的内嵌资源（ADC 作为输入，PWM 作为输出）集成。

STM32F101xx 和 STM32F103xx 微控制器系列产品具有 12 位分辨率的多通道 ADC 和 1 μ s 的转换时间，以及 16 位定时器并能产生 PWM。这些特点使其能够实现低成本且高性能的音频应用。

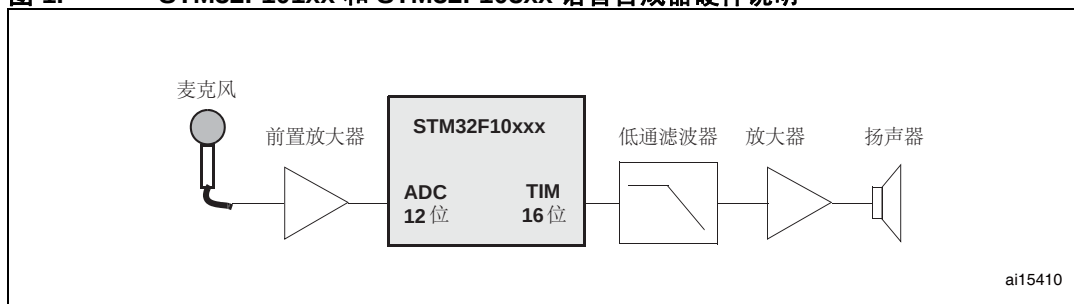
2.2 硬件说明

硬件由两个音频级组成：一个输入和一个输出，分别进行采集和播放语音（请参看图 1）。

本应用笔记中，语音采集由微控制器的内部 ADC 实现，该 ADC 必须通过一个前置放大器连接到麦克风上。

语音合成由微控制器的内部定时器实现，该定时器为 PWM 模式。PWM 输出信号进行低通滤波，消除了高频分量，然后经过放大并发送到扬声器。

图 1. STM32F101xx 和 STM32F103xx 语音合成器硬件说明



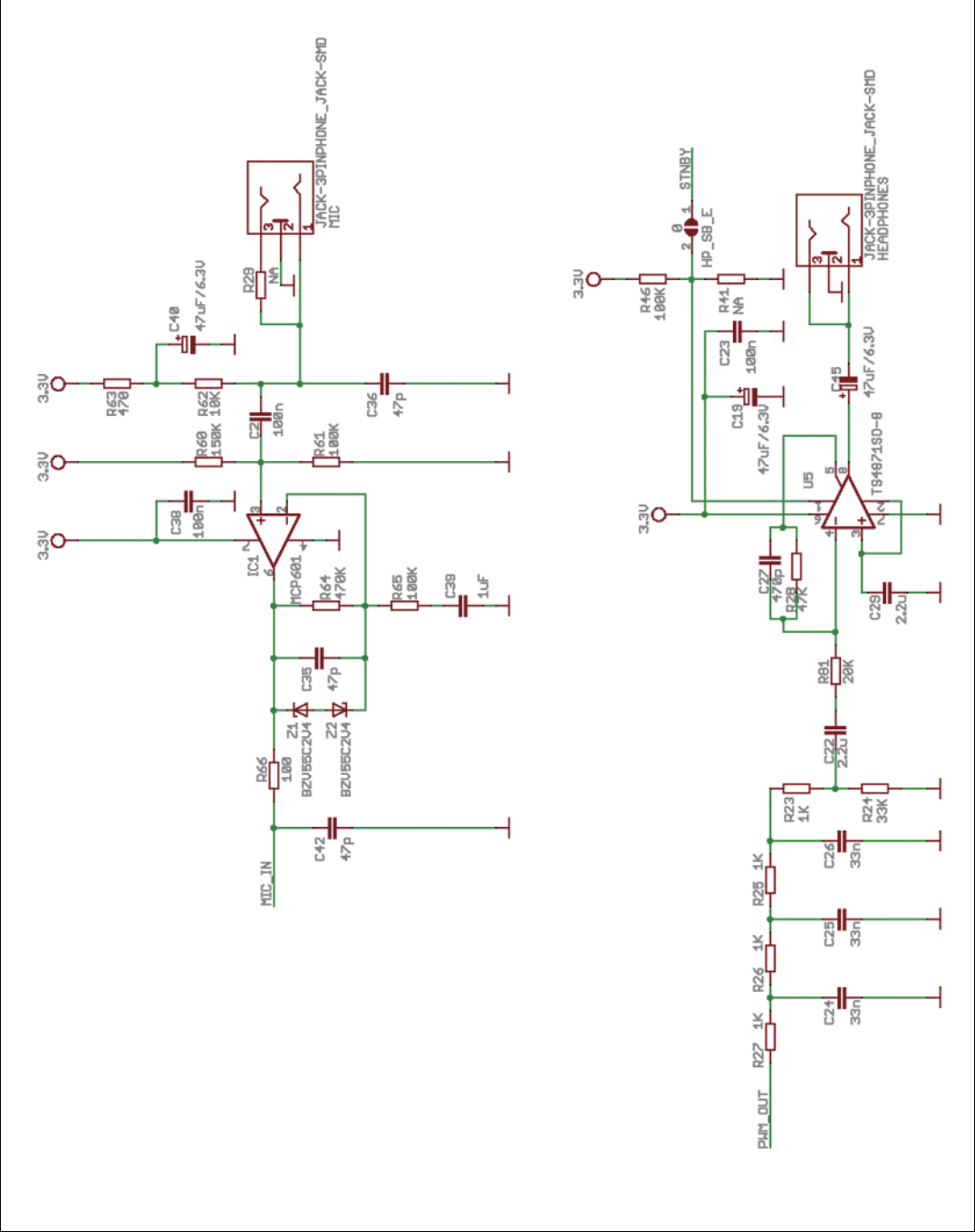
此应用说明适用于 OLIMEX（OLIMEX 网址为 www.OLIMEX.com）的 STM32F103-STK 板。

该板具有一个输入音频接口和一个输出音频接口，它们连接到两个插座上，而插座连接到耳机（或扬声器）和麦克风（参考图 2）。

应用于 OLIMEX 板上的装置为 STM32F103RBT6，它具有 128 KB 的 Flash 和 20 KB 的 RAM。

关于 STM32F103-STK 板的更多信息，请参考下面的网址：<http://www.olimex.com/dev/stm32-103stk.html>。

图 2. OLIMEX STM32F103-STK 音频输入 / 输出级



2.3 Speex 编解码器实现

2.3.1 Speex 编解码器设置

在此应用笔记中，Speex 编解码器支持窄带语音信号（8 kHz 采样），并实现改善 8 kbps 比特率的编码的作用。给定压缩比为 16:1。

Speex 编解码器用一个“质量”的参数来控制比特率。该参数在编码设置中进行配置。

要将比特率设为 8 kbps，质量参数须等于 4。由于此处 Speex 具有恒定的比特率（等于 8 kbps），质量参数也必须保持不变。

在本实现模型中，质量参数为 4，编码器的复杂度只能设为一个小于等于 2 的值。实际上，复杂度值大于 2 会导致编码解码器占用额外的 CPU 循环，这与期望的质量水平并不相符。为了获得最佳品质，建议将复杂度设为小于等于 2 的值。

2.3.2 Speex 编解码器优化

为了减少内存使用率并提高性能，Speex 编解码器进行了改进，保持 8 kbps 的恒定比特率且复杂度值小于等于 2，这样即可优化 STM32F101xx 和 STM32F103xx 微控制器资源。

还移除了不使用模式的文件、结构和常数表，并且修改了一些函数以适应该实现模型。

另一方面，为了充分利用 ARM Cortex™-M3 核（thumb-2 指令集）的优势，下面的函数（需要占用大量 CPU 负载周期的函数）以汇编语言写成：

- filter_mem16
- fir_mem16
- iir_mem16
- inner_prod
- vq_nbest

在本应用笔记适用范围内，STM32F101xx 和 STM32F103xx 上的 Speex 编解码器实现利用两种不同的工具链进行开发和验证：

- 面向 ARM（5.11 版本）的 IAR 嵌入式 Workbench
- Keil μVision3 RealView MDK-ARM（3.22 版本）

IAR EWARM 工具链利用 ARM IAR 汇编器开发汇编代码，Keil RVMDK 工具链则利用 ARM 嵌入式汇编器进行开发。

所有改进和增加的文件，包括汇编代码，在固件包的 **SpeexLib/stm32** 目录中均有提供。

注： 该实现可以很容易地经过调整而适应其他工具链（编译器或汇编），如 GNU。

2.3.3 Speex 编解码器要求

Speex 编码不能在 STM32F101xx 基本型产品上运行，这是由于该系列产品的最高频率为 36 MHz，而该频率下进行编码的 CPU 负载约为 95%。因此目前的固件版本无法在 STM32F101xx 基本型产品上实现编码器。
本应用笔记中，Speex 编解码器的编码在 STM32F103xx 增强型产品上实现。

在 STM32F103xx 上以 72 MHz 运行时，Speex 编解码器所需要的资源情况在表 1 中给出。

表 1. Speex 实现要求

参数	资源需求 ⁽¹⁾
编码器 Flash 大小（以字节计）	31844
编码器 RAM 大小（以字节计）	6456
解码器 Flash 大小（以字节计）	31636
解码器 RAM 大小（以字节计）	3680
编解码器 Flash 大小（以字节计）	31904
编解码器 RAM 大小（以字节计）	7216
编码 CPU 负载（%）	52
解码 CPU 负载（%）	8

1. 这些值是利用 Keil RVMDK 3.22 工具链在 -O3 优化级别且不使用 microlib 的情况下测量的。
这些值来自于仅整合了 Speex 编码器和解码器（而无其他附加处理）的测试演示。Speex 编解码器设置为窄带模式，质量参数设为 4，复杂度等于 1。

2.4 语音合成器固件说明

该固件具有三种应用，说明了语音合成器的典型用途：

- 1. 回放应用：播放编码格式存储的信息。
- 2. 录音和播放：
 - a) 编码和存储一段语音：录音阶段
 - b) 解码和播放存储的信息：播放阶段
- 3. 环回应用：全双工编解码器的基本实现。语音信息被捕获并编码。然后立即实时解码和播放。

这些应用集合在一个演示用例中，该演示在 OLIMEX 的 STM32F103-STK 板上运行。

2.4.1 回放应用

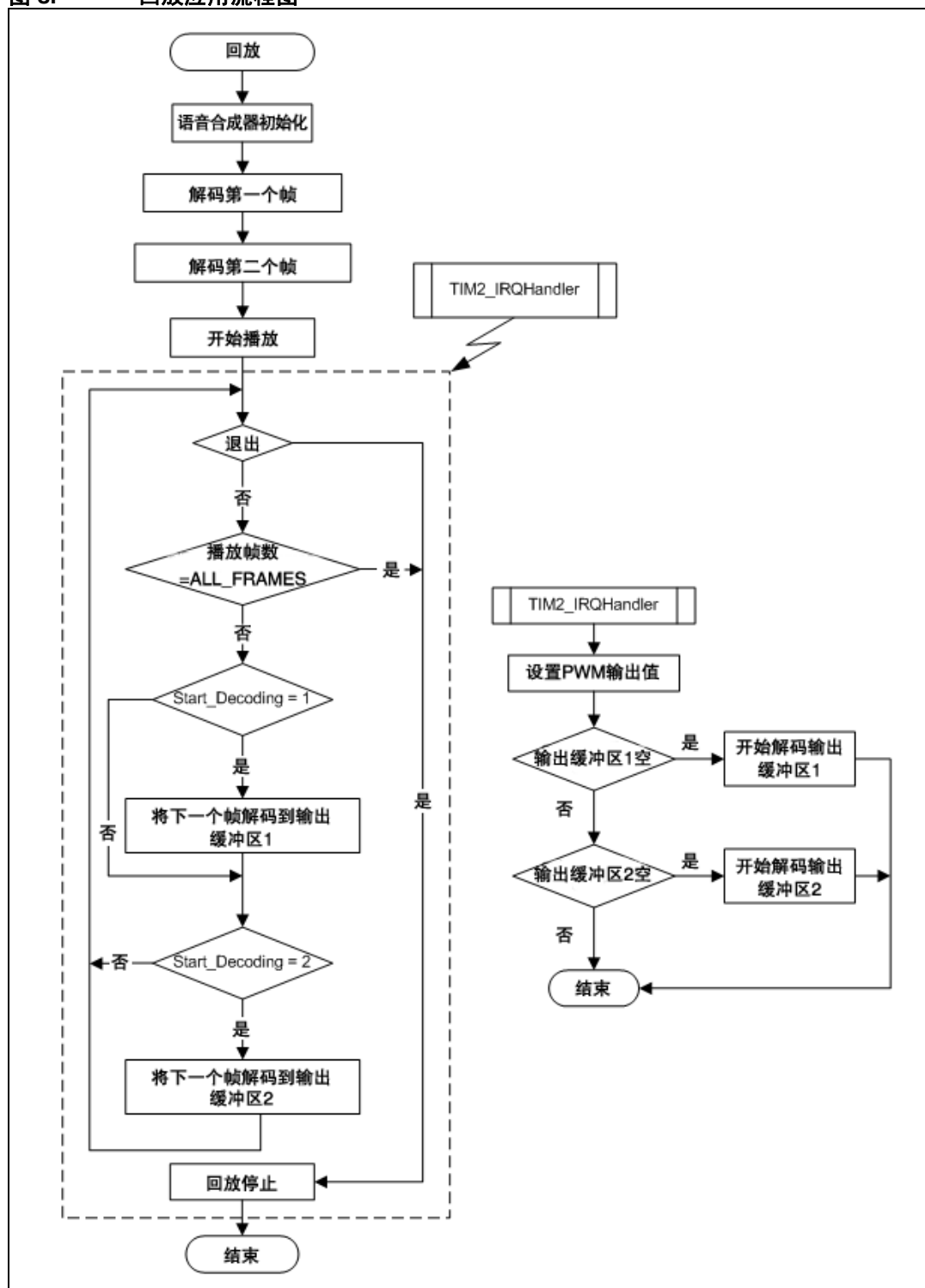
回放应用即解码和播放存储在内置的 STM32F101xx 和 STM32F103xx Flash 中的 Speex 编码信息。

回放应用的主体部分是利用两个缓冲区来解码存储的信息。当一个缓冲区（的信息）已经播放完毕，该应用就重新利用此缓冲区来解码另一帧（信息）。

利用 TIM2 的处理中断函数来播放解码的数据，该函数以 8 kHz 速率被调用。

[图 3: 回放应用流程图](#)详细显示了回放应用的原理。

图 3. 回放应用流程图



2.4.2 录音和播放应用

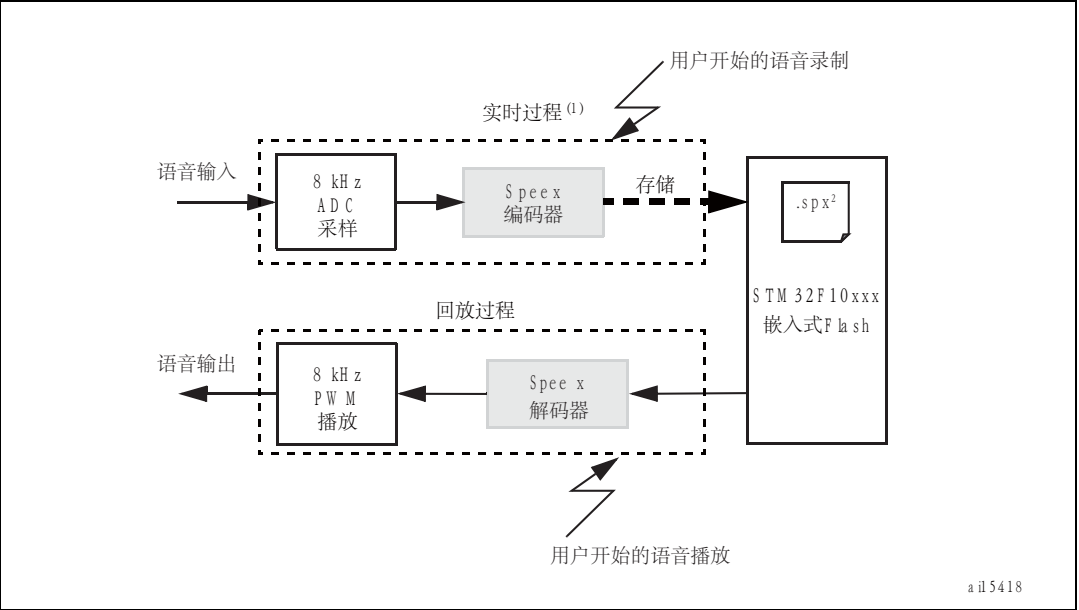
此应用仅用很少的内存来存储语音。用户可将一段语音记录和存储到 Flash（嵌入到 STM32F103xx）中，然后进行重播。

利用该演示，您可以将长达 1 分钟的语音存入内嵌的 Flash 中。为此，固件需要 59 KB 的 Flash 作为数据存储 EEPROM。

作为对比，仅用原始 PCM 格式（不使用 Speex 编解码器）存储 1 分钟的语音时，需要 938 KB 的内存空间。这意味着需要利用外部存储器，也意味着需要额外的成本。

录音和播放应用框图如图 4 所示。录音和播放过程是独立的，因此仅需进行一次录音即可实现多次播放。

图 4. 录音 & 播放框图



- 1. 捕获、编码和存储均是随着发声过程实时进行的。
- 2. .spx 是指通过 Speex 编解码器压缩的语音信息。

录音和播放应用的录音和播放主代码很相似。

录音主代码对数据进行编码并将其存储到内嵌的 Flash 中，而播放主代码则从内嵌的 Flash 中读取并解码数据。

TIM3 中断处理函数（TIM3_IRQHandler）控制输入 / 输出级的语音输入和输出，包含了 ADC 和 PWM 过程。

将 TIM3 设置为每隔 125 μs（8 kHz）产生一次中断。每次调用时，TIM3_IRQHandler 要检查其是被调用于语音记录还是用于语音播放。根据检查结果，TIM3_IRQHandler 选择通过 ADC 进行语音捕获或通过 PWM 进行语音播放。

两个模块（Speex编解码器和语音输入/输出级）之间的交互利用两个不同的缓冲区进行管理。当编解码器在第一个缓冲区工作时，语音 I/O 级则使用第二个缓冲区。当 I/O 达到缓冲区的末端，编解码器和 I/O 级交换缓冲区，编解码器利用第二个缓冲区而 I/O 级使用第一个缓冲区，反之亦然。

图 5、图 6 和图 7 所示的流程图中完整描述了该原理，图中给出了录音和播放的主代码以及 TIM3 中断处理程序。

图 5. 主要的录音应用流程图

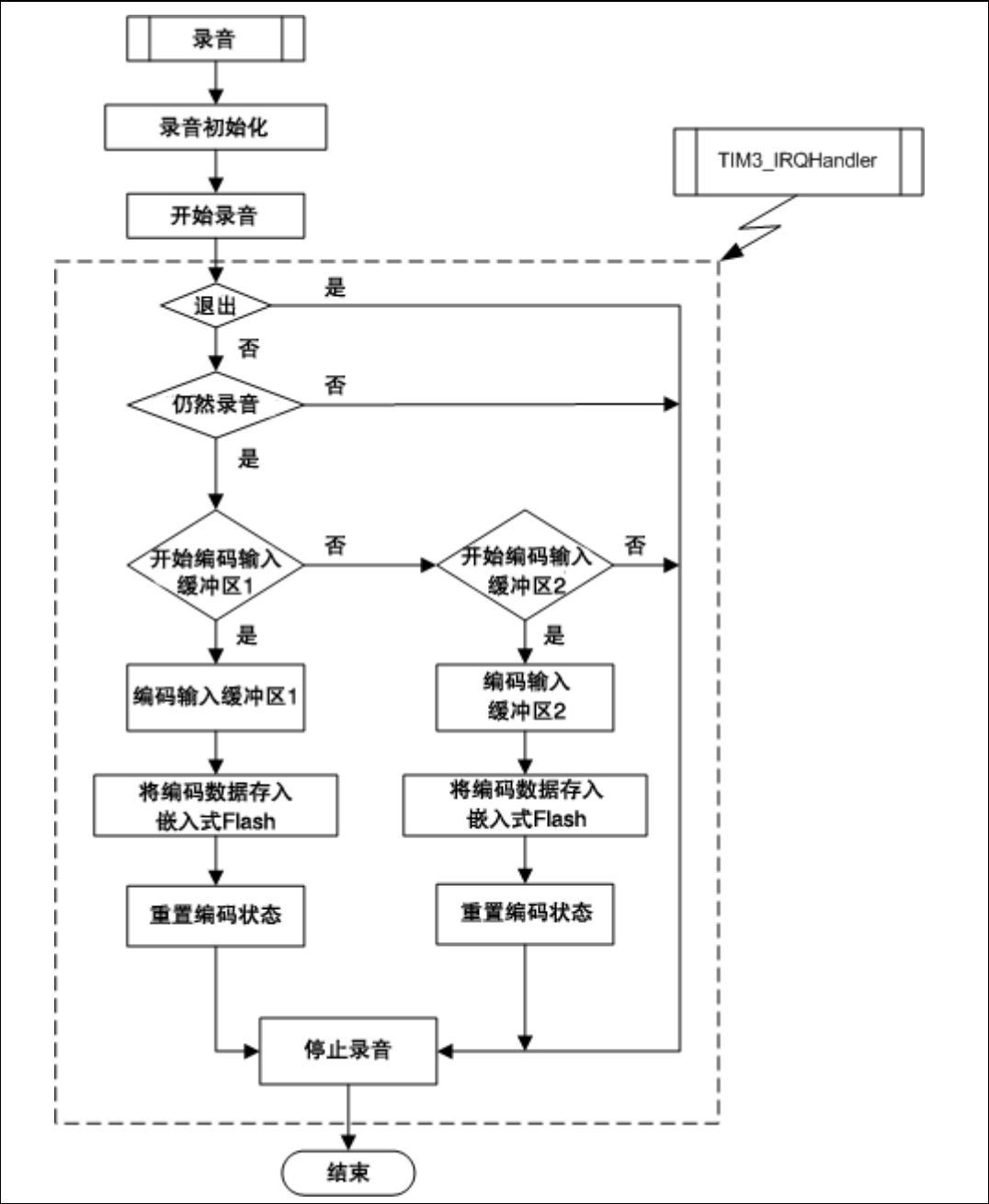


图 6. 主要的播放应用流程图

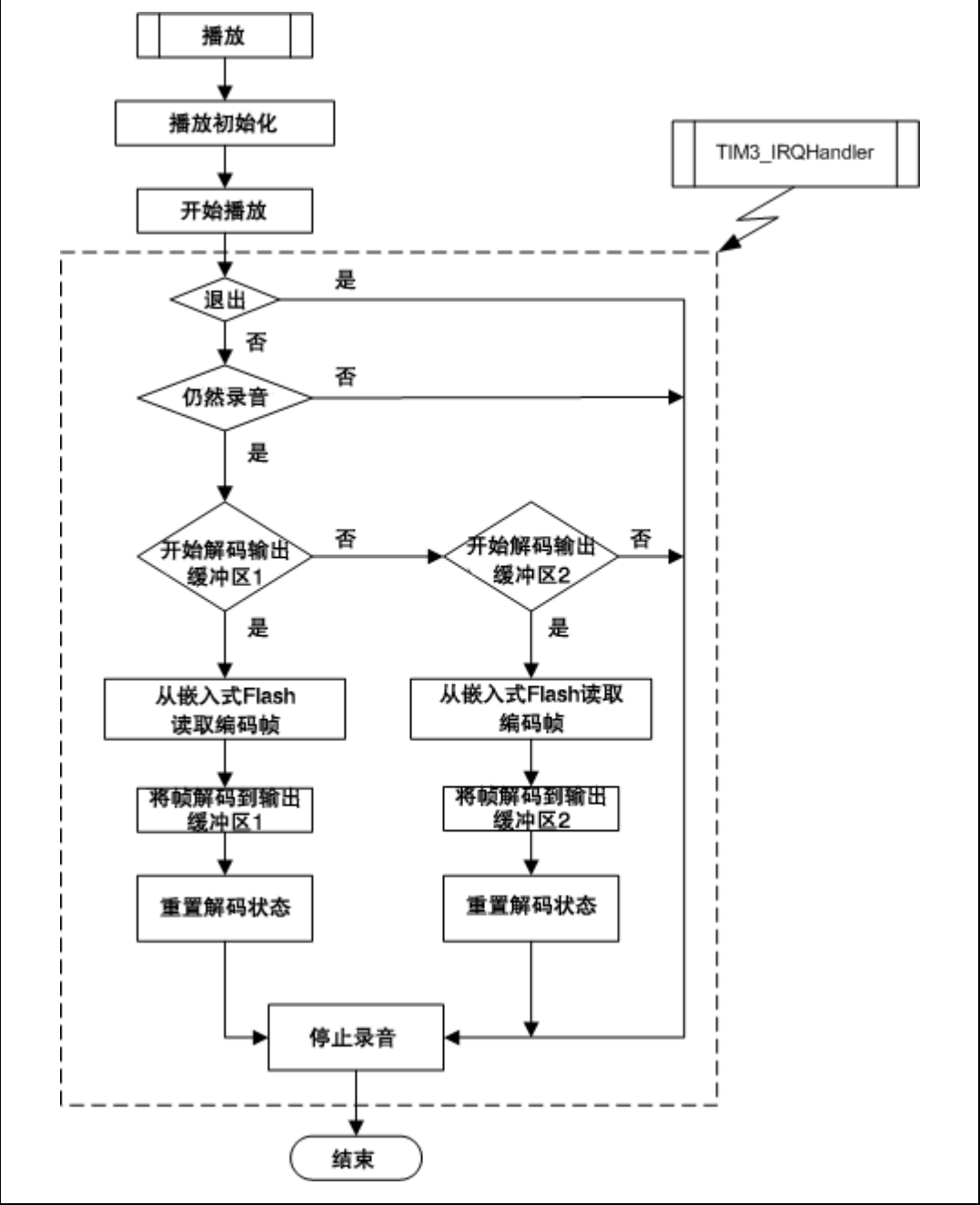
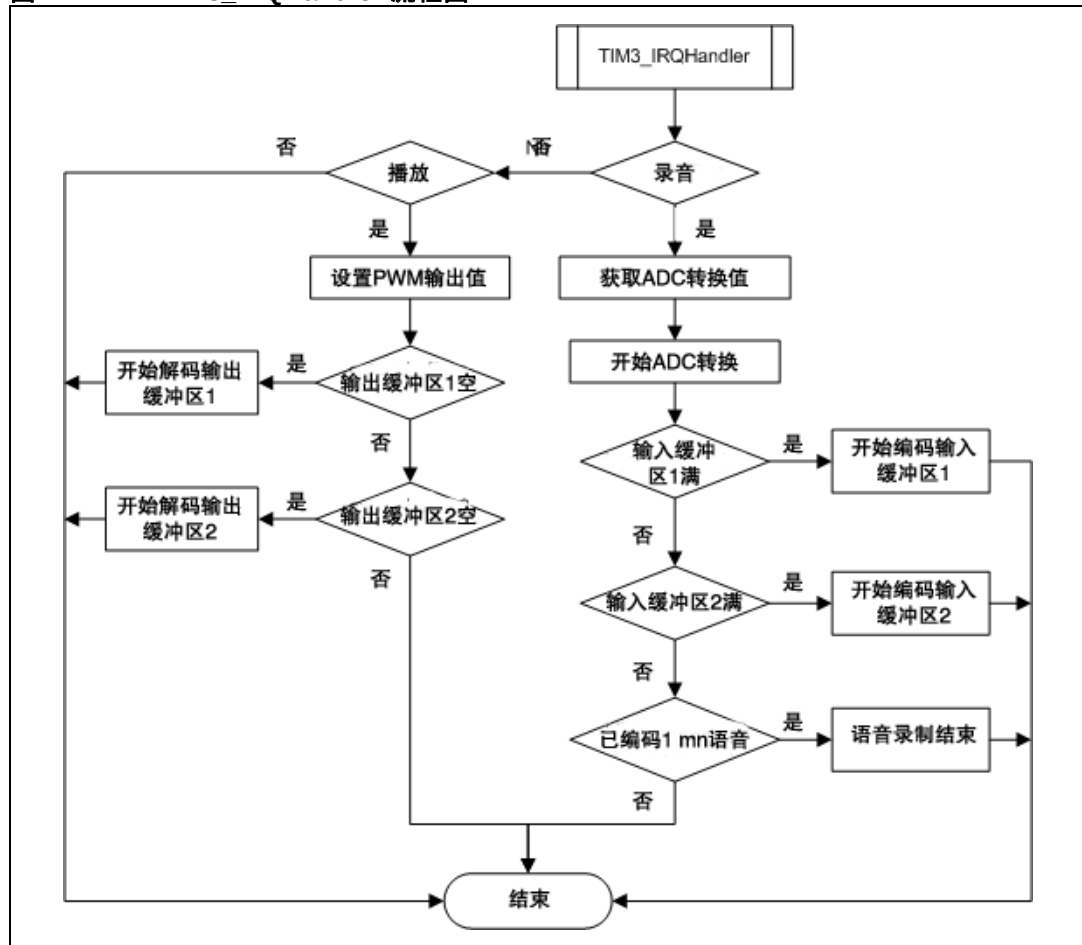


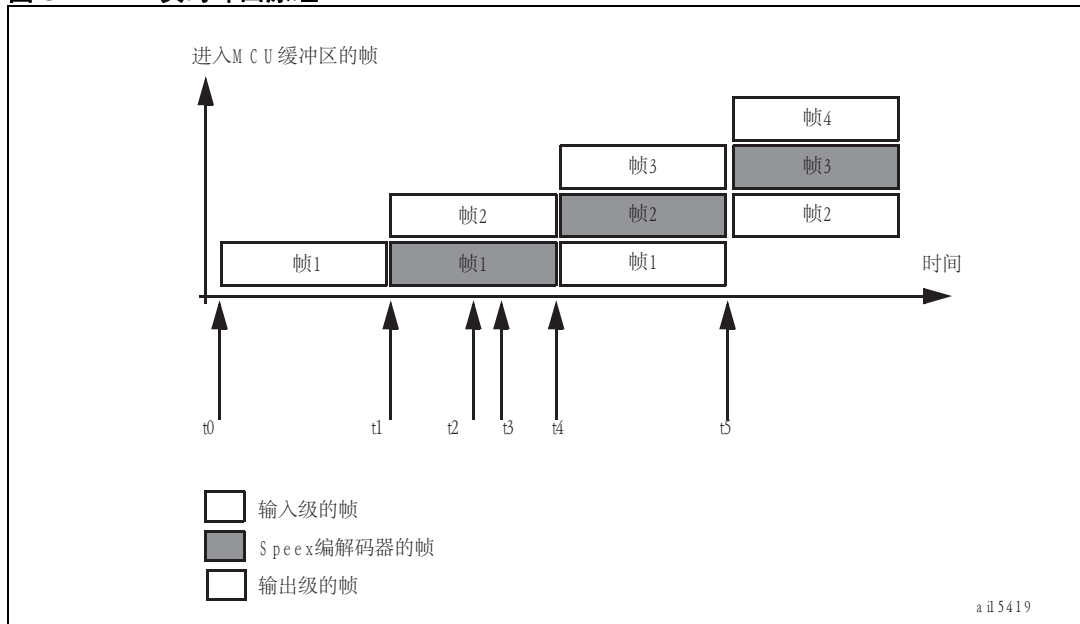
图 7. TIM3_IRQHandler 流程图



2.4.3 环回应用

环回应用说明了如何利用 Speex 编解码器的全双工模式。每一帧语音中，编码和解码过程很难实时进行。[图 8](#) 描述了环回应用中的帧管理原理。

图 8. 实时环回原理



1. t0 时刻:
 - Frame1 开始被记录,
2. t1 时刻:
 - Frame1 在输入缓冲区中准备就绪并开始进行编码,
 - Frame 2 开始被记录。
3. t2 时刻:
 - Frame1 经过了编码并开始进行解码,
4. t3 时刻:
 - Frame1 经过了解码并在输出缓冲区准备就绪,
5. t4 时刻:
 - Frame 1 开始被播放,
 - Frame 2 在输入缓冲区中准备就绪并开始进行编码,
 - Frame 3 开始被记录。
6. t5 时刻:
 - Frame1 已经播放过,
 - Frame 2 经过了编码 / 解码并进行播放,
 - Frame 3 在输入缓冲区中准备就绪并开始进行编码,
 - Frame 4 开始被记录。

如上所述, 在连续操作过程中, 在 RAM 存储缓冲区中总是存在三帧语音。

因此, 我们需要对输入和输出缓冲区进行区分。为了改进实时环回应用, 我们总共需要 4 个 160 采样 (有符号短整型数据) 的缓冲区。

TIM2 中断处理程序改进了输入 / 输出操作，该函数以 16 kHz 速率被调用。实际上，对于每次调用，TIM2_IRQHandler程序要么通过获取ADC转换值进行输入操作，或者通过设置PWM脉冲实现输出操作。这使得输入输出操作均具有 8 kHz 的执行速率。

图9和图10给出了描述主要环回操作和中断环回操作的流程图。

图 9. 环回流程图

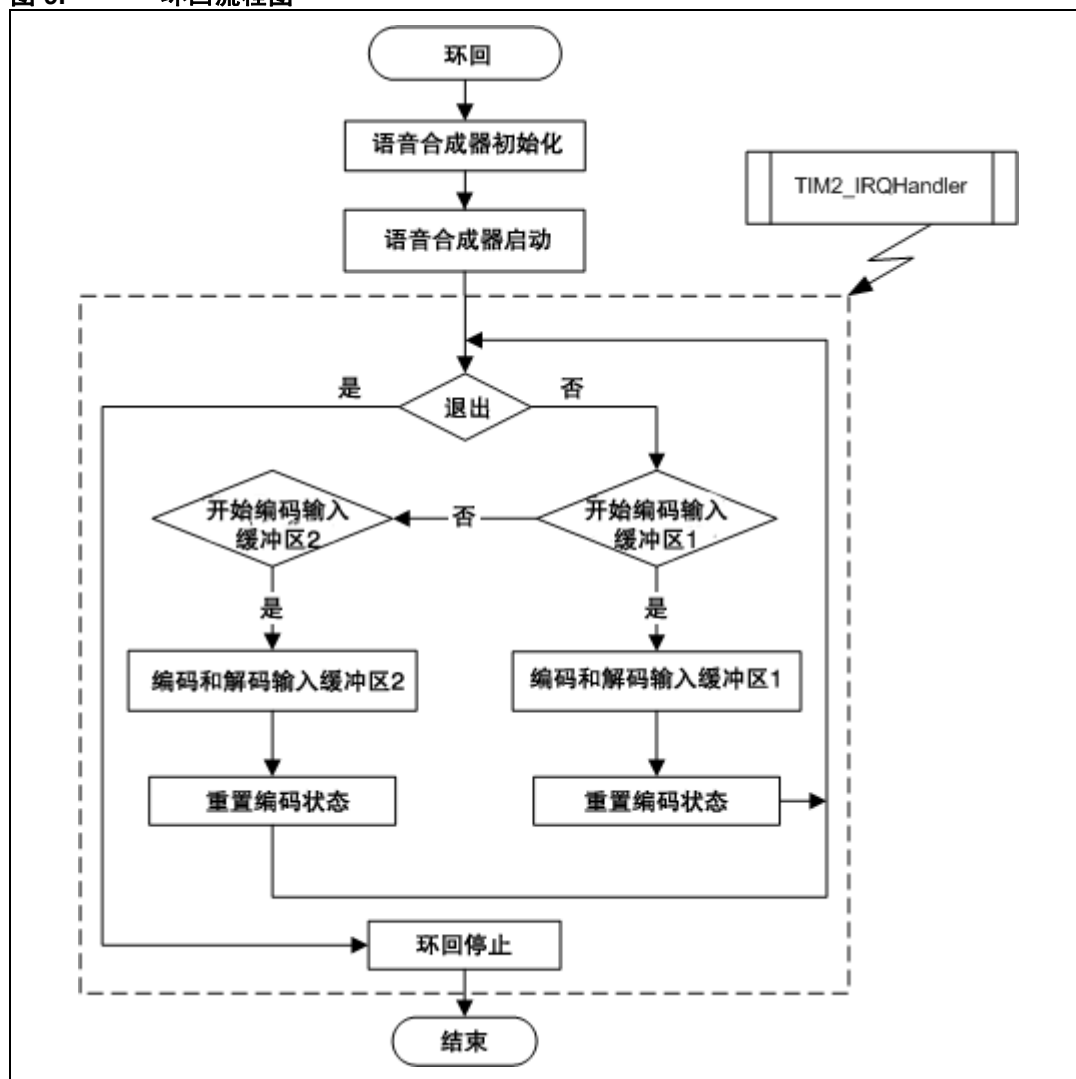
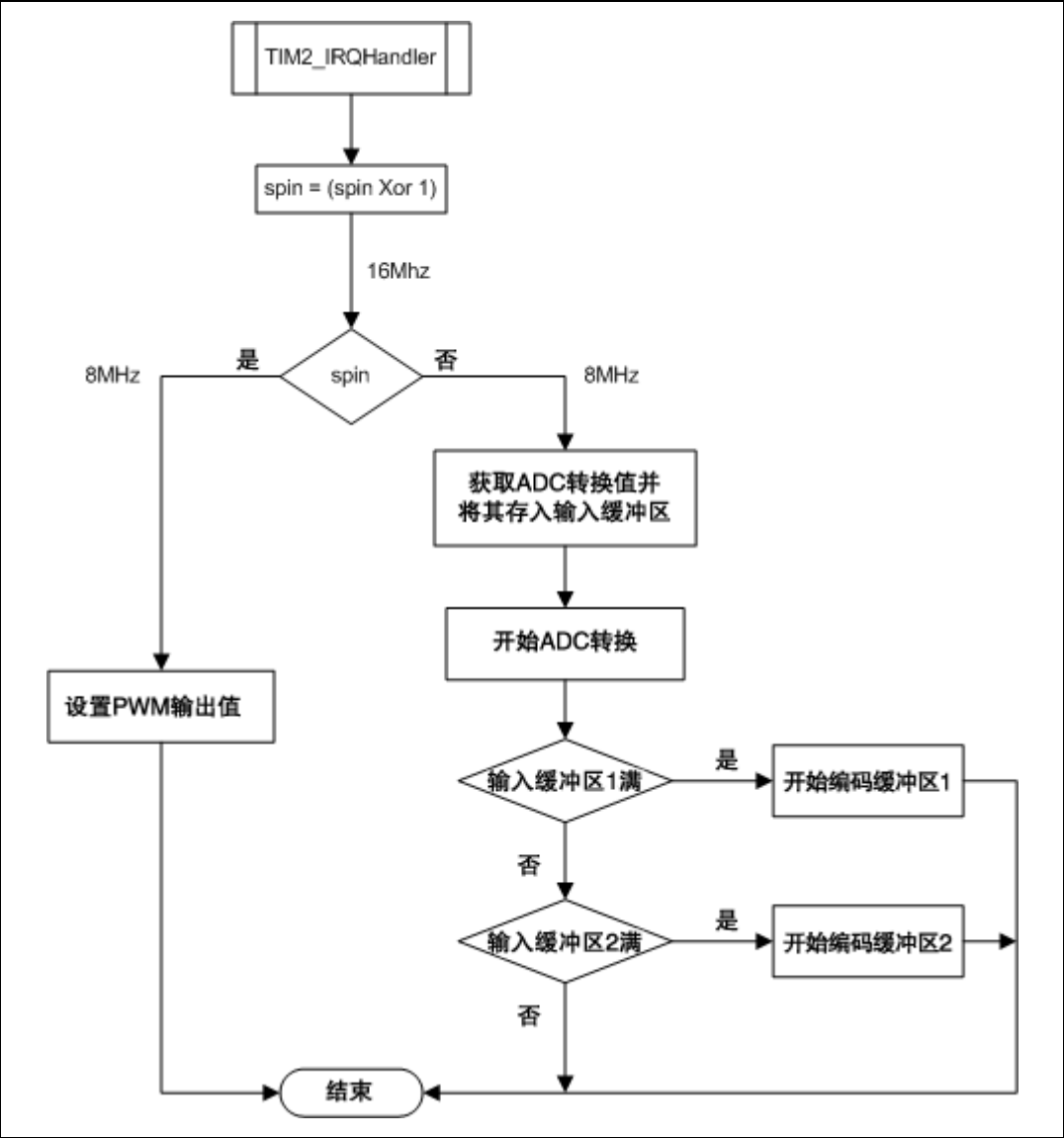


图 10. TIM2_IRQHandler 流程图



3 结论

此应用笔记说明了如何在微控制器的 STM32F101xx 和 STM32F103xx 产品上，利用 Speex 软件编解码器实现语音合成器解决方案：

提供了三种不同的应用：

- 回放应用：播放编码的文件
- 录音和播放应用：在内存中存储信息并重播
- 环回应用：运行实时语音合成器

注：由于目前的固件版本无法在 STM32F101xx 基本型器件上实现编码器，因此这些器件（其频率限制在 36 MHz 内）仅可用于回放应用。STM32F103xx 系列产品则不受此限制。

STM32F101xx 和 STM32F103xx 微控制器产品为创建音频应用提供了其他解决方案。例如，大容量 STM32F101xx 和 STM32F103xx 器件提供了能够用于音频输出的嵌入式 12 位 DAC。它们还提供了 I²S 接口，可以用来连接外部编解码器，而无需额外的硬件。

可以利用这些选择来改进语音合成器和音频应用的功能，使其保持合理成本，并且除控制 MCU 之外，不再需要高端器件或专用 DSP 引擎。

4 修订历史

表 2. 文档修订历史

日期	修订	变更
2008 年 9 月 15 日	1	初始版本。
2008 年 10 月 09 日	2	修改了 表 1: Speex 实现要求 。 删除了“语音合成器演示存储器要求”表。

重要通知 - 请仔细阅读

意法半导体公司及其子公司（“ST”）保留随时对 ST 产品和 / 或本文档进行变更、更正、增强、修改和改进的权利，恕不另行通知。买方在订货之前应获取关于 ST 产品的最新信息。ST 产品的销售依照订单确认时的相关 ST 销售条款。

买方自行负责对 ST 产品的选择和使用，ST 概不承担与应用协助或买方产品设计相关的任何责任。

ST 不对任何知识产权进行任何明示或默示的授权或许可。

转售的 ST 产品如有不同于此处提供的信息的规定，将导致 ST 针对该产品授予的任何保证失效。

ST 和 ST 徽标是 ST 的商标。所有其他产品或服务名称均为其各自所有者的财产。

本文档中的信息取代本文档所有早期版本中提供的信息。

© 2015 STMicroelectronics - 保留所有权利