

# 东芝语音触发软件解决方案

**TOSHIBA**

逻辑LSI 部门  
2018年9月

# 东芝语音触发解决方案

通过语音控制获得快速响应

## 什么是语音触发技术？

这项技术将持续监控某个字的发生（大约10个）并在检测时通知系统。



## 突出特点

### 1 快速响应

反应很快，在0.1秒的范围内。

### 2 高准确度

因为我们在有环境噪声的情况下对字典进行了训练，所以准确度很高。  
(在SNR 0db下达到97.5%(\*))

### 3 小资源

东芝的语音触发器MW需要较小的计算资源。  
(75MIPS, ROM 200KB, RAM 230KB(\*))  
这可以在嵌入式系统上实现。

(\* 这些值来自于东芝的评估。可能因环境而变化。)

## 用例

### 用于游戏

- 当玩家想要施展魔法时，他们会呼叫“魔法咒语”。
- 当玩家忙着用手控制时，他/她将通过声音控制角色。

你必须服从！  
“火球！”



### 用于智能手机

- 用户通过语音拍照
- 用户通过语音而不是手指来控制某些设置

让我们笑一下！  
“说，茄子！”



# 东芝语音触发和传统语音识别

	东芝语音触发器	传统语音识别
响应	<b>快速</b> ( 0.1秒内) 因为按音位进行比较, 所以不需要等待输入短语结束。	<b>不快</b> (约有1秒延迟) 因为需要确认输入短语结束, 并逐字比较所有候选词。
噪声检测比	<b>高</b> (在SNR 0dB下达到97.5% ) 因为不需要确认输入短语结束。	<b>不高</b> 因为可能无法确认输入短语是否结束, 从而导致检测错误。
检测启动信号	<b>不要求</b> 音位比较始终有效。	<b>要求</b> 需要一个明确的信号来启动检测。
需要硬件资源	<b>较少</b> 例如: 75MIPS, ROM200KB, RAM230KB on ARM 9	<b>较多</b>

1. 快速响应

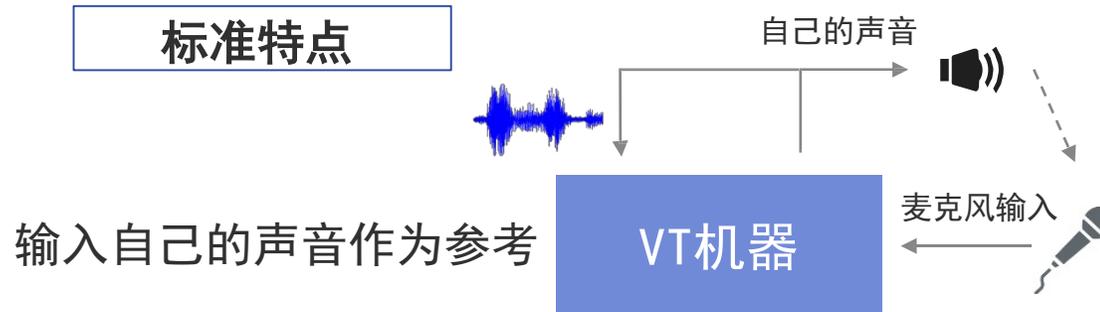
2. 高检测率

3. 小资源

# 独特的高检测率技术

## 改进因扬声器输出引起的错误接收

为了避免检测到自己的声音，可以比较麦克风输入信号和参考信号。



## 相似词的区分

使用原始算法标记关键字列表中的相似词以进行区分。

**标准特点**

举例. 相似关键词

Turn **R**ight

Turn **L**ight on

## 词适应

提高特定关键字的检测率。

**可选\***

E n g l i s h

English

对于常规词语，  
按音位学习

对于特殊关键词，  
学习词语或短语

\*需要额外费用

# 语音触发器应用举例

## ■ (举例1) 启动应用触发器

- “S” 启动XXX”
- 提供词适应 (\*选项)

“启动XXX”

## ■ (举例2) 角色扮演

- 我是一个巫师。
- “发射火球!”

“火球!”



## ■ (举例4) 关键字自定义

- “Abrakadabra”
- 创建自己的魔法咒语

“ Abrakadabra”

## ■ (举例3) 辅助命令

- “治疗药水!”
- 当你的手忙得不可开交时，语音触发器可以作为一个宏命令帮助你。

“治疗药水!”



## ■ (举例5) 照相

- “说，茄子!”
- 控制功能的语音命令。

“说，茄子!”



**TOSHIBA**

**谢谢!**

