

# 用 C# 开发基于 Microsoft Speech SDK 的语音应用程序

熊 凯

(山东莱阳农学院信息科学与工程学院, 山东 青岛 266109)

**摘 要:** 对 Microsoft Speech SDK 5.1 语音控制编程接口 SAPI 进行了分析, 详细描述用 C# 语言结合这些接口开发语音应用程序的流程步骤, 并最终通过实例程序展示了共享型语音控制程序的开发流程。

**关键词:** Microsoft Speech SDK; 语音合成; 语音识别; C#; COM

## 0 引言

一直以来, 让计算机听懂人们说的话, 从而让人与计算机用自然语言方便地进行交流, 是语音识别技术追求的目标。

本文在分析 Microsoft Speech SDK 应用程序开发接口的基础上, 阐述应用 C# 如何开发语音应用程序。

## 1 Microsoft Speech SDK

Microsoft Speech SDK 5.1 是微软提供的软件开发包, 其中包含了语音识别和合成引擎相关组件、帮助文档和例程, 它是一个语音识别和合成的二次开发平台。我们可以利用这个平台, 在自己开发的软件里嵌入语音识别和合成功能, 从而使用户可以用声音来代替鼠标和键盘完成部分操作, 例如: 文字输入、菜单控制等, 实现真正的“人机对话”。

## 2 .NET 与 COM 组件

SDK 中提供的组件采用 COM(Component Object Model)标准开发, 底层协议都以 COM 组件的形式完全独立于应用程序层, 为应用程序设计人员屏蔽掉复杂的语音技术, 语音相关的一系列工作由 COM 组件完成, 程序员只需专注于自己的应用, 调用相关的语音应用程序接口(SAPI)就可实现语音功能。COM 组件为在 Windows 平台上开发跨语言程序提供了可能, 但传统 COM 编程的一系列技术较为复杂, 对程序员水平的要求较高。Microsoft 提供的 .NET 平台对此作了很大改进, 最终大大简化了 .NET - COM 的互操作过程。例如, 利用 Visual Studio .NET 开发平台引用 SDK 提供的 COM 组件非常简单, 在菜单选择中 工程 | 添加引用, 然后点击 COM 标签, 选择 Microsoft Speech Object Library, 即可完成对该组件的引用。

实际上, 这一做法, 意味着项目引用了下面路径中的程序集: C:\Program Files\Common Files\Microsoft Shared\Speech\sapidll (具体路径与操作系统相关)。

引用 COM 组件后, 在 C# 程序中就可以像使用 .NET 自身的组件一样使用 COM 组件。

## 3 Microsoft Speech API 5(SAPI 5)

Microsoft Speech SDK 提供的 Speech API(SAPI)主要包含 API for Text-to-Speech (TTS) 和 API for Speech Recognition(SR)。其中 API for Text-to-Speech 是微软 TTS 引擎的接口, 通过它可以很容易地建立功能强大的文本语音程序。API for Speech Recognition 是与 TTS 相对应的语音识别

引擎接口, 通过这些接口能够获取音频流并给出相应的识别结果。应用程序对底层语音数据的控制处理正是通过这些接口实现的。SAPI 框架结构如图 1 所示。

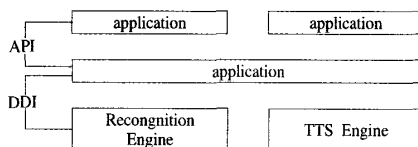


图 1 SAPI 框架结构

### 3.1 语音合成 API

实现语音合成功能, 最主要的是使用 ISpVoice, 该接口的主要功能就是实现文本到语音的转换。在接口中, 有许多成员函数, 通过这些成员函数可以实现对发音频率、发音音量等合成属性进行调整。在 C# 中, 我们通过实现 ISpVoice 接口的 SpVoice 类来实现语音合成。具体步骤如下:

(1) 首先创建 SpVoice 类的对象;

(2) 利用对象成员函数设置语音属性, 如果不设置, 则按默认方式处理;

(3) 调用对象成员函数 Speak 函数来朗读指定文本。Speak 函数需要两个参数, 第一个参数指明需要朗读的文本, 第二个参数指明发音方式。

关于 SpVoice 类及其成员函数的介绍可以参考 SDK 开发文档, 本文不再赘述。

### 3.2 语音识别 API

语音识别涉及的接口较多, 也较复杂, 本文介绍几个较为重要的接口。

**ISpRecognizer 接口。**该接口负责与底层的 Recognition engine 打交道, 它扮演语音识别引擎代言人的角色。有两种语音识别器可供选择: 共享式和独占式。SpSharedRecognizer 类实现了共享式识别器接口, 独占式识别器接口由 SpInProcRecognizer 类实现。

**ISpRecoContext 接口。**此接口是完成语音识别的主要接口, 它主要用于发送和接收与语音识别有关的消息通知、创建语法规则对象。有两个类实现了该接口, SpSharedRecoContext 类和 SpInProcRecoContext 类, 但在创建它们的实例时, 前者会同时创建一个相关联的 SpSharedRecognizer, 后者创建一个 SpInProcRecognizer 与之关联。

ISpRecoGrammar 接口。这是语法器接口,用于创建、载入和激活语音识别时用的语法规则。语法规则也有两种类型:一种是听写(Dictation)型,允许使用引擎中的大量短语,但识别效率较低;另一种是控制命令(Command and Control)型,能最大限度识别自定义的短语,故识别效率高。

在 C# 中,实现语音识别的具体步骤如下:

(1) 首先创建上下文类 SpRecoContext 的实例对象。应根据实际需要,决定是创建共享型上下文 SpSharedRecoContext 类还是独占型 SpInprocRecoContext 类对象。由于在创建 SpSharedRecoContext 类和 SpInprocRecoContext 类时会自动生成相关联的语音识别引擎,因此,程序中不需要显式地创建识别器。

(2) 为上下文对象增加相关事件处理机制。上下文对象能处理多个事件,其中最常用的是 Recognition Event。当语音识别引擎识别出语音内容后触发该事件,并通过 ISpeechRecoResult 接口返回识别的结果。

(3) 为上下文对象创建语法器。SpRecoContext 类的 CreateGrammar 函数可以创建并返回一个语法器。

(4) 为语法器装载语法规则。根据实际需要,如果是听写型语法器,用 ISpRecoGrammar::DictationLoad 函数装载,如果是控制命令型,则应该用类似于 ISpRecoGrammar::cmdLoadFromxxx 的函数装载。

(5) 激活语法器。对于听写型应该用 ISpRecoGrammar::DictationSetState, 对于控制命令型, 用 ISpRecoGrammar::SetRuleState 或 ISpRecoGrammar::SetRuleIdState。

经过以上步骤,即可以建立一套语音识别机制。上述过程中涉及到的函数具体用法可以参考 SDK。

#### 4 应用实例

下面通过完整实例介绍用 C# 语言开发共享型语音控制程序。此实例程序的主要功能是:用户在文本框内输入文本,程序朗读指定文本;用户选择“听写”后,程序根据用户的语音输入,输出识别的结果。

##### (1) 工程准备

如果没有安装 Microsoft Speech SDK 5.1, 应先安装。在 Visual Studio C# .NET 平台下建立一个 Windows 应用程序工程项目,并按上文所述方式为工程添加 Microsoft Speech Object Library 引用。

##### (2) 界面设计

如图 2 所示,在窗体上添加一个多行文本框,取名为 textbox;两个按钮,分别取名 speakbtn 和 dictbtn。

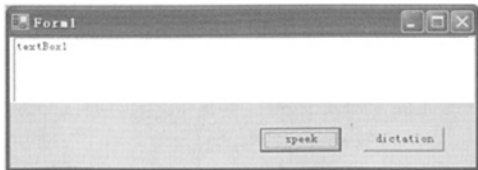


图 2 实例程序窗体

##### (3) 添加代码

转到代码编辑页面。

添加 using 指令:using SpeechLib;

为类添加以下成员声明:

```
private SpVoice voice; // 语音合成对象
private SpSharedRecoContext objRecoContext;
// 共享型上下文对象
private ISpeechRecoGrammar grammar; // 上下文语法对象

添加成员函数:

public void RecoContext_Recognition(int StreamNumber, object
StreamPosition, SpeechRecognitionType RecognitionType,
ISpeechRecoResult Result)
{ // Recognition 事件的响应处理
    textbox.Text+=Result.PhraseInfo.GetText(0, -1, true)+" ";
    // 输出语音识别结果
}

private void InitializeSpeech()
{ // 语音识别初始化
    objRecoContext=new SpSharedRecoContext (); // 初始化共享型上下文对象
    objRecoContext.Recognition += new
    _ISpeechRecoContextEvents_RecognitionEventHandler(
    RecoContext_Recognition); // 为上下文添加事件处理函数
    grammar=objRecoContext.CreateGrammar(0); // 创建语法器
    grammar.DictationLoad("",SpeechLoadOption.SLOStatic);
    // 装载语法规则
    grammar.DictationSetState(SpeechLib.SpeechRuleState.
    SGDSActive); // 激活语法器
    objRecoContext.State=SpeechRecoContextState.
    SRCS_Disabled; // 停止语音识别
}

private void Form1_Load(object sender, System.EventArgs e)
{ // 窗体装载
    voice=new SpVoiceClass(); // 初始化语音合成对象
    this.InitializeSpeech(); // 初始化语音识别
}

private void speakbtn_Click(object sender, System.EventArgs e)
{ // 用户要求程序朗读指定文本
    // 先关闭上下文,即停止语音识别
    objRecoContext.State=SpeechRecoContextState.
    SRCS_Disabled;
    // 语音合成器朗读文本框文本
    voice.Speak(textbox.Text ,SpeechVoiceSpeakFlags.
    SVSFlagsAsync );
}

private void dictbtn_Click(object sender, System.EventArgs e)
{ // 用户要求进行语音识别
    objRecoContext.State=SpeechRecoContextState.
    SRCS_Enabled; // 激活上下文
}
```

##### (4) 编译运行程序

程序运行时,在文本框中输入内容后,按“speak”按钮,则程

# 基于 Dialogic 卡的办公自动化系统研究

童亚拉

(湖北工业大学理学院, 湖北 武汉 430068)

**摘要:** 从系统的概述、结构、工作层次、软硬件支持、功能、程序设计事项方面对基于 Dialogic 电话卡、具有语音合成与语音识别功能的办公自动化系统进行了详细描述,并给出了测试结果与尚待解决的问题。

**关键词:** 办公自动化系统; 语音合成; 语音识别; Dialogic 电话语音卡

## 0 引言

传统的计算机通过键盘和鼠标输入,通过显示器和打印机输出。这些设备操作不方便,学习成本较高,需要我们设计出操作方便、实时性与可靠性较强、功能齐全、成本低廉的办公自动化系统。语音技术正迅速成为人与计算机间信息交流的有效工具,该技术的发展将为实现办公自动化提供更优越的条件。目前语音识别和文语转换技术已从实验室逐步走向成熟,并在市场中得以应用。本文正是根据所开发的这样一套办公自动化系统而撰写的。

## 1 系统概述

带语音功能的办公自动化系统通过电话语音卡将计算机和公用电话网紧密结合起来,充分利用电话方便快捷和计算机海量存储的特点,在语音提示下进行交互,从而提供信息查询、事务处理等服务<sup>[1]</sup>。系统的信息查询业务是指拨打一个电话号码后通过 IVR(Interactive Voice Response)获取存放在计算机数据库的各种信息和服务。如,学校查分系统,学生或学生家长拨通呼叫中心电话后,在语音的提示下手工输入或说出如学号或预先分配的序号,即可查询该生的在校成绩、在校表现等,同时呼叫中心用普通话报出相关信息。又如主动呼出服务,即呼叫中心自动定时外拨某个用户电话通知有关事项,如基层单位任务下达、语音转发、会议通知、需商议重要事情的预约通知、进行催缴及代收项目通知等。对方的语音回复将被自动记录备查,保证通知的有效性和及时性,使通知方不必为对方接不通、不在家而烦恼,也使通知完成情况有据可查。

## 2 系统结构

如图 1 所示,系统由电话语音卡、计算机、办公自动化系统软件、ASR 和 TTS 引擎等构成。只要将语音卡放置于计算机的扩展槽中,并把接口与申请该项业务的电话机相连接,再运行此系统软件就可实现办公自动化系统中的语音功能。可根据需要设置 4 路、8 路、12 路或更多路的并行呼入,提高线路的利用率。

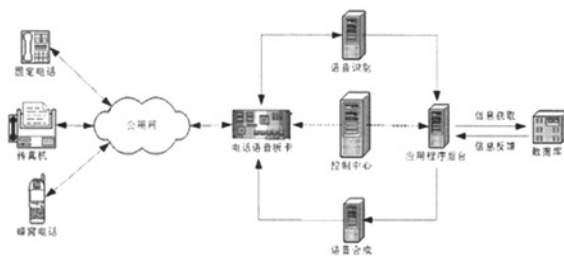


图 1 系统结构示意图

## 3 工作层次

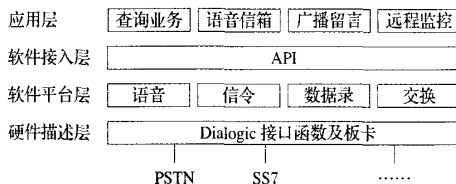


图 2 工作层次图

系统的工作层次如图 2 所示。应用层处理查询业务、语音信箱、

序朗读文本框内容。按“dictation”按钮时,用户通过话筒说话,程序在文本框内追加识别结果。

为了突出说明技术要点,实例程序只给出语音合成及识别的必要代码,没有描述异常处理和代码优化等细节。

## 5 结束语

Microsoft Speech SDK 5.1 软件开发包提供了很好的语音识别和语音合成二次开发平台,.NET 平台为操纵 COM 组件屏蔽了复杂的技术细节,C# 语言与.NET 平台与生俱来的良好结合,这些先进技术大大简化了语音控制应用程序的开发,促进了语音控制技术的应用与推广。本文阐述的 SAPI 语音控制接

口的工作原理和 C# 语音应用程序开发流程,对开发语音应用程序能够起到抛砖引玉的作用,具有一定参考价值。

以上程序示例在 Windows XP SP2,Microsoft Speech SDK 5.1,Visual Studio C#.NET Version 7.1 环境下调试通过。

## 参考文献:

- [1] Microsoft Inc.Microsoft speech SDK help[EB/OL]. <http://www.microsoft.com/downloads>,2001.8.8.
- [2] 李英才,左友东,郑青清等.基于 Speech SDK 的语音控制应用程序的设计与实现[J].计算机应用,2004.24(6):114~117
- [3] 高敬忠,姜子敬,胡金铭.基于 Speech SDK 的语音应用程序实现[J].广西科学院学报,2005.21(3):169~172



# 用C#开发基于Microsoft Speech SDK的语音应用程序

作者: [熊凯](#)  
作者单位: [山东莱阳农学院信息科学与工程学院, 山东, 青岛, 266109](#)  
刊名: [计算机时代](#)  
英文刊名: [COMPUTER ERA](#)  
年, 卷(期): 2007, "" (2)  
被引用次数: 0次

## 参考文献(3条)

1. [Microsoft Inc](#) [Microsoft speech SDK help](#) 2001
2. [李禹才](#), [左友东](#), [郑秀清](#) [基于Speech SDK的语音控制应用程序的设计与实现](#) [期刊论文] - [计算机应用](#) 2004 (06)
3. [高敬惠](#), [姜子敬](#), [胡金铭](#) [基于Speech SDK的语音应用程序实现](#) [期刊论文] - [广西科学院学报](#) 2005 (03)

## 相似文献(10条)

1. 学位论文 [赵小刚](#) [基于VoiceXML和Microsoft Speech SDK的呼叫中心的设计及实现](#) 2005

本文对基于VoiceXML和Microsoft Speech SDK的呼叫中心的设计及实现进行了研究。文章在分析了传统的呼叫中心发展趋势之后,对目前流行的VoiceXML进行了仔细研究,认为利用VoiceXML技术不但可以克服传统呼叫中心的缺陷,而且可以简化开发人员的工作。文章详细介绍了VoiceXML的体系结构,结构模型,工作原理,支撑平台;系统介绍了基于VoiceXML的语音浏览器的工作流程,开发步骤,交互过程;集中介绍了VoiceXML的基本语法内容:包括窗体、事件、FIA(窗体解析算法)、语音识别语法等。文章同时结合发票语音查询系统项目,分析了项目的背景、需求。根据在项目开发中遭遇的实际开发困难,提出了用VoiceXML和MicrosoftSpeech来实现整个呼叫中心的方法,并清楚的定义了各部分功能的分配。

2. 期刊论文 [陈永煊](#), [鲍鸿](#), [张晶](#), [CHEN Yongxuan](#), [BAO Hong](#), [ZHANG Jing](#) [基于SPEECH SDK的中文学习系统](#) - [电脑编程技巧与维护](#) 2010, "" (2)

介绍了Microsoft Speech SDK的功能以及相关的一系列API函数,分别利用其语音识别(Speech Recognition)引擎和语音合成(Text to Speech)引擎实现英文语音识别和中文文本语音合成功能,采用Access 2003数据库的查询操作完成翻译功能,进而实现整个英文到中文的学习系统,该系统具有很高的商业价值。

3. 期刊论文 [尹惠玲](#), [杨帆](#), [于虹](#), [吴涛](#), [YIN Hui-ling](#), [YANG Fan](#), [YU Hong](#), [WU Tao](#) [基于COM的智能TTS系统的设计与实现](#) - [微计算机信息](#) 2009, 25 (15)

本文利用Microsoft Speech SDK 5.1语音开发包提供的语音合成引擎COM组件,在Matlab环境下开发了一个能够朗读中英文混排文档的智能TTS系统,极大地缩短了语音合成应用程序的开发周期。实验证明该系统实用可靠,有很大的应用价值,可广泛应用于盲人计算机、即时信息服务、语音报警提示等领域。

4. 学位论文 [宋冰](#) [语音技术在英语学习软件中的应用](#) 2009

随着社会的发展和对外交往的日益增加,英语逐渐成为必须掌握的技能。而传统的英语教学,忽视对“听”和“说”的训练。本课题研究通过实用语音技术,开发一款具有语音识别和合成的英语学习软件iEnglish。<br>

语音技术是近半个世纪发展起来的新型科学,包括语音识别和语音合成。语音识别目的是让机器把语音信号转化可以接受的文本文件或者控制命令。语音合成则是将文本转化为人类可以理解的声音信号。语音技术在工业、军事、交通、医学、民用等各方面有广阔的应用前景,蕴含着巨大的社会、经济效益,受到各国日益重视。<br>

本文是主要阐述通过利用Microsoft Speech Sdk和Microsoft Agent,在英语学习软件实现语音的识别和合成。利用Speech Sdk的Speech Recognition API和Agent的语音识别方法,可以用语音命令代替键盘进行操作,进行人机交互。同时也可以通过语音识别来进行语音留言,更加方便用户的使用。利用Speech Sdk的TTS API和Agent的TTS技术,可以实现语音的合成,为英语听力训练的实现提供了可能。<br>

课题重点使用Agent的TTS技术。利用Speak方法,可以实现文本到语音的转换。通过设置Agent的语言属性和语音引擎,可以进行多语言的发音。通过设置Agent的气泡属性和发音控制符号,给用户提供的多样的、适合的、可自主选择的听力训练模块。<br>

iEnglish不但具有语音的识别和合成,而且还具有词典、网络查词、屏幕取词、生词本等多种实用功能。用户可以维护和修改词库,建立自己需要的生词本,更加便于用户的使用。<br>

最后,对课题主要完成的工作进行了总结和展望。课题的研究对于将语音技术应用于英语学习软件具有重要的实用价值。

5. 期刊论文 [杨世强](#), [梁丁宏](#), [傅卫平](#), [YANG Shi-qiang](#), [LIANG Ding-hong](#), [FU Wei-ping](#) [智能机器人语音远程控制系统的的设计](#) - [计算机工程与应用](#) 2009, 45 (25)

为使智能机器人远程控制更加方便、快捷、人性化,设计并实现了一种智能机器人的语音远程控制方案。该方案利用微软语音开发包Microsoft Speech SDK,构建基于听写模式的大词汇量语音识别模块和语音合成模块,利用海量中文智能分词组件构建关键词检测模块,结合VFW(Video For Windows)技术与无线网络技术构建信息传输模块。实验表明,该系统语音识别准确率高,识别范围广,语音输入灵活。

6. 期刊论文 [童强](#), [TONG Qiang](#) [用TIS技术开发具有文本朗读功能的的应用程序](#) - [现代计算机 \(专业版\)](#) 2003, "" (12)

TTS技术已经越来越广泛地应用到现代计算机的各个领域,本文介绍了利用微软的Microsoft Speech SDK 5.1提供的TTS(text-to-speech)技术开发具有语音朗读功能的应用程序。

7. 学位论文 [许芹](#) [视觉语音合成技术在英语发音辅导中的应用探究](#) 2007

随着全球一体化进程的迅速推进,我国与世界各地之间的交流日益频繁,英语作为国际通用的工作语言越来越受到人们的重视。但是,由于多年只重视书面教学,和缺乏良好的口语学习环境,致使我国当前的英语口语教学收效甚微。虽然计算机技术在我国发展的如火如荼,但是我国的计算机辅助语言学习(CALL, Computer-Assisted Language Learning)却仍然停留在起步阶段。

针对这一现状,笔者将视觉语音(Visual Speech)技术应用于英语初学者的语音教学。本文参考美国大力推广的Phonics教学法,开发了一个唇形一语音同步的英语发音辅导系统,希望从以下两方面帮助英语初学者学习语音:一是根据语音的双模态特性,视觉语音可以帮助用户更好的观察、模仿唇



部发音动作，有助于用户理解、记忆语音；二是借助于视觉语音技术呈现的用户界面更加友好，人机交互更加和谐、自然，这样对于缓解英语初学者的压力，提高学习者的学习积极性有很大帮助。

本文所做的工作主要有以下几点：

- 基础标准层面，对MPEG-4定义的“人脸对象”进行介绍并以该定义中人脸动画的参数(FAP)为基础开展后面的工作；
- 技术要素层面，对本文采用的Microsoft Speech SDK 5.1中的TTS引擎进行研究和实践；
- 系统架构层面，对本文提出的视觉语音合成系统(TTVS)的框架结构、进行介绍和分析；
- 具体算法层面，详细介绍实现视觉语音动画合成系统的步骤和算法等；
- 系统应用层面，将详细介绍“EP Tutor”系统的知识结构、各模块功能及其应用场景；
- 工作展望层面，将对EP Tutor系统进一步的发展做出展望。

8. 期刊论文 [秦铁, QIN Tie 利用微软用TTS引擎实现语音报警系统 -黑龙江气象2009, 26 \(2\)](#)

1 微软的TTS语音引擎概述

Microsoft Speech SDK提供了一套关于语音处理的应用程序编程接口SAPI (Speech Application Programming Interface). SAPI实现了文字到语音的转换和语音识别的方法. 语言引擎通过DDI层 (设备驱动接口) 和SAPI进行交互, 应用程序通过API层与SAPI进行通信. 通过使用这些API, 用户可以快速开发在语音识别和语音合成上面的应用程序. 其结构功能见图1. SAPI最基本的语音引擎就是TTS (Text To Speech), 也就是文本到语音的转换, TTS通过语音合成来实现文本字符串或文本文件的朗读.

9. 期刊论文 [陈景帅, 周凤余, CHEN Jing-shuai, ZHOU Feng-yu 基于Speech SDK的机器人语音交互系统设计 -北京联合大学学报（自然科学版）2010, 24 \(1\)](#)

介绍了一种基于Microsoft Speech SDK 5.1的机器人语音交互系统, 利用Speech SDK5.1提供的应用程序编程接口SAPI进行语音识别, 对识别结果在编辑程序中处理, 使用Interphonic 5.0语音合成技术替代TTS技术来合成语音, 实现了AHRH- I 接待机器人的语音对话和语音控制.

10. 学位论文 [刘遥峰 基于情感交互的仿人头部机器人研究 2009](#)

智能机器和情感的结合是一个重要的课题。情感交互的实现应该是任何智能机器想要和人类沟通的中心方面。为了与人自然的交流，机器人必须既能读懂用户的情感，又能表达自身的情感。这就需要机器人通过视觉感知用户的面部表情，通过听觉感知用户的话语、声音的韵律、分析话语的情感。人类在赋予机器人情感感知能力的同时还必须赋予机器人能够表达情感的面部表情，可以转动的眼睛，有韵律的声音和做出各种姿势的身体。本研究的目的是设计一台机器人，使它可以与人互动，并在日常生活中和常见的地方协助人类。为了完成这些任务，机器人必须友好地显示出一些情感，表现出友好的特点和个性。依据仿生学，研制了一台仿人头部机器人，建立了机器人的行为决策模型。该机器人具有人类的6种基本面部表情，以及人脸检测、语音情感识别与合成、情感行为决策等能力，能够通过机器视觉、语音交互、情感表达等方式与人进行有效的情感交互。

本文的创新点如下：

- 1) 设计了一个具有面部表情的情感机器人，旨在通过表情、计算机视觉、语音交互等方式与人类进行自然交流。情感机器人通过红外传感器、麦克风、USB摄像头检测外部环境的信息。PC对收集到的外部环境信息首先进行情感特征提取，分析出语音情感和检测到的人脸的面部表情。语音输出模块，机械头的身体语言和机械头的面部表情构成了情感机器人的情感表达方式。PC通过串口通讯向单片机发送要调用的面部表情数据包和身体语言数据包。PC在经过语音识别和语音合成后通过音响设备输出具有情感的语音。机器人的机械设计是参照成年男性头部，以1:1的比例进行设计，有14个自由度。
- 2) 设计并实现了一种具有语音情感分析和情感语音表达的机器人语音交互系统。在设计时针对目前人与机器人的交流仅限于一些固定词句和简单行为方式的缺陷，分析现有语音识别软件的优点和不足，采用了中科院Pattek ASR V3.4和微软Microsoft Speech SDK相结合的语音识别方法，保证了语音识别的准确率，同时提供了机器人进行一定场景范围内自然语言识别的能力。利用语义分析算法，分析人对机器人语音输入的情感，并通过情绪转换模型根据上一次对话情感和当前输入情感确定机器人输出情感，通过改变机器人语音合成的音调、语速、音量来实现不同情感的表达。最后，将带情感控制符号的语音合成内容编码，通过串口发送给语音合成芯片实现带有情感的语音合成。在整个过程中，机器人将分析用户的命令，如果需要进行适当的动作响应，则通过网络通信向ARM控制板发送动作指令控制机器人做适当的动作。
- 3) 建立了能够改变机器人个性特点的情感行为决策模型。在人机交互中，个性是机器人最重要的特点之一，它决定了机器人能否以友好的方式与人类互动，而不让人类感到不舒服。我们通过实施情感行为决策模型来提供机器人自己的个性和特点。利用这个模型，机器人能做出不同行为的决定。如果机器人的个性改变，情感行为决策模型就会使机器人在相同的条件下做出不同的行为决策。情感行为决策模型由5个要素组成：反应动态，内部动态，情感动态，行为动态和个性。机器人的情感模型是线性动态系统，它的5个要素由矩阵来表示。

关键词：仿人头部机器人，情感交互，语音情感交互，情感模型

本文链接：[http://d.g.wanfangdata.com.cn/Periodical\\_jsjsd200702017.aspx](http://d.g.wanfangdata.com.cn/Periodical_jsjsd200702017.aspx)

授权使用：上海交通大学(shjtdxip)，授权号：1a3f2689-58ce-4c09-b73d-9e14014a78b5

下载时间：2010年10月19日