

利用 Speech SDK 实现中英文全文朗读

蒋铁海

(广东外语外贸大学英文学院信息管理系,广州 510420)

摘 要: Speech SDK 是 Microsoft 提供的计算机语音开发工具包,它使我们在程序或网页中实现文字声音的播放或将声音转换成电子文本。本文通过实例介绍了在 VB 程序和动态网页设计中,利用 Speech SDK 来实现中英文全文朗读的方法。

关键词: Speech SDK ;TTS ;语音识别(SR);VB ;DHTML

1 Speech SDK 简介与安装

计算机系统的语音能力指用声音来播放文字的能力(指文字到语音或 TTS,也称为语音合成),或者指将声音转换成电子文本的能力(指语音识别或 SR)。语音识别和语音合成是当今计算机应用的一个非常重要也是非常活跃的领域,它们在自动控制、机器阅读和识别、人机对话、多媒体、计算机辅助教学等领域里有着十分广泛的应用。如果你想使用 Microsoft 的编程语言来开发具有语音识别或语音合成功能的应用软件,那么你就可以使用 Microsoft 提供的开发工具 Speech SDK (Software Development Kit)或 .NET Speech SDK (后者目前只有 Version 1.0 Beta2 版,并且必须在 Visual Studio .NET 中使用)。为此,你必须先安装 Speech SDK,目前 Speech SDK 的最高版本是 5.1 版,可以到 Microsoft 指定的官方网站 <http://www.microsoft.com/speech/download/sdk51/> 上去免费下载。为了能够使程序或网页能够朗读中文,你还必须下载 Language Pack,其中就有简体中文语言包。下载后先安装 Speech SDK 5.1,然后再安装简体中文语言包,这样你的计算机就具有中英文语音能力了。安装完之后,可以看到 Windows 的“控制面板”中多了一个“语音”的控制项目,在这个项目中你可以设置语音的各种属性。安装好 Speech SDK 之后,电脑上也同时安装好了语音编程接口 Win32 Speech

API(SAPI),这样,你就可以在编程语言或动态网页中使用 Win32 Speech API(SAPI)来开发应用程序了。语音识别牵涉的因素很多,因此,其编程相对讲来也比较复杂,所以本文只讲述利用 Speech SDK 实现中英文全文朗读的方法。

2 在 VB 程序中实现中英文全文朗读

图 1 就是一个能够实现中英文全文朗读的 VB 程序。在这个程序中,你可以直接输入一些词语,或者使用“打开文件”命令打开一个文本文件(注意:只能是文本文件);如果文本框里的内容是英文,则你可以选择多种声音(或者说语音、音色或角色)来朗读,如果文本框里的内容是中文,那么你就应该选择“Microsoft Simplified Chinese”,再单击“朗读”命令按钮就可以朗读出文本框里的内容。

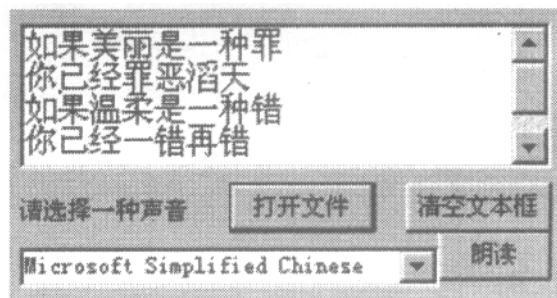


图 1 在 VB 程序中实现全文朗读

为了实现上述功能,在启动 VB 之后请选择

“工程”菜单中的“引用”，然后选择“Microsoft Speech Object Library”，这样你就可以使用 Win32 Speech API 来编程了。先定义二个窗体级的对象变量 sp 和 speakFlags，它们的具体定义如下：

```
Dim sp As SpVoice
Dim speakFlags As SpeechVoiceSpeakFlags
```

SpVoice 是 Speech API 中最重要的对象，其主要的属性有 :Voice 和 GetVoices。GetVoices 是一个数组，其中的元素就是声音的种类，Voice 属性表示当前所使用的声音。SpVoice 最常用的方法是：Speak，利用 SpVoice 对象的 Speak 方法就可以朗读文本文字的内容。

图 1 所对应程序的窗体装载事件的代码是：

```
Private Sub Form_Load( )
Set sp = New SpVoice
Dim Token As ISpeechObjectToken
For Each Token In sp.GetVoices
Combo1.AddItem (Token.GetDescription())
Next
Combo1.ListIndex = 0
End Sub
```

这样，程序一运行时就会在组合框里把所有的声音(或者说语音)列出来。接着，我们就要设计程序的主要部分了。

```
Private Sub Combo1_Click( )
Set sp.Voice = sp.GetVoices( )Item(Combo1.ListIndex)
End Sub
Private Sub Command1_Click( ) ' 朗读
sp.Speak Text1.Text, speakFlags
End Sub
Private Sub Command2_Click( ) ' 打开文件
On Error GoTo out:
Dim strfilename As String
CommonDialog1.InitDir = App.Path
CommonDialog1.Filter = "TextFiles(*.txt)*.txt|AllFiles(*.*)
|*.*|"
CommonDialog1.ShowOpen
strfilename = CommonDialog1.FileName
If strfilename <> "" Then
Open strfilename For Input As #1
Do Until EOF(1)
Line Input #1, s
Text1.Text = Text1.Text + s + Chr(13) + Chr(10)
Loop
```

```
Close #1
End If
out:
Exit Sub
End Sub
```

要注意的是 Speech SDK 目前还不能很好地处理有中英文混合的文本，如果中文中夹有英文单词，则把英文单词分成一个个字母来朗读。如果英文中夹有中文字或词语，则不朗读其中的中文内容。

3 在网页中实现中英文全文朗读

只要你的电脑安装好了 Speech SDK5.1，还可以把 Win32 Speech API 应用到 DHTML 中，这样就可以在网页中实现中英文全文朗读了。图 2 就是在网页中实现的效果图。用浏览器打开网页后，在图 2 的文本框中输入文字内容，然后从组合框里选择一种声音，再单击“朗读”，就可以把文本框里的内容朗读出来。

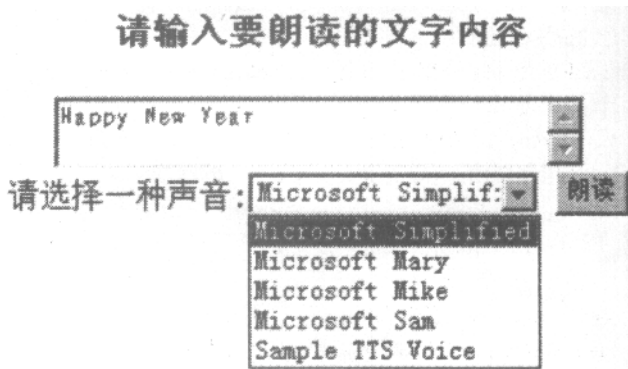


图 2 在网页中实现全文朗读

这实际上是 ActiveX 的应用，如果用 JavaScript 来实现，则其代码的主要部分如下：

```
<SCRIPT LANGUAGE="JavaScript">
var VoiceObj = new ActiveXObject("Sapi.SpVoice");
function ChangeVoice( ) {
var i = parseInt( idsVoices.value );
VoiceObj.Voice = VoiceObj.GetVoices( )Item(i);
}
function SpeakText( ){
if( idbSpeakText.value == "朗读" )
{
try
{
VoiceObj.Speak( idTextBox.value, 1 );
```

```

    }
    catch(exception)
    {
        alert("Speak error");
    }
}
else if( idbSpeakText.value == "停止")
{
    VoiceObj.Speak( "", 2);
}
}
</SCRIPT>
<SCRIPT FOR =" window" EVENT =" OnQuit ( ) LAN-
GUAGE="JavaScript">
    delete VoiceObj;
</SCRIPT>
<H3 align=center>请输入要朗读的文字内容</H3>
<P align =center><TEXTAREA ID=idTextBox COLS=35
ROWS=2 WRAP=VIRTUAL>Happy New Year</TEXTAREA>
<br>
<Align=center>请选择一种声音:<SELECT id=idsVoices
name = Voices onchange = ChangeVoice ( )style = "WIDTH:
148px"></SELECT>
<STRONG><BUTTON id=idbSpeakText onclick=Speak-
Text( )>朗读
</BUTTON></STRONG></p>
<SCRIPT LANGUAGE="JavaScript">
    InitializeControls( );
    function InitializeControls( )
    {
        var VoicesToken = VoiceObj.GetVoices( );
        var AudioOutputsToken = VoiceObj.GetAudioOutputs( );
        for( var i=0; i<VoicesToken.Count; i++)
    
```

```

    {
        var oOption = document.createElement("OPTION");
        idsVoices.options.add(oOption);
        oOption.innerText = VoicesToken.Item(i).Get
        Description( );
        oOption.value = i;
    }
}
function VoiceObj::StartStream( ){
    idbSpeakText.value = "停止";
}
function VoiceObj::EndStream( ){
    idbSpeakText.value = "朗读";
}
</SCRIPT>

```

同样,在网页中实现文本朗读时仍然要注意中英文混合的情况。

结 语

使用 Speech SDK 可以在程序和网页中实现中英文全文朗读的功能,其朗读效果也是目前语音技术中最好者之一,只要我们充分利用 Speech API 中的各种对象的属性和方法,我们就能设计出各种复杂的语音处理程序。

参考文献

- [1]http://www.microsoft.com/speech
- [2]陈涌泉. Visual Basic 第三方控件大全. 北京:中国电力出版社,2002

(收稿日期 2003-06-23)

Realization of Speaking Chinese or English Full-text With Speech SDK

JIANG Tie-hai

(Guangdong University of Foreign Studies, Guangzhou 510420 China)

Abstract Speech SDK is a computer speech developing package which provided by Microsoft. We can realize Text-to-Speech (TTS) or speech recognition in program or home page with Microsoft speech SDK. This paper introduces the method of realization of speaking Chinese or English full text in VB program or DHTML with speech SDK.

Key words :Speech SDK ;TTS ;Speech Recognition ;VB ;DHTML