

Text to Audio

1. Purpose of the experiment

Realize text-to-audio synthesis. (This experiment requires an internet connection to proceed!)

2. Experimental path source code

Enter the dog's system, end the dog program, enter "ip (ip is the dog's ip):8888" in the browser, enter the password "yahboom"



Password: yahboom

Log in

and log in. Enter the path of `cd ~/DOGZILLA_Lite_class/2.Base Control/04.Text synthesized audio` and run `dog_tts.ipynb`.

3. Experimental Phenomenon

After running the code, the dog will synthesize an audio result.wav based on the added characters and save the audio in the file /home/pi/xgoVideos/. After the synthesis is completed, the synthesized audio will be played automatically. (This experiment requires an Internet connection to proceed!)

```
[2]: XGO_edu.SpeechSynthesis("你好你好") #chinese 生成的文件在 /home/pi/xgoMusic/result.wav (需要联网)
#XGO_edu.SpeechSynthesis("Hello,Hello") #English The generated file is in /home/pi/xgoMusic/result.wav (Requires network)

%E4%BD%A0%E5%A5%BD%E4%BD%A0%E5%A5%BD
test on Web Browserhttp://tsn.baidu.com/text2audio?tok=24.0b8d8a64b3f0c4ee9bbd903168584fbb.2592000.1744188516.282335-35511852&tex
=%25E4%25BD%25A0%25E5%25A5%25BD%25E4%25BD%25A0%25E5%25A5%25BD&per=0&spd=5&pit=5&vol=5&aue=6&cuid=123456PYTHON&lan=zh&ctp=1
result saved as :result.wav
MPlayer 1.4 (Debian), built with gcc-10 (C) 2000-2019 MPlayer Team

Playing /home/pi/xgoMusic/result.wav.
libavformat version 58.45.100 (external)
Audio only file format detected.
Load subtitles in /home/pi/xgoMusic/
=====
Opening audio decoder: [pcm] Uncompressed PCM audio decoder
AUDIO: 16000 Hz, 1 ch, s16le, 256.0 kbit/100.00% (ratio: 32000->32000)
Selected audio codec: [pcm] afm: pcm (Uncompressed PCM)
=====
AO: [pulse] Init failed: Connection refused
Failed to initialize audio driver 'pulse'
AO: [alsa] 48000Hz 1ch s16le (2 bytes per sample)
Video: no video
Starting playback...
A: 1.2 (01.2) of 1.0 (01.0) 0.2%

Exiting... (End of file)
```

4. Main source code analysis

The parameter passed by `SpeechSynthesis("Hello,Hello")` is the string of audio to be synthesized.

