**linux环境下用python将文字转换成语音（不调用api）**

[gswyhq](https://www.zhihu.com/people/gswyhq)

[**gswyhq**](https://www.zhihu.com/people/gswyhq)

​关注他

4 人赞同了该文章

一、原理

将每个汉字转换成对应的拼音，再通过拼音调用对应的音频文件，拼接成最后的音频文件；缺陷是合成的结果似在无感情朗读。

二、准备工作

1、下载单字的音频文件，下载页面：[eGuideDog free software for the blind](https://link.zhihu.com/?target=https%3A//sourceforge.net/projects/e-guidedog/files/Ekho%2520Voice%2520Data/0.2/" \t "_blank)

找到一个前缀是‘pinyin’音频数据下载，比如下载[eGuideDog free software for the blind](https://link.zhihu.com/?target=https%3A//sourceforge.net/projects/e-guidedog/files/Ekho%2520Voice%2520Data/0.2/pinyin-yali-44100-v9.tar.bz2/download" \t "_blank)

下载音频文件解压；

gswyhq@gswyhq**-**pc**:~**$ tar xJvf pinyin**-**huang**-**44100**-**wav**-**v2.tar.xz

...

gswyhq@gswyhq**-**pc**:~**/pinyin$ ls

a1.wav che3.wav diao4.wav gou1.wav jin4.wav li1.wav mo4.wav pao1.wav rong2.wav shuo4.wav wo1.wav zang3.wav

a2.wav che4.wav diao5.wav gou2.wav jin5.wav li2.wav mo5.wav pao2.wav rong3.wav shuo5.wav wo2.wav zang4.wav

...

2、安装程序包：

# sudo apt-get install ffmpeg

# sudo apt-get install pydub

# sudo pip3 install pypinyin

注：*AudioSegment原生就支持wav和raw，如果其他文件需要安装ffmpeg。*

如果安装了 [jieba](https://link.zhihu.com/?target=https%3A//github.com/fxsjy/jieba" \t "_blank) 分词模块，pypinyin会自动调用，以处理多音字的问题。

三、过程及完整代码

#!/usr/bin/python3

*# coding: utf-8*

**import** os

**from** pypinyin **import** lazy\_pinyin, TONE3

**from** pydub **import** AudioSegment

*# 中间插入空白静音500ms。*

silent **=** AudioSegment**.**silent(duration**=**500)

*# 单字音频文件的路径*

PINYIN\_VOICE\_PATH **=** '/home/gswyhq/pinyin'

*# 最终合成的音频文件路径*

EXPORT\_PATH **=** '/home/gswyhq/data/新文件夹'

**def** **load\_voice\_dict**():

voice\_file\_list **=** [f **for** f **in** os**.**listdir(PINYIN\_VOICE\_PATH) **if** f**.**endswith('.wav')]

voice\_dict **=** {}

**for** voice\_file **in** voice\_file\_list:

name **=** voice\_file[:**-**4]

song **=** AudioSegment**.**from\_wav(os**.**path**.**join(PINYIN\_VOICE\_PATH, voice\_file))

voice\_dict**.**setdefault(name, song)

**return** voice\_dict

VOICE\_DICT **=** load\_voice\_dict()

**def** **txt\_to\_voice**(text, name**=**'test'):

"""

将文字转换为音频

:param text: 需要转换的文字

:param name: 生成的音频文件名

:return:

"""

pinyin\_list **=** lazy\_pinyin(text, style**=**TONE3)

new **=** AudioSegment**.**empty()

**for** piny **in** pinyin\_list:

piny\_song **=** VOICE\_DICT**.**get(piny)

**if** piny\_song **is** **None** **and** piny **and** piny[**-**1] **not** **in** '0123456789':

*# 没有音调*

piny **=** piny **+** '5'

piny\_song **=** VOICE\_DICT**.**get(piny, silent)

*# 交叉渐入渐出方法*

*# with\_style = beginning.append(end, crossfade=1500)*

*# crossfade 就是让一段音乐平缓地过渡到另一段音乐，crossfade = 1500 表示过渡的时间是1.5秒。*

crossfade **=** min(len(new), len(piny\_song), 1500)**/**60

new **=** new**.**append(piny\_song, crossfade**=**crossfade)

*# new += piny\_song*

new**.**export(os**.**path**.**join(EXPORT\_PATH, "{}.mp3"**.**format(name)), format**=**'mp3')

**def** **main**():

text **=** ''' 红海早过了。船在印度洋面上开驶着。但是太阳依然不饶人地迟落早起侵占去大部分的夜。

夜仿佛纸浸了油，变成半透明体；它给太阳拥抱住了，分不出身来，也许是给太阳陶醉了，

所以夕照霞隐褪后的夜色也带着酡红。到红消醉醒，船舱里的睡人也一身腻汗地醒来，洗了澡赶到甲板上吹海风，

又是一天开始。这是七月下旬，合中国旧历的三伏，一年最热的时候。在中国热得更比常年利害，

事后大家都说是兵戈之象，因为这就是民国二十六年【一九三七年】。'''

txt\_to\_voice(text)

**if** \_\_name\_\_ **==** '\_\_main\_\_':

main()

参考资料：

[Python实现文字转语音功能](https://zhuanlan.zhihu.com/p/26726297)