

حامد زارعی

تکلیف سوم پردازش گفتار رقمی

اول

برای نصب و استفاده از ابزار htk بر روی لینوکس از لینک زیر و با استفاده از داکر این کار را انجام دادم.

<https://github.com/loretoparisi/htk>

config_file مورد نیاز برای MFCC در htk مانند زیر خواهد شد

```
# Feature configuration
TARGETKIND = MFCC_E_D_A
TARGETRATE = 100000.0
WINDOWSIZE = 250000.0
USEHAMMING = T
PREEMCOEF = 0.97
NUMCHANS = 24
NUMCEPS = 12
# input file format (headerless 8 kHz 16 bit linear PCM)
SOURCEKIND = WAVEFORM
SOURCEFORMAT = NOHEAD
SOURCERATE = 1250
```

و فایلی که لیست فایل‌های صوت با مسیر فایل ویژگی‌ها مانند:

```
data/train/tr_f_01_01.wav features/train/f_01_01_mfcc
```

و به همین ترتیب برای LPC فقط برای قسمت TARGETKIND باید LPC_E_D_A قرار دهیم.

```
HCopy -C config_file_mfcc -S file_list_mfcc
```

```
HList features/train/f_01_01_mfcc
```

دوم

بعد از اضافه کردن مسیر voicebox به مسیر متلب قابلیت اجرای دستورهایمانند زیر را خواهیم داشت.

- V_readwav
- V_melcepst

```
addpath('E:\University\Digital_Speech_Recognition\sap-voicebox\voicebox');
v1 = v_readwav('data/train/tr_f_01_01.wav');
v1_mel = v_melcepst(v1, 7000, 'EdD', 12);
```

برای train_01_01

خروجی htk

```
-12.139 -4.039 -1.920 6.949 -7.643 3.846 9.557 4.891 8.217 -2.834
0.241 7.965 0.988 1.536 2.224 4.418 -0.942 2.120 0.533 -1.555
-1.500 -2.084 0.783 0.342 -2.593 -0.066 0.040 0.036 -0.320 -0.116
0.043 -0.206 0.375 -0.128 0.095 0.346 0.296 0.402 0.005
```

برای متلب

```
8.21816254082468    2.53083334490085    2.88023069221089    -0.207168469279081
0.524081555550490    0.440583239776224    -1.04460505358118    -0.423716480728678
-1.13560613439527    -1.15281615040190    -1.26715807662754    -1.53791752096781
-1.32302197350227    0.158324452242927    -0.767113816884346    -0.128849693642646
-0.109072052789220    -0.238277199973421    0.255596770326694    0.0116701650764305
-0.0436216920619230    0.0175234994450618    0.0481563095417257    0.0329182723724817
0.0691406279471068    -0.106895094541920    0.107292506338658    -0.252499572367142
-0.179008549534354    0.0437270218525535    -0.0360731622848838    0.0720426837942674
0.0275004359307543    -0.0742842827518237    0.0430297302898902    0.0383447029132190
0.0248339760191854    0.0445784927452263    -0.0256017641050127
```

یکسان بدست نیامده است که دلایلی مانند پارامترهای ثابت در تابع‌های فراخوانی شده می‌تواند باشد.

مانند SOURCERATE, TARGETRATE,...