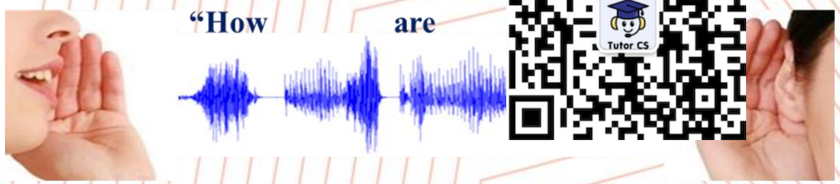
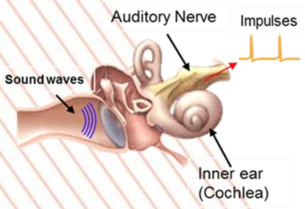
程序代写代做 CS编程辅导

ELEC3104: Mini-Project – Cochlear Signal Processing





WeChat: cstutorcs

TLT – Level 5 (High Distinction Level): Using the Level 2 FIR/ IIR cochlear filter bank model, develop a method for pitch detection of speech signals. Then, use the pitch information to extract a speaker's voice from a mixture of two voices at the input of the filter bank.

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

Complete TLT-Levels 2, 3 and 4 first, and ensure that you are on the right track before proceeding to TLT – Level 5

https://tutorcs.com

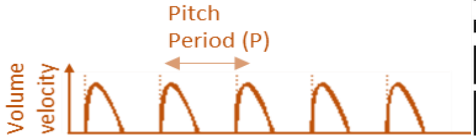
Prof. Eliathamby Ambikairajah, School of EE&T

Term 3, 2024

Classification of Speech Sounds: divided into two broad classes

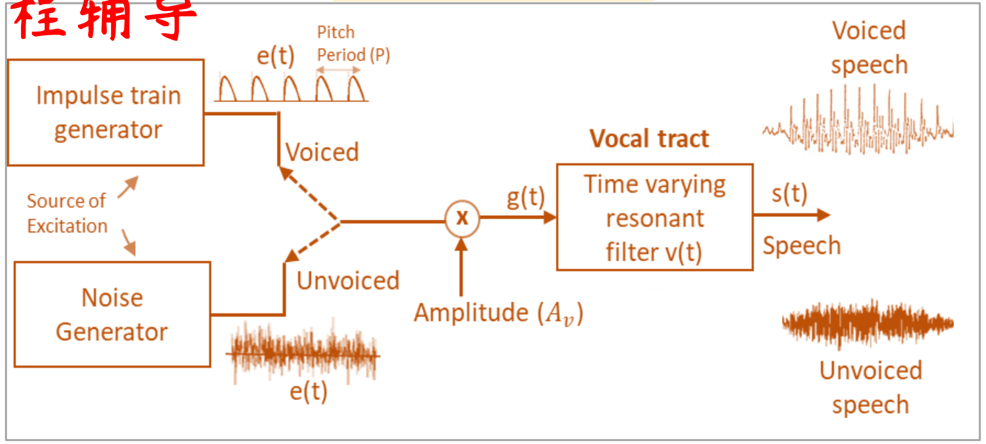
Voiced Sound

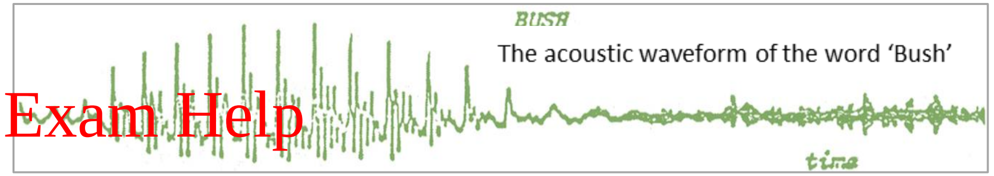
- ✓ VOICED sound (e.g. /a/, /e/, /i/ …) is produced when the vocal cords vibrate during the production of a sound.
- ✓ Voiced speech occurs when the air flows throug…rds into the vocal tract in discrete "puffs" rather than as a c…v.



- ✓ The vocal cords vibrate at a particular frequency, which is called the fundamental frequency (defined as pitch) of the voiced sound
  - 50 : 200 Hz for male speakers
  - 150 : 300 Hz for female speakers
  - 200 : 400 Hz for child speakers

Speech Production Model



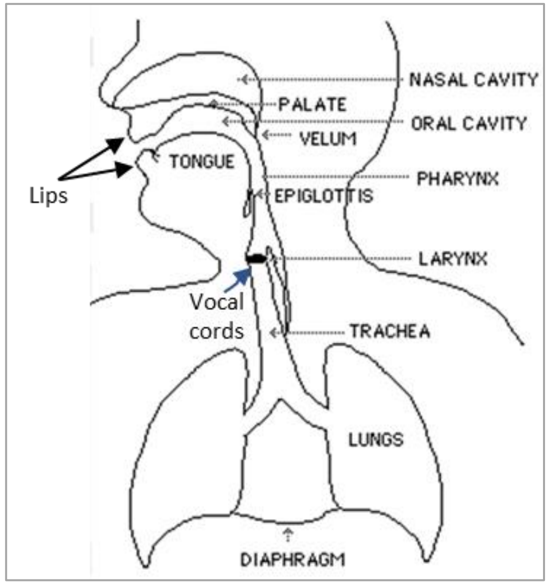


Unvoiced Sound

- ✓ UNVOICED speech (e.g. /s/, /f/…) is produced when vocal cords are not vibrating and they are held open and the air flows continuously through them.

Vocal tract

- ✓ Vocal tract is a non-uniform acoustic tube that is terminated at one end by the vocal cords and at the other end by the lips
- ✓ The vocal tract changes shape rather slowly in continuous speech and it is reasonable to assume that the vocal tract has a fixed characteristic over a time interval of the order of 10 - 20ms
- ✓ Thus once every 10 -20ms, the vocal tract configuration is varied, producing new vocal tract parameters (resonant frequencies)



Speech Production Mechanism

程序代写代做 CS编程辅导
WeChat: cstutorcs
Assignment Project Exam Help
Email: tutorcs@163.com
QQ: 749389476
https://tutorcs.com

Example: The estimation of the pitch period, in the time domain

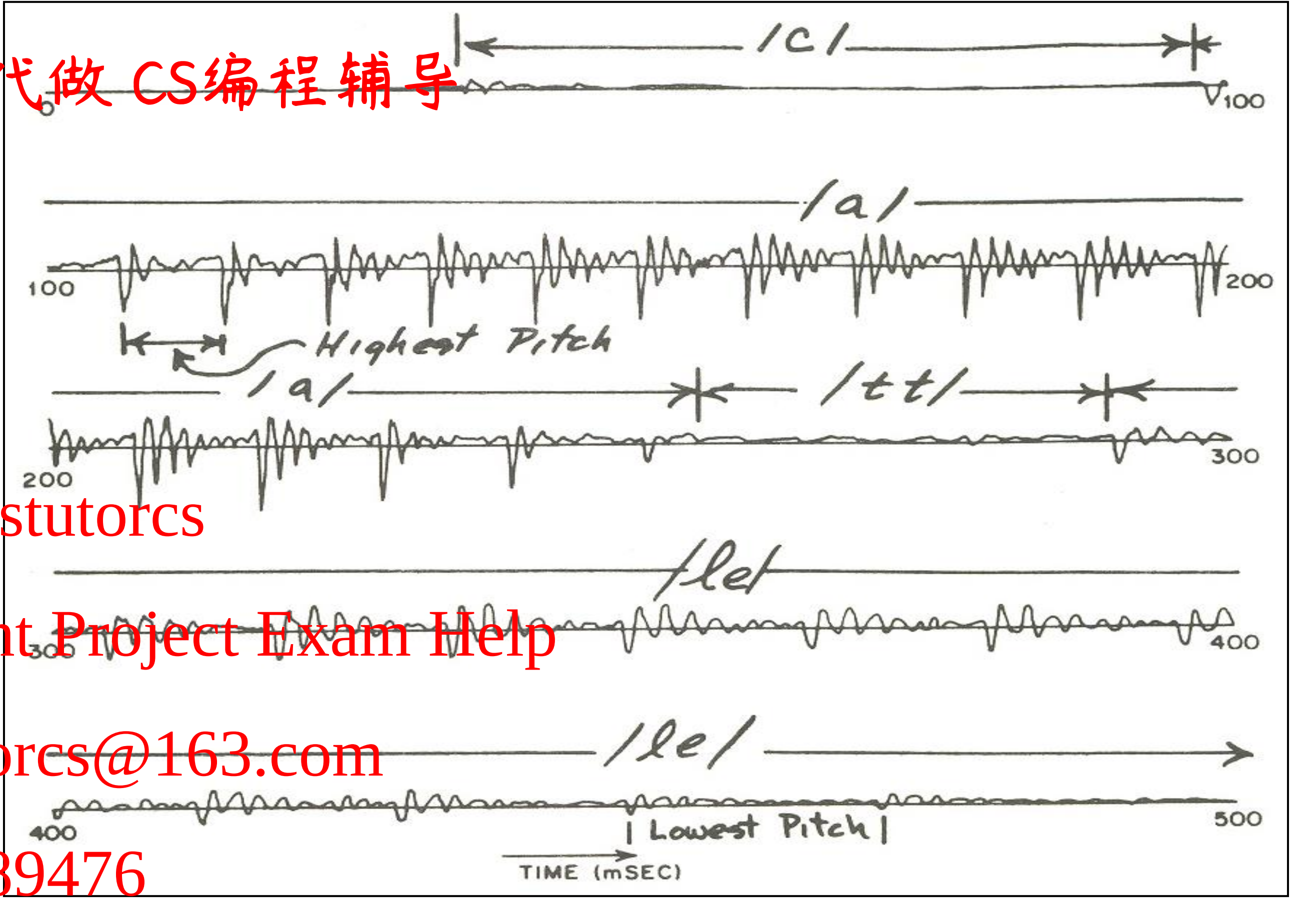
Example:

The waveform on the right, is for the word “cattle”. Note that each line of the plot corresponds to 100 ms of the signal:

(a) the point where the voice pitch frequenc… highest; and (ii) the lowest, is indicated. W… the approximate pitch frequencies at the… points?

(b) Is the speaker most probably a male, or a child? How do you know?

The lowest pitch has a period of about 21.5 ms corresponding to the frequency 46 Hz. This low pitch indicates the speaker is probably male



Speech waveform of the word ‘Cattle’

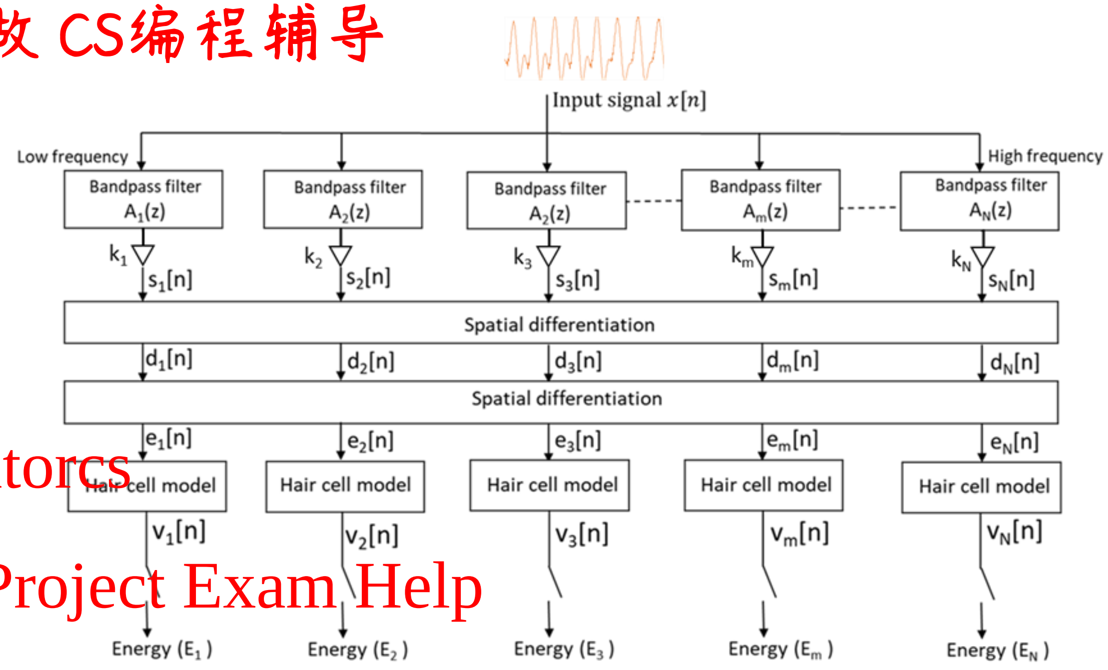
程序代写代做 CS编程辅导
WeChat: cstutorcs
Assignment Project Exam Help
Email: tutorcs@163.com
QQ: 749389476
https://tutorcs.com

Pitch period estimation of Voiced Speech using the Parallel filter bank model

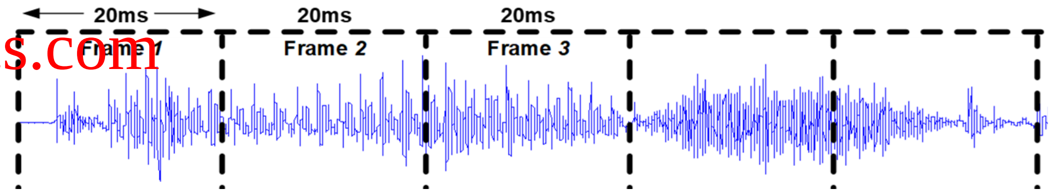
Pitch Detection

程序代写代做 CS编程辅导

- The pitch period estimation of voiced speech is normally carried out in the time domain using an Autocorrelation Function or an Average Magnitude Difference function (AMDF)
- In this mini-project, we will determine the pitch period of a voiced speech segment using a frequency domain technique, i.e. using the spectrum analyzer that you designed and built in the TLT-LEVEL 3 mini-project.
- Periodicity in the time domain, results in useful impulses in the frequency domain at the fundamental frequency and its harmonics. If the speech signal is periodic or quasi periodic, the spectrum analyzer will show peaks in multiples of the fundamental frequency f_0.
- Download the voiced speech file 'voiced1_16k.wav' (in Moodle: Resources-> 'Data for Project' folder) and divide the voiced speech signal into small frames (20ms each, see below) and process (filtering) each segment through the spectrum analyser that is shown on the right. (You may use the filter command [Y, Z_f] = filter(B,A,X,Z_i) and make sure that you use the initial/final condition Z)
- The outputs ($v_1[n]$ to $v_N[n]$) of the hair-cell model can be used to extract the fundamental frequency f_0 of the input voiced speech frame.



Spectrum Analyser



Voiced speech

WeChat: cstutorcs
Assignment Project Exam Help
Email: tutorcs@163.com
QQ: 749389476
https://tutorcs.com

Pitch period estimation using the cochlear model

Pitch Detection continued…

程序代写代做 CS编程辅导

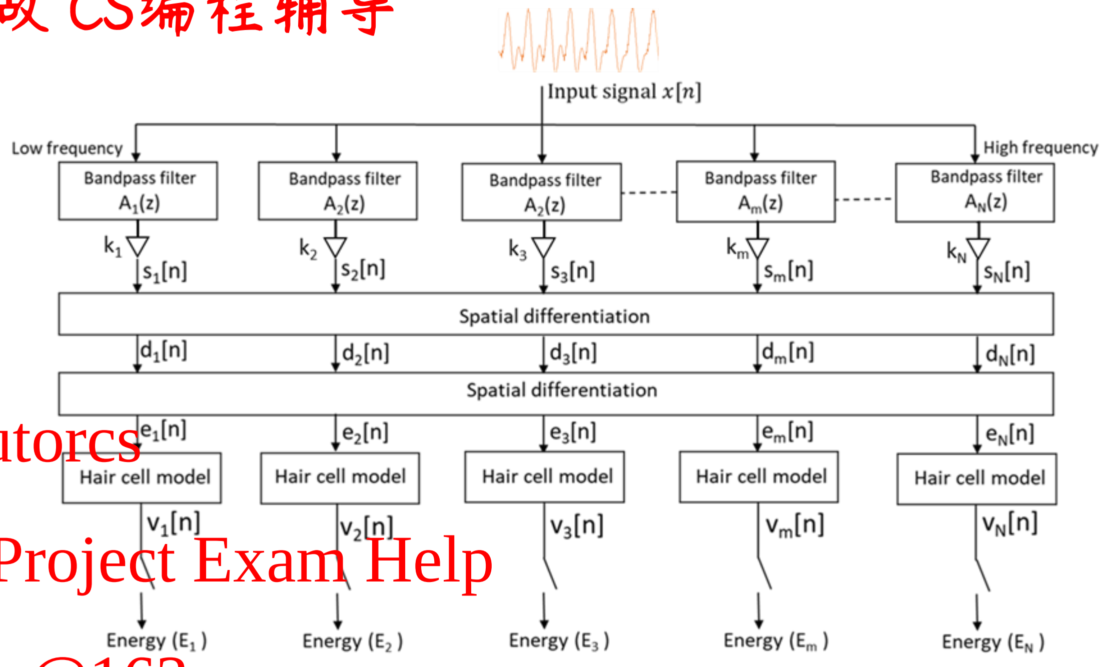
- It is assumed that the **fundamental frequency** of the input speech signal (Male/Female) lies in the range of **102Hz (5^{th} filter) to 303 Hz (36^{th} filter),** so only 32 filters cover thi
- The inner hair-cell output of each filter and the c filters corresponding to its H^{th} harmonics are m together in order to get the harmonic product sp):

$$HPS_i = \prod_{k=1}^{H} v_{ik}$$

where *i* is the filter number (i = 5, 6,…….36), H is the number of harmonics (use H=4 from the table 1), v_{ik} is the hair cell output corresponding to the k^{th} harmonic.

WeChat: cstutorcs
Assignment Project Exam Help
Email: tutorcs@163.com
QQ: 749389476
https://tutorcs.com

- This is carried out for all of the 32 filters within the **fundamental frequency range 102Hz to 303Hz.**
- Table 1 contains the number of filters corresponding to the fourth harmonics of the centre frequency of each of the 32 filters. Using this table, calculate HPS_i for all the 32 filters.
- In order to determine the fundamental frequency, HPS_i is plotted against the filter number (5 to 36).
- A peak-picking algorithm is used to determine the position of the highest peak and that position (filter number) corresponds to the fundamental frequency of the input voiced speech frame.



Spectrum Analyser

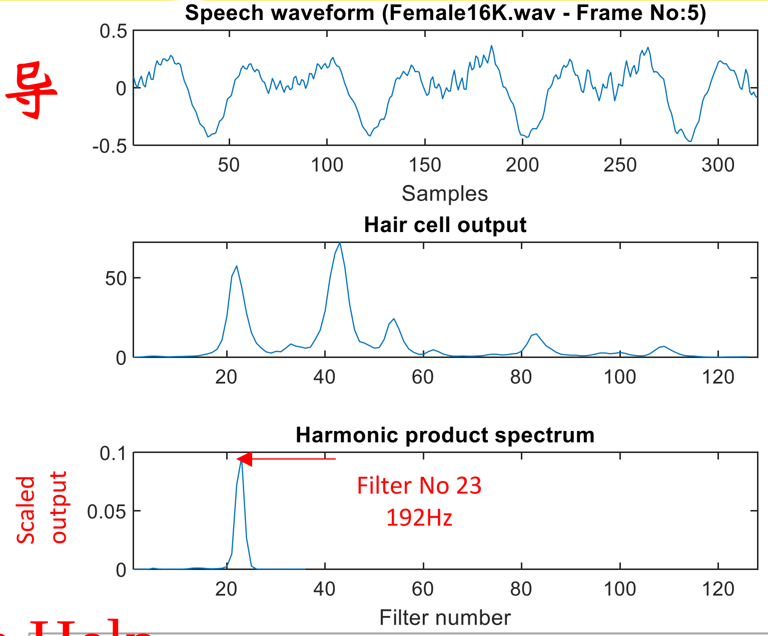
Table 1: 2^{nd} to 6^{th} harmonics of the filters 1 to 36

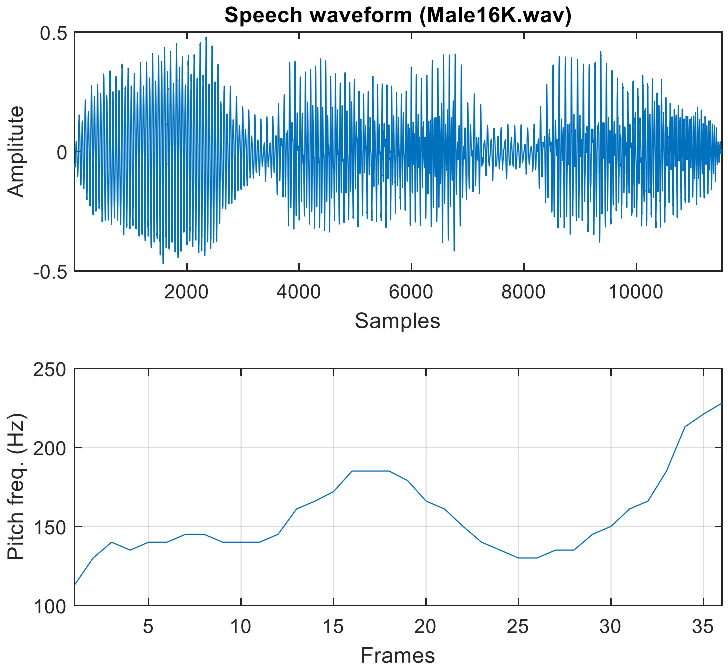
Fundamental frequency (f0) Hz	Filter No	2nd harmonic (2f0) Filter No	3rd harmonic (3f0) Filter No	4th harmonic (4f0) Filter No	5th harmonic (5f0) Filter No	6th harmonic (6f0) Filter No
88	1	21	32	40	47	52
91	2	22	33	41	48	53
95	3	23	34	42	49	54
98	4	24	35	43	50	55
102	5	25	[illegible]	44	51	56
105	6	26	[illegible]	45	52	57
109	7	27	[illegible]	46	53	58
113	8	28	[illegible]	47	54	59
117	9	29	[illegible]	48	55	60
121	10	30	[illegible]	49	56	61
126	11	31	[illegible]	50	57	62
130	12	32	43	51	58	63
135	13	33	44	52	59	64
140	14	34	45	53	60	65
145	15	35	46	54	61	66
150	16	36	47	55	62	67
155	17	37	48	56	63	68
161	18	38	49	57	64	69
166	19	39	50	58	65	70
172	20	40	51	59	66	71
179	21	41	52	60	67	72
185	22	42	53	61	68	73
192	23	43	54	62	69	74
198	24	44	55	63	70	75
205	25	45	56	64	71	76
213	26	46	57	65	72	77
221	27	47	58	66	73	78
228	28	48	59	67	74	79
237	29	49	60	68	75	80
245	30	50	61	69	76	81
253	31	51	62	70	77	82
263	32	52	63	71	78	83
272	33	53	64	72	79	84
282	34	54	65	73	80	85
292	35	55	66	74	81	86
303	36	56	67	76	82	87

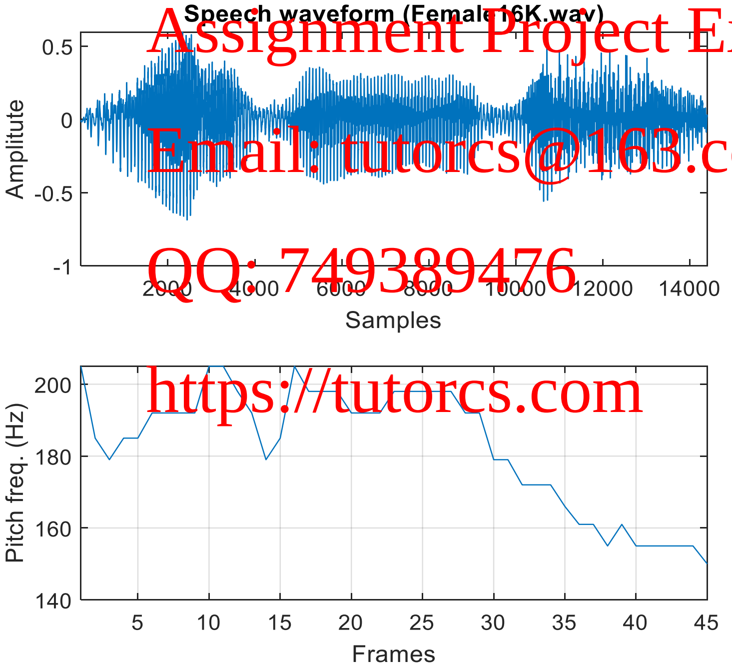
程序代写代做 CS编程辅导
WeChat: cstutorcs
Assignment Project Exam Help
Email: tutorcs@163.com
QQ: 749389476
https://tutorcs.com

Are you on the right track?

- ✓ The figure on the right shows one frame (20ms – [illegible] samples) of voiced speech, the corresponding hair cell output and the harmonic product spectrum.
- ✓ It can be seen from the hair cell output, that it is difficult to locate the fundamental frequency position.
- ✓ When the hair cell output is processed using the harmonic product spectrum (HPS), the amplitude corresponding to the fundamental frequency is enhanced, while the peaks corresponding to other frequencies are suppressed. The highest peak position (filter number or frequency) corresponds to the fundamental frequency of the voiced speech frame.





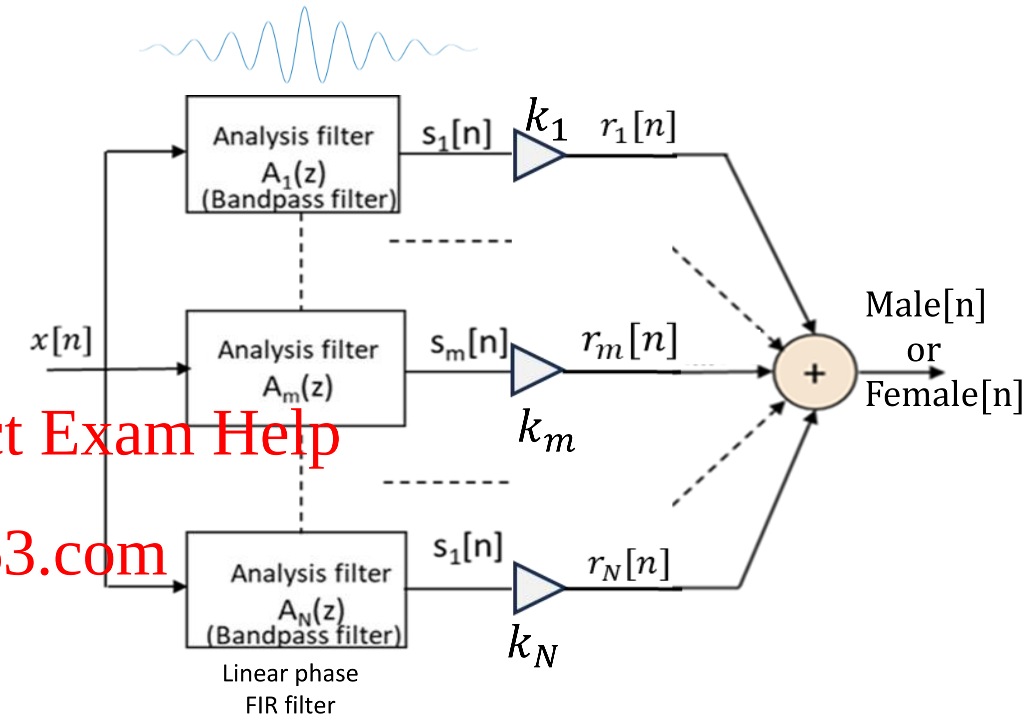


- ✓ The figure on the left shows the speech waveforms for a female and a male speaker. The fundamental frequency, extracted using the Harmonic Product Spectrum (HPS) method for each 20 ms frame, is also displayed. Both speakers produced the same utterance, **'We were away**' which consists entirely of voiced frames, facilitating the extraction of the pitch period for each utterance
- ✓ Download the 'Male16K.wav' file and 'Femal16K.wav' file and implement your HPS algorithm and you should get the similar output as shown in the figure on the left.

程序代写代做CS编程辅导
WeChat: cstutorcs
Assignment Project Exam Help
Email: tutorcs@163.com
QQ: 749389476
https://tutorcs.com

Final Implementation -Speaker Separation

- ✓ Add the two waveforms to obtain the input signal $x[n]$ and pass it through the linear phase analysis filter to obtain the outputs $\{s_1[n]$ to $s_N[n]$ } (see diagram on the right)
- ✓ Locate the filter corresponding to the male pitc[illegible]rom the previous part, as well as the filters correspondi[illegible]rmonic frequencies.
- ✓ Set the gain k corresponding to all these filters to 1 and set the gain k for all other filters to zero (see figure on the right)
- ✓ Sum the output and use the sound command to check if you can hear only the male speaker
- ✓ Repeat this process for the female speaker and check if you can hear only the female speaker
- ✓ You may also experiment with different scalar values to find which values k_1 to k_N are suitable for separating the speakers.
- ✓ Show your reconstructed signal to your lab demonstrator.



程序代写代做 CS编程辅导
WeChat: cstutorcs
Assignment Project Exam Help
Email: tutorcs@163.com
QQ: 749389476
https://tutorcs.com