Test Information

Description: After you completed the assignment, you may get a mark (probably 0), it is a dummy result and not the real result because the true answers have not been input to the systems yet. The real result will be published 2 weeks or more after the assignment deadline.

Instructions: Hard deadline: late submission is not allowed.

Please click on the "save and submit" button at the end of the page after you completed the assignment.

If your answer is not an integer, give it to the nearest 3 decimal places.
If your answer is an integer, add decimal point has no harm. E.g. 3 = 3.0 = 3.00 = 3.000 etc.

Please complete the assignment before the deadline. Multiple attempts are allowed.

Multiple Attempts: This test allows multiple attempts.

Force Completion: This test can be saved and resumed later.

Your answers are saved automatically.

Question Completion Status:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

QUESTION 1

10 points Save Answer

A speech signal (single channel) is sampled at 44100 Hz and the resolution is 16 bits. How many bytes are used to store 70 seconds of this signals?

QUESTION 2

10 points Save Answer

A speech signal is sampled at 44100 Hz . The frame size is 477 ms and the overlapping ratio is 70% . Calculate the number of overlapped samples of adjacent frames. Your answer may not be an integer, give your answer in decimal points, however, our accepted answer is +/-1 sample.

QUESTION 3

10 points Save Answer

A speech signal is sampled at 1433 Hz . A low pass filter (from 0 to X Hz) should be placed between the microphone amplifier and the analogue-to-digital converter to avoid antialiasing (failure to sample). Calculate the value of X.

QUESTION 4

10 points Save Answer

The transformation from signal to Cepstral coefficients is achieved by the following 4 processors. Please select the order of execution on the left side of the processor shown below (order 1 is the first being processed etc.) .

- IDFT
- DFT
- windowing
- LOG|X(m)|

QUESTION 5

10 points Save Answer

A frame segment of a sound signal X has a sequence of time domain samples: X=[1,3,6,1,6,7,1,8,3,6,6,7,9,5,8]

The whole sequence is one segment, no frame blocking and framing is required here. You can calculate the auto-correlation parameters r0, r1, r2, r3 and r4 from X. Find (r0 + r1 + r2 + r3 + r4) .

QUESTION 6

10 points Save Answer

Linear predictive coding LPC coefficients--

A frame segment of a sound signal X has a sequence of time domain samples: X=[1,3,6,1,6,7,1,8,3,6,6,7,9,5,8]

The linear predictive coding LPC coefficients of order 4 are a1, a2 , a3 and a4. Find a4.

QUESTION 7

10 points Save Answer

Given a 2-D data set of

X=[2.2 1.3
2.3 0.7
0.2 0.6
2.4 5.6
4.5 7.8
2.3 2.6
5.5 5.6
5.7 8.9
1.2 3.4
4.5 4.7];

Use the Binary-split K-means method with perturbation e = 0.01 to cluster the data. The first centroid is found to be c =(cx,cy) . In the first stage, the data X is then split into two sets, the centroids of the two sets (set 1 and set2) are (ex1,ey1) and (ex2,ey2), respectively. Find ex1+ey1+ex2+ey2.

QUESTION 8

10 points Save Answer

Dynamic programming
The codes of a reference sound segment, and an unknown input segment are listed below.
reference = [1 5 7 8 4 2 5 1 5 7]; %reference sound segment
input = [1 4 6 8 4 3 4 2 5 6] %input segment

Note: The distortion measure between two sound of code A and B is (A – B)^2.
Use the dynamic programming method to find the optimal distortion between the two sounds.

QUESTION 9

10 points Save Answer

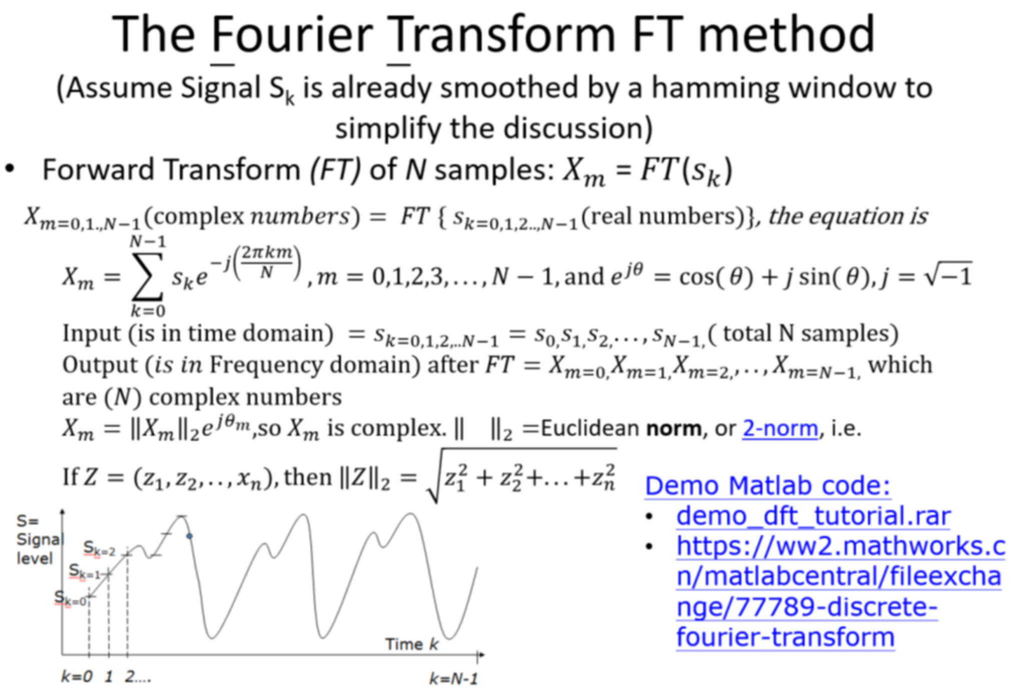
The frequencies of two signals are 220 and 533 Hz. Find the difference between the two signals in MEL scale.

QUESTION 10

10 points Save Answer

Fourier Transform
A signal s= [2 1 4 7 2 3 1 5 7 8 5 4 3 6 7 8 7 4 0 2 3 6]

Using Fourier transform shown in the diagram to find ||Xm||, where m=2 (the energy of the frequency at m.)
Hint, you may use the link shown below or in the diagram to find the
code for the calculation,
Demo Matlab code: demo_dft_tutorial.rar (http://www.cse.cuhk.edu.hk/~khwong/www2/cmsc5707/demo_dft_tutorial.rar) ,



QUESTION 11

10 points Save Answer

speech signal frame blocking
A speech signal is 2.25 seconds in duration. A frame blocking method is applied to obtain the frames, the frame size is 30ms and non-overlapping region is 20ms. Estimate how many frames (+/-1 accuracy allowed) will you obtain for this signal.

QUESTION 12

10 points Save Answer

MFCC (multiple answer question)

- ☐ The term MFCC0 (index 0) is smaller than other MFCC terms.
- ☐ The term MFCC0 (index 0) is bigger than other MFCC terms.
- ☐ The term MFCC0 (index 0) represents the highest frequency term of the signal.
- ☐ MFCC0 (index 0) is not used in speech recognition because it represents the energy term and irrelevant to the frequency content of the signal.

QUESTION 13 20 points Save Answer

Dynamic programming

An accumulated distortion score matrix (using the format same as the lecture notes) of matching 2 sequences is shown in the table below.

The optimal distortion score of this accumulated distortion score matrix is P. The number of elements of the optimal path is N (including the first and last elements of the path). Find N+P.
Hints: The diagonal path is preferred (see lecture notes for the definition), and assume we do not apply any restricted region in the optimal search.

60	60	69	70	34	49	37	46	26	26
59	71	107	91	30	24	21	25	22	46
55	71	99	75	29	20	21	21	21	61
39	35	39	39	20	38	27	39	12	13
35	51	79	47	11	2	3	3	12	61
19	15	31	11	2	11	15	24	24	40
18	6	10	2	6	31	47	72	76	75
17	5	1	5	30	79	89	120	95	71
1	1	10	11	15	40	56	75	70	71
0	4	20	24	25	41	50	66	67	76

QUESTION 14 20 points Save Answer

Autocorrelation programming exercise --
You need to use MATLAB for the programming exercises.
Obtain the data from http://www.cse.cuhk.edu.hk/~khwong/www2/cmsc5707/data57074_1.zip
Unzip the files and follow the instructions in readme4b.docx. Run the code "demo_sign1_question".
You will obtain a signal frame "x_frame", the autocorrelation parameters of x_frame are r0,r1,r2,r3,r4 etc. Find r2.

QUESTION 15 20 points Save Answer

MFCC programming exercise--
Obtain the data from http://www.cse.cuhk.edu.hk/~khwong/www2/cmsc5707/data57074_1.zip
Study the matlab documentation of the function "mfcc". Run the following code

```
%% matlab code starts here %%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear
[audioIn,fs] = audioread("s4.wav");
%the source .wav sound is stereo, the following uses one channel of the sound
[coeffs,delta,deltaDelta,loc] = mfcc(audioIn(:,1),fs);
% If W=window size in samples
% M=overlapping sample
% N=number of frames obtained.
% P=number of MFCC parameters obtained.
```

Find W+M+N+P . Hints: +/- 3 is acceptable.

QUESTION 16 20 points Save Answer

MFCC programming exercise--
Based on the above questions, extraction
xa= the 40-th frame of sound file "s4.wav"
xb= the 50-th frame of sound file "s5.wav"
Find the mfcc_distortion between the MFCC vectors of xa and xb.
Hints:
(1) In the output "coeffs" of the MATLAB mfcc function, the n-th frame output is the n-th row of coeffs.
E.g. the 40th row of coeffs is the mfcc vector of frame 40 of the sound file.
(2) Assume the two mfcc vectors (coeffs) are [LogEnergy_m0, m0,m1,m2...,m12], and [LogEnergy_n0,n0,n1,n2,..,n12]. The mfcc_distortion=sqrt [(m1-n1)^2+(m2-n2)^2+,...,+(m12-n12)^2].
Note that LogEnergy and m0 and n0 terms are not used in the calculation. See details in
https://ww2.mathworks.cn/help/audio/ref/mfcc.html?overload=mfcc+false
(3) % How to index a matrix in matlab,

```
M=[1 2 3
   4 5 6
   7 8 9]
M(1,1) %=1
M(2,2) %=5
M(1,3) %=3
```

Click Save and Submit to save and submit. Click Save All Answers to save all answers.

Save All Answers Save and Submit