Advanced Topics in AI Assignment 1(Written question)

Name: Wenli SONG Student ID: 1155114524

1. Signal Pre-processing

1.1.(a)

16bits/8 \* 22200Hz \* 25s = 1110000 Bytes

1.1.(b)



N = 22200Hz \* 30/1000 s = 666 samples

M = N \* (1-75%) = 666 \* 0.25 = 166.5 samples

Overlapped samples = N – m = 666 -166.5 = 499.5 samples

(Method2: 22200Hz \* 30ms \* 75% = 499.5)

1.1.(c)

22200Hz / 2 = 11100Hz

1.1.(d)

Use Discrete Fourier Transform converting time domain X(n) to frequency domain X(w).

The use log, get log|X(w)|.

At laster use Inverse Discrete Fourier Transform converting log|X(w)| to C(n)

1.1.(e)

In cepstrum, the glottal excitation can be easily removed and it is not useful for recognition, we just interested in filter. Apart from this, using LPC is more efficient than Fourier transform.

1.2

When window size is big, the number of samples in each frame is large, the frequency resolution becomes high, but we get few frames of a certain length of signal, and time resolution becomes low.

Similarly, When window size is small, the number of samples in each frame is few, the frequency resolution becomes low, but we get many frames of a certain length of signal, and time resolution becomes high.

So we should trade off window size to make both time and frequency resolution good enough.

1. Linear Predictive Coding

2.1

r0 = 3\*3+1\*1+4\*4+1\*1+5\*5+9\*9+2\*2+6\*6+5\*5+3\*3+5\*5+8\*8+9\*9+7\*7+9\*9

= 9+1+16+1+25+81+4+36+25+9+25+64+81+49+81 = 507

r1 = 1\*3+4\*1+1\*4+5\*1+9\*5+2\*9+6\*2+5\*6+3\*5+5\*3+8\*5+9\*8+7\*9+9\*7

= 3+4+4+5+45+18+12+30+15+15+40+72+63+63 = 389

r2 = 4\*3+1\*1+5\*4+9\*1+2\*5+6\*9+5\*2+3\*6+5\*5+8\*3+9\*5+7\*8+9\*9

= 12+1+20+9+10+54+10+18+25+24+45+56+81 = 365

r3 = 1\*3+5\*1+9\*4+2\*1+6\*5+5\*9+3\*2+5\*6+8\*5+9\*3+7\*5+9\*8

= 3+5+36+2+30+45+6+20+40+27+35+72 = 331

r3 = 5\*3+9\*1+2\*4+6\*1+5\*5+3\*9+5\*2+8\*6+9\*5+7\*3+9\*5

= 15+9+8+6+25+27+10+48+45+21+45 = 259

2.2



1. Matching Technique

1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Distortion matrix for ONE** | | | | | |
| **4** | 2 | 1 | 1 | 1 | 0 |
| **2** | 0 | 3 | 3 | 3 | 2 |
| **6** | 4 | 1 | 1 | 1 | 2 |
| **4** | 2 | 1 | 1 | 1 | 0 |
| **1** | 1 | 4 | 4 | 4 | 3 |
|  | **2** | **5** | **5** | **5** | **4** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Accumulated matrix for ONE** | | | | | |
| **4** | 9 | 7 | 7 | 7 | 6 |
| **2** | 7 | 6 | 6 | 6 | 6 |
| **6** | 7 | 3 | 3 | 4 | 6 |
| **4** | 3 | 2 | 3 | 4 | 4 |
| **1** | 1 | 5 | 9 | 13 | 16 |
|  | **2** | **5** | **5** | **5** | **4** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Distortion matrix for TWO** | | | | | |
| **7** | 5 | 2 | 2 | 2 | 3 |
| **0** | 2 | 5 | 5 | 5 | 4 |
| **1** | 1 | 4 | 4 | 4 | 3 |
| **7** | 5 | 2 | 2 | 2 | 3 |
| **2** | 0 | 3 | 3 | 3 | 2 |
|  | **2** | **5** | **5** | **5** | **4** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Accumulated matrix for TWO** | | | | | |
| **7** | 13 | 10 | 12 | 13 | 14 |
| **0** | 8 | 11 | 11 | 11 | 12 |
| **1** | 6 | 6 | 6 | 8 | 9 |
| **7** | 5 | 2 | 4 | 6 | 9 |
| **2** | 0 | 3 | 6 | 9 | 11 |
|  | **2** | **5** | **5** | **5** | **4** |

Total minimum accumulated distortion for “ONE” is 4, and is less than that of “TWO”, which is 9, so the unknown input is “ONE”.