You must turn in your code as well as output files. Please generate a report that contains the code and output in a single readable format.

Visit the book website to download companion software, including all the example problems.


1. (KLT 9.10)

Solution

The energy estimation is plotted in Fig. 1. Notice these values have been normalized for visualization. The frame length is 256 and the window lengths are $\alpha_s = 1/16$ and $\alpha_l = 1/128$. Please see the end for code.

![Figure 1: KLT 9.10](image)

2. (KLT 9.11)

Solution

The noise floor estimation is plotted in Fig. 2.

3. (KLT 9.12)

Solution

The voiced areas is plotted in Fig. 3.

4. (KLT 9.13)

Solution

(a) The different window sizes are required by the VAD algorithm to respond quickly to the start of speech ($E_s$) while providing loudness consistency during a word utterance ($E_l$).
(b) During the onset of speech, the output of $E_s$ is larger.

(c) During the offset of speech, the output of $E_l$ is larger and takes longer to decay.
Speech Code

```matlab
s = load('TIMIT.ASC');
x = s/max(s) + 0.02*randn(size(s));
scale = 8;
fs = 8000;
win = 2^scale;
alphas = 1/2^(scale-4);
alphal = 1/2^(scale-1);
betal = 5;

% fft params
N = round(length(x)/win);
delf = fs/win;
K1 = find(delf*(0:N-1)>=300,1);
K2 = find(delf*(0:N-1)>1000,1)-1;

VAD = zeros(1,N+1); Es = zeros(1,N+1); El = zeros(1,N+1);
Nf = zeros(1,N+1); En = zeros(1,N+1);
for f=1:N
    % get frame
    it = f*win;
    ib = (f-1)*win+1;
    xc = x(ib:it);

    % compute frame energy
    Xc = fft(xc);
    En(f+1) = sum(abs(Xc(K1:K2)).^2);

    % signal energy
    Es(f+1) = (1-alphas)*Es(f) + alphas*En(f+1);
    El(f+1) = (1-alphal)*El(f) + alphal*En(f+1);

    % noise floor
    if(Nf(f)<Es(f+1))
        Nf(f+1) = (1-alphal)*Nf(f)+alphal*En(f+1);
    else
        Nf(f+1) = (1-alphas)*Nf(f)+alphas*En(f+1);
    end

    % threshold
    Tr(f+1) = Nf(f+1)/(1-alphal) + betal;

    % VAD
    if(En(f+1) > Tr(f+1)), VAD(f+1) = 1; end
end
```
The results of the VAD algorithm can be plotted together (with normalization for visualization) in Fig. 4.

Figure 4: Voice Activity Detection (VAD) Overlay