HW2 LMS Filter and DWT Filter design

VLSI DSP HW2

Shun-Linag Yeh, NCHU Lab612

3/19/2023

INDEX

Adaptive FIR Low pass filter

Problem

Q1. LMS filter design

For a least mean square (LMS) adaptive filter, assume the filter is of the form finite impulse response (FIR) and 15-tap long (i.e., with 15 coefficients b_0~b_{14} for x(n)~x(n-14)). Given an input signal consisting of 2 frequency components

$$s(n) = \sin(2\pi*n/12)+\cos(2\pi*n/4)$$

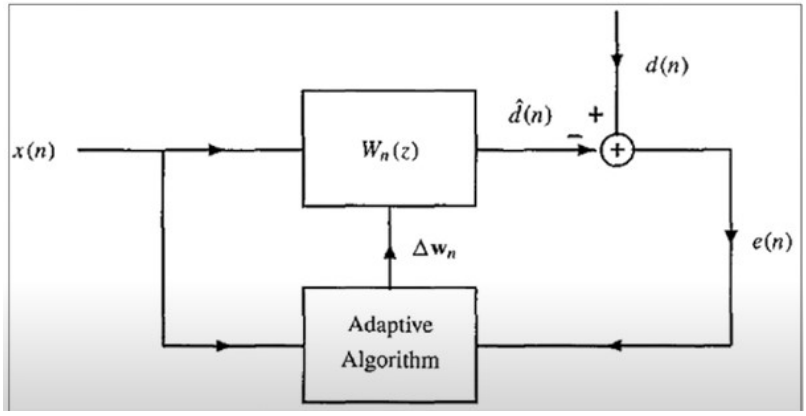
develop an adaptive low pass filter design

Set the target as a low pass filter to remove the high frequency component cos(2 π*n/4) and use sin(2π*n/12) as the desired (or training) signal for LMS adaptation. Assume the step size μ is 10^{-2}.

- write a Matlab code to simulate the LMS based adaptive filtering. Calculate the RMS (root mean square) value of the latest 16 prediction errors
  (i.e., r = sqrt((e^2(n) + e^2(n-1) + ….+ e^2(n-15))/16) and the adaptation is considered being converged if this value is less than 10% of RMS (root mean square) value of the desired signal, which equals 0.1/sqrt(2).
- Show the plot of "r" versus "n" and indicate when the filter converges, i.e. how many training samples are required
- Show the plot of filter coefficients b_i(n), for i = 0~14, versus "n" and see if the values of filter coefficients remain mostly unchanged after convergence
- Apply a 64-point FFT to the impulse response of the converged filter and verify the filter is indeed a low pass one. Note that the input vector to the 64-point FFT is
  (b_0, b_1, ….., b_{14}, 0,0,…..,0) with 49 trailing zeros.
- Change the step size μ to 10^{-4} and see how the behavior of the adaptive filter changes.
- Conduct simulation with a sufficiently large number of samples to see how small the value of "r" can be (the convergence bias)

Derivation steps

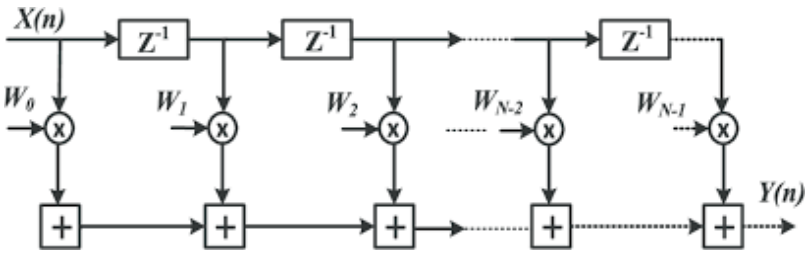
Adaptive Filter specification



1. x(n) is the input signal, wn(z) is the adaptive filter block with coefficients of wn.
2. d_hat(n) is the generated system response and d(n) is the desired signal.
3. e(n) is the error between d_hat(n) and d(n)
4. The adaptive algorithm block determines which kind of policy we should use to find the suitable filter coefficients. In this HW, LMS algorithm is chosen.

The adaptive FIR filter

$$\hat{d}(n) = \sum_{k=0}^{p} \omega_n(k) x(n-k) = \boldsymbol{w}_n^T \boldsymbol{X}(n)$$



- The desired output is genereated through the p-tap FIR filter design, where wn is the coefficients that gets updated on the fly.

Error function

$$e(n) = d(n) - \hat{d}(n)$$
$$= d(n) - w_n^T \boldsymbol{X}(n)$$

$$E\{e(n)x^*(n-k)\} = 0 \;;\; k = 0,1,\dots,p$$

- Error function simply is the difference between the desired signal and the generated system response.
- Ultimate goal is to minimize the autocorrelation between error vector and input signal.

LMS algorithm

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu e(n)\boldsymbol{X}^*(n)$$

$$\omega_{n+1} = \omega_n(k) + \mu e(n)\boldsymbol{X}^*(n-k)$$

- mu is the step sizes for the algorithm, which governs the variability of the coefficients in each iteration.
- e(n)X*(n) is the factor of auto-correlation between the input signal and the error function.

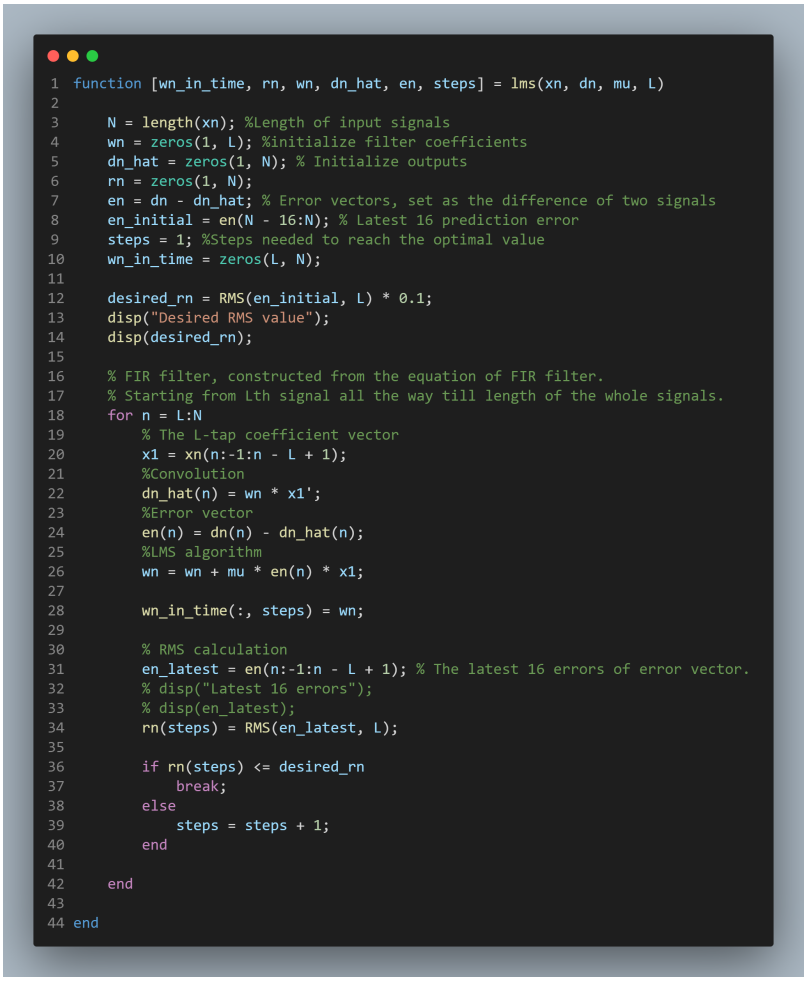
RMS(Root mean square)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$

- Root mean square used to find the norm of the error vector, we hope that this value be as small as possible s.t. the system is converged.

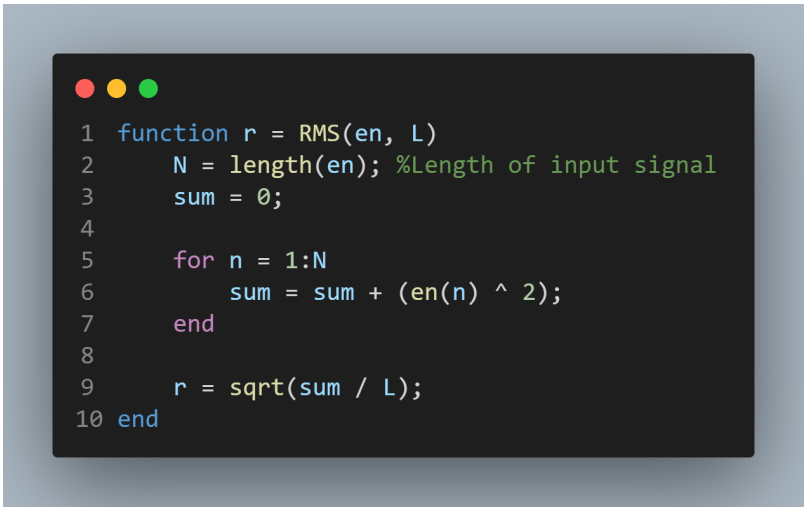
Code

Adaptive Filters



```m
function [wn_in_time, rn, wn, dn_hat, en, steps] = lms(xn, dn, mu, L)

    N = length(xn); %Length of input signals
    wn = zeros(1, L); %initialize filter coefficients
    dn_hat = zeros(1, N); % Initialize outputs
    rn = zeros(1, N);
    en = dn - dn_hat; % Error vectors, set as the difference of two signals
    en_initial = en(N - 16:N); % Latest 16 prediction error
    steps = 1; %Steps needed to reach the optimal value
    wn_in_time = zeros(L, N);

    desired_rn = RMS(en_initial, L) * 0.1;
    disp("Desired RMS value");
    disp(desired_rn);

    % FIR filter, constructed from the equation of FIR filter.
    % Starting from Lth signal all the way till length of the whole signals.
    for n = L:N
        % The L-tap coefficient vector
        x1 = xn(n:-1:n - L + 1);
        %Convolution
        dn_hat(n) = wn * x1';
        %Error vector
        en(n) = dn(n) - dn_hat(n);
        %LMS algorithm
        wn = wn + mu * en(n) * x1;

        wn_in_time(:, steps) = wn;

        % RMS calculation
        en_latest = en(n:-1:n - L + 1); % The latest 16 errors of error vector.
        % disp("Latest 16 errors");
        % disp(en_latest);
        rn(steps) = RMS(en_latest, L);

        if rn(steps) <= desired_rn
            break;
        else
            steps = steps + 1;
        end

    end

end
```

RMS



```m
function r = RMS(en, L)
    N = length(en); %Length of input signal
    sum = 0;

    for n = 1:N
        sum = sum + (en(n) ^ 2);
    end

    r = sqrt(sum / L);
end
```

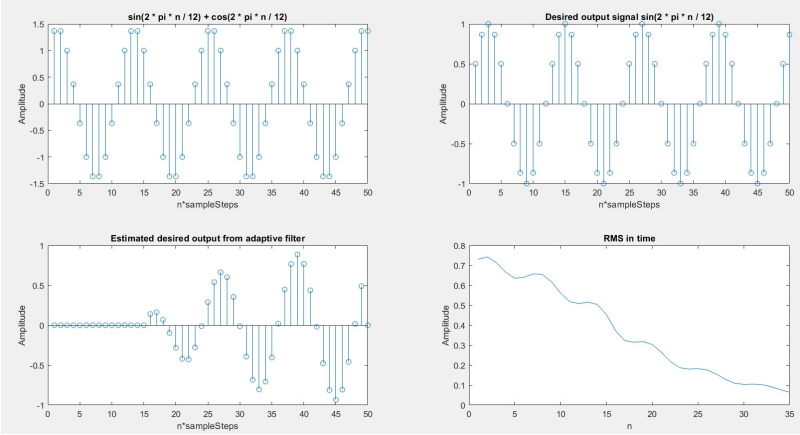
Main drivers

```m
clear all;

% Note sequence starting from 1~60
Sample_size = 3000;
n = 1:Sample_size;
L = 15;
mu = 0.0001;

n_s = 1:100;

%Input signal xn(n)
xn = sin(2 * pi * n / 12) + cos(2 * pi * n / 12);
dn = sin(2 * pi * n / 12);

%plot of original signal for 50 equally sampled value
figure(1);
sampleSteps = 25;
coefficient_steps = 50;
coefficient_spaced = 1:coefficient_steps:Sample_size;
n_spaced = 1:sampleSteps:Sample_size;
subplot(2, 2, 1);
stem(xn(n_spaced));
title('sin(2 * pi * n / 12) + cos(2 * pi * n / 12)'); xlabel('n*sampleSteps'); ylabel('Amplitude');

subplot(2, 2, 2);
stem(dn(n_spaced));
title('Desired output signal sin(2 * pi * n / 12)'); xlabel('n*sampleSteps'); ylabel('Amplitude');

% Adaptive Filter, Look for the minimum samples to reach 10% of LMS
[wn_in_time, rn, wn, dn_hat, en, steps] = lms(xn, dn, mu, L);

% Plot of dn_hat
subplot(2, 2, 3);
stem(dn_hat(n_spaced));
title('Estimated desired output from adaptive filter'); xlabel('n*sampleSteps'); ylabel('Amplitude');

% Plot of rn
subplot(2, 2, 4);
plot(rn(n_spaced));
title('RMS in time'); xlabel('n*sampleSteps'); ylabel('Amplitude');

disp("RMS final value");
disp(rn(steps));
```

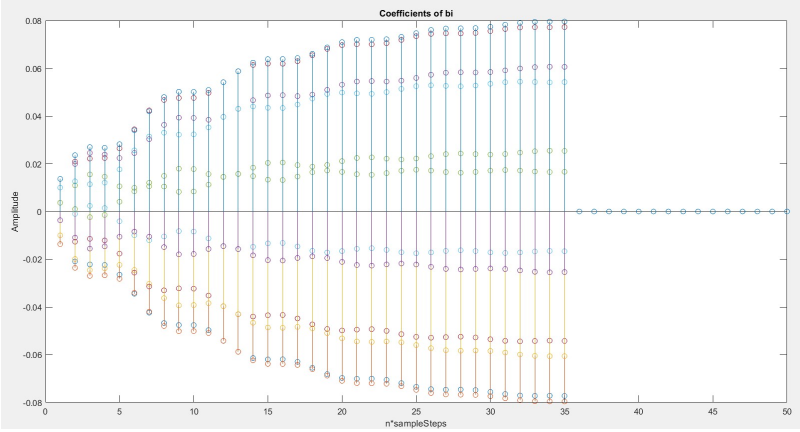
```m
disp("Total Steps needed to reach 10% of RMS");
disp(steps);
% Plot of Coefficients v.s. steps
figure(2);
wn_in_time = wn_in_time';
stem(wn_in_time(coefficient_spaced, :));
title('Coefficients of bi'); xlabel('n*coefficient steps'); ylabel('Amplitude');

% FFT for the impulse response of converged filters.
N = 64;
wn_padded = zeros(1, N);
wn_padded(1:L) = wn;

figure(3);

Y = fft(wn_padded, N);
stem(Y); % Note must use stem.
```

Results

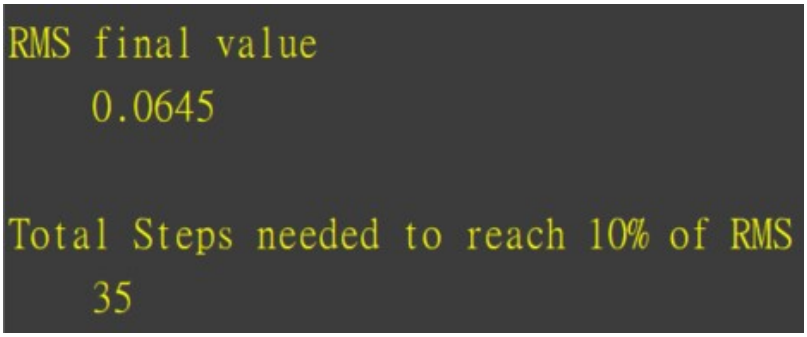
Adaptive Filter Response n=100, mu = 0.01, sampleSteps = 1 and RMS over time



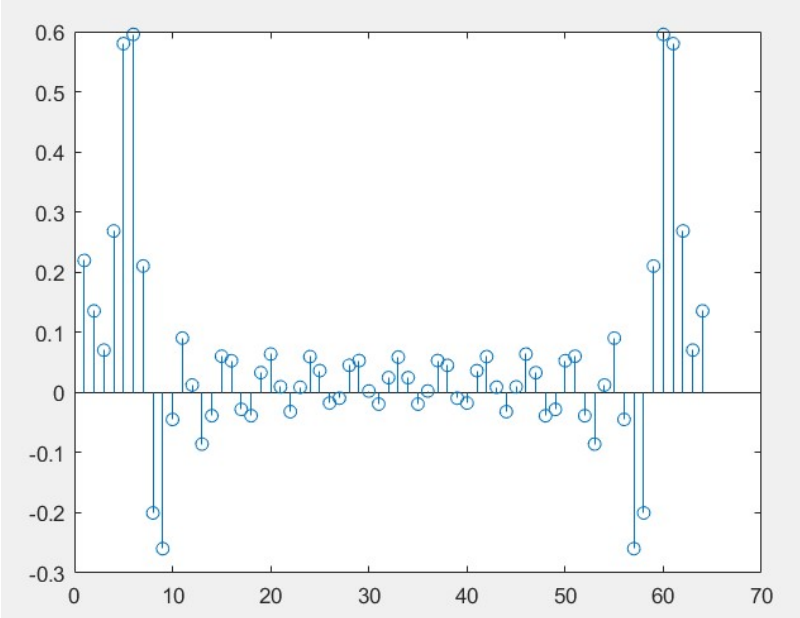
Filter Coefficients over time



Converged steps

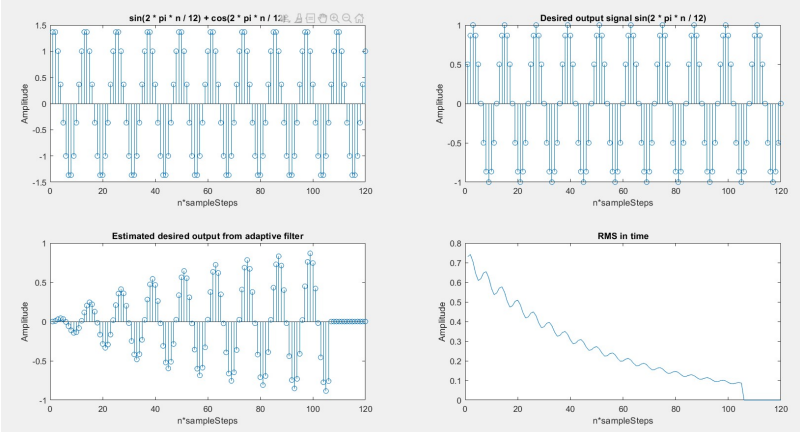


64-point FFT spectrum

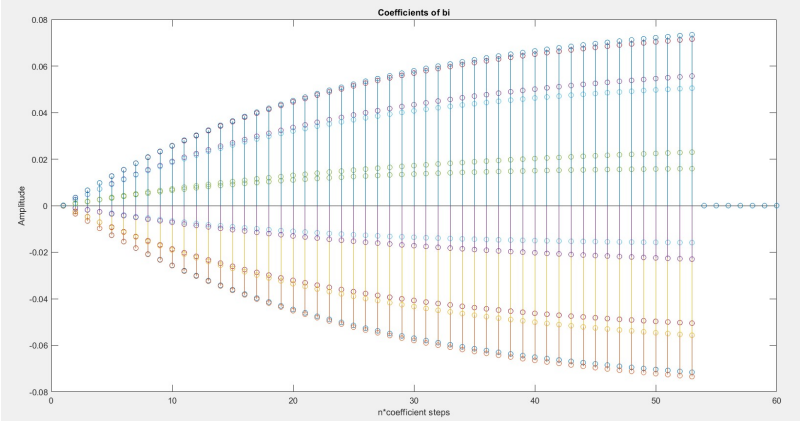


- The response is indeed a Low-Pass filter response.

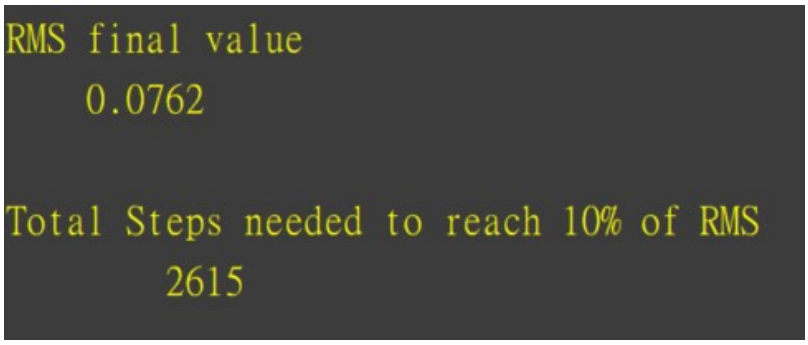
Adaptive Filter Response n=3000, mu = 0.0001, samepleSteps = 25 and RMS over time



Filter Coefficients over time

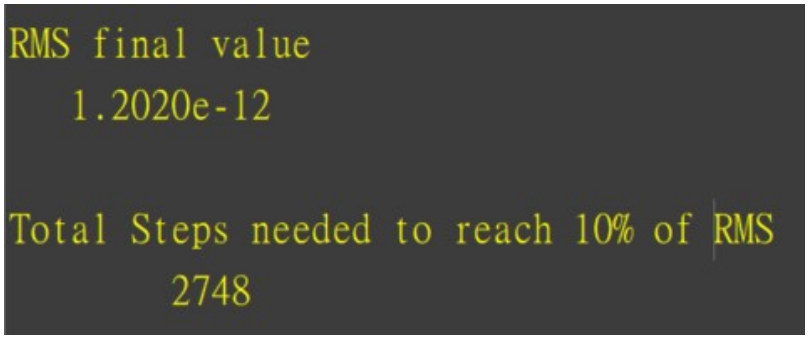


Converged steps



- Notice the changing of the RMS value responses more drmatically due to smaller step difference. Also it takes longer and more sample points for it to converge.

RMS with large sample size n = 10000



- The convergence bias is found, the filter cannot converge any further, it keeps on oscillating between within the convergence bias.

Note

- Due to the fact that sample sizes are large for smaller mu, plotting all of the signals makes analysis hard, thus samples_steps is defined s.t. only a certain multiple of signal sample_steps are selected for plotting.
- The latest 16 prediction errors should be selected for caculation, selecting more than that might yield the wrong results, and the filter would never converge.

Discrete Wavelet transform

Problem

For a discrete wavelet transform (DWT) adopting (9/7) filters, i.e. the low pass filter $h(i)$ is 9-taped and the high pass filter $g(i)$ is 7-taped. Both filters are liner phased and have symmetric coefficients. The filter coefficients are given in Table 1. For a corresponding inverse discrete wavelet transform, the low pass filter $q(i)$ is 7-taped and the high pass filter $p(i)$ is 9-taped. The filter coefficients are given in Table 2.

Table 1. Analysis filter coefficients for the floating point 9/7 filter

Analysis Filter Coefficients		
i	Lowpass Filter h_i	Highpass Filter g_i
0	0.852698679009	-0.788485616406
±1	0.377402855613	0.418092273222
±2	-0.110624404418	0.040689417609
±3	-0.023849465020	-0.064538882629
±4	0.037828455507	

Note: the high pass and low pass filter notations here are opposite to those in the lecture note

Table 2. Synthesis filter coefficients for the floating point 9/7 filter

Synthesis Filter Coefficients		
i	Low pass Filter q_i	High pass Filter p_i
0	0.788485616406	-0.852698679009
±1	0.418092273222	0.377402855613
±2	-0.040689417609	0.110624404418
±3	-0.064538882629	-0.023849465020
±4		-0.037828455507

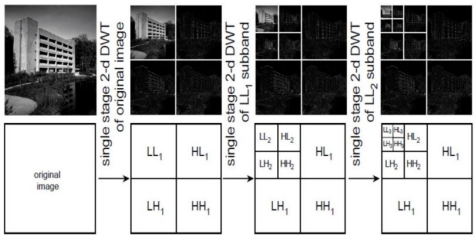
- Goal is to build a DWT filter by adopting the (9/7) filters composed of high pass and low pass components.

Problem a

a) For a 512×512 gray scale image (will be provided along with the homework assignment), please conduct a 2-D 3-level DWT transform (as shown in Figure 1) and show the transformed result. Then conduct a 2-D 3-level IDWT to convert it back. Please compare if the reconstructed image (after IDWT) is same as the original image by calculating its PSNR value.

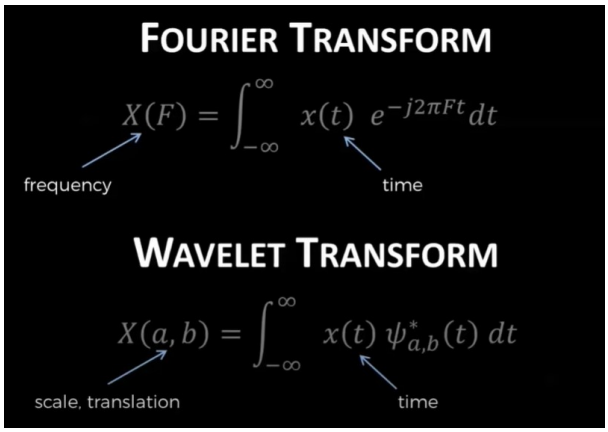
Problem b

b) By setting all three level 1 sub-bands HL1, LH1 and HH1 coefficients to zeros and perform IDWT. See how the reconstructed image is different from the original one and calculate its PSNR value.

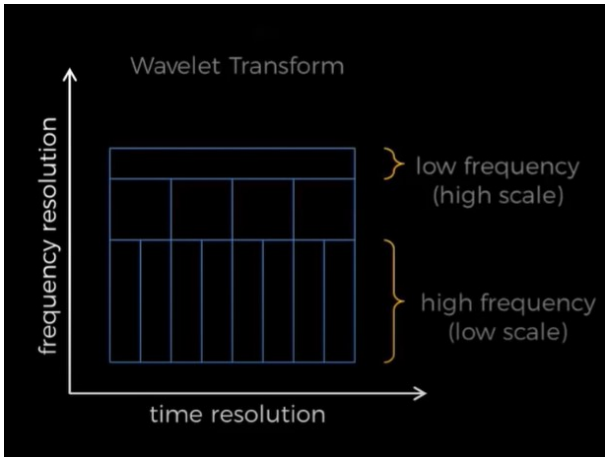


Derivation steps

Wavelet transform

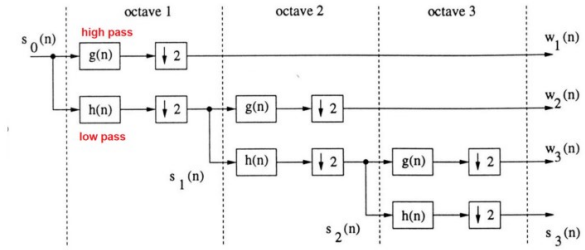


- For a fourier transform, the orthnormal basis is selected as e^j2piFt as analyzing function, however for the wavelet has orthonormal basis of wavelet analyzing fuction phi(t).
- Fourier transform ouputs frequency, yet wavelet outputs a translation and scaled autocorrelation of the input.



- Wavelet would has better time localization for high frequency yet worse time localization for low frequency.
- Wavelet is used to tackle with problem of time and frequency resolution since you cannot have both great time resolution and frequency resolution at the same time. Wavelet is the balance between time and frequency resolution.
- It is extensively used for compressing images.

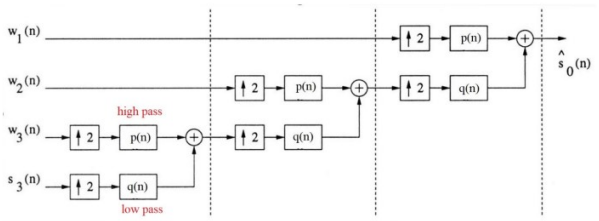
3-level DWT transform



(a) One-dimensional 3-level DWT transform

- The 3 level DWT structure has 3 octaves, where after passing each octaves the image becomes 2 times smaller than the original image due to down-Sampler.
- When doing downsampling, low pass filter sample the odd number output data, while high pass filters keep the even number output data.

3-level IDWT transform



(b) One-dimensional 3-level IDWT transform

- The 3 level IDWT structure has 3 octaves also, where after passing each octaves the image becomes 2 times larger than the original image due to up-Sampler. During the upSampling process, action should be taken to combat the loss of information when doing up Sampling.

Symmetric extension scheme

Figure 3. Symmetric extension scheme for boundary pixels

- Due to the boundary condition when doing filtering, one must extend the original signal length s.t. the boundary would not get convolved into 0 values.

PSNR(Peak signal to noise ratio) and MSE(Mean square error)

$$\mathrm{MSE} = \frac{1}{\mathrm{M}\cdot\mathrm{N}}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(I(i,j)-\hat{I}(i,j)\right)^2$$

$$PSNR = 10\log_{10}\left(\frac{MAXI^2}{\mathrm{MSE}}\right)$$

- Used to check whether the reconstructed image is close to the original image or not.

Code

FIR Filter(Symmetric extended)

```
function yn = filterSystem(xn, wn, N)
    % N signals
    % xn is the input signals
    % wn is the coefficient vector of filterSystem
    yn = zeros(1, N);
    L = size(wn);
    L = L(2);

    %Symmetric extension
    xn = [xn(2:L), xn, xn(N - L:N - 1)];

    for n = L:N + L - 1
        x = xn(n:-1:n - L + 1);
        yn(n - L + 1) = wn * x';
    end

end
```

g(n) High pass filter and h(n) low pass filter

```
function filtered_img = gn_HPF(raw_img)
    wn = [-0.064538882629 0.040689417609 0.418092273222 ...
              -0.788485616406 ...
              0.418092273222 0.040689417609 -0.064538882629];

    [h, w] = size(raw_img);
    filtered_img = raw_img;

    for i = 1:h
        row_img = raw_img(i, :)
        filtered_row = filterSystem(row_img, wn, w);
        filtered_img(i, :) = filtered_row;
    end

end

function filtered_img = hn_LPF(raw_img)
    wn = [0.037828455507 -0.023849465020 -0.110624404418 0.377402855613 ...
              0.852698679009 ...
              0.377402855613 -0.110624404418 -0.023849465020 0.037828455507];

    [h, w] = size(raw_img);
    filtered_img = raw_img;

    for i = 1:w
        row_img = raw_img(i, :)
        filtered_row = filterSystem(row_img, wn, h);
        filtered_img(i, :) = filtered_row;
    end

end
```

p(n) High pass filter and q(n) low pass filter

```
function filtered_img = qn_LPF(raw_img)
    wn = [-0.064538882629 -0.040689417609 0.418092273222 ...
              0.788485616406 ...
              0.418092273222 -0.040689417609 -0.064538882629];

    [h, w] = size(raw_img);
    filtered_img = raw_img;

    for i = 1:w
        row_img = raw_img(i, :)
        filtered_row = filterSystem(row_img, wn, h);
        filtered_img(i, :) = filtered_row;
    end

end

function filtered_img = pn_HPF(raw_img)
    wn = [-0.037828455507 -0.023849465020 0.110624404418 0.377402855613 ...
              -0.852698679009 ...
              0.377402855613 0.110624404418 -0.023849465020 -0.037828455507];

    [h, w] = size(raw_img);
    filtered_img = raw_img;

    for i = 1:w
        row_img = raw_img(i, :)
        filtered_row = filterSystem(row_img, wn, h);
        filtered_img(i, :) = filtered_row;
    end

end
```

Down Sampler & UpSampler

```
function downSampledimg = downSampler(img, stride, odd)
    [h, w] = size(img);
    % downSampledimg = img;
    downSampledimg = zeros(h);

    if odd == 0
        %even for HPF
        downSampledimg(1:w / 2, 1:h / 2) = img(2:stride:w, 2:stride:h);
    else
        %odd for LPF
        downSampledimg(1:w / 2, 1:h / 2) = img(1:stride:w, 1:stride:h);
    end

end
```

```
function upSampledimg = upSampler(img, stride, n)
    % n means nth octave
    [h, w] = size(img);
    partition = 2 ^ n;
    % upSampledimg   = img;
    upSampledimg = zeros(h);

    % Simply replicates two rows
    upSampledimg(2:stride:w / (partition / 2), 2:stride:h / (partition / 2)) = img(1:w / partition, 1:h / partition);
    upSampledimg(1:stride:w / (partition / 2), 1:stride:h / (partition / 2)) = img(1:w / partition, 1:h / partition);

    % Closest neighboring algorithm
    % upSampledimg(2:stride:w / (partition / 2), 2:stride:h / (partition / 2)) = img(1:w / partition, 1:h / partition);
    % upSampledimg(1:stride:w / (partition / 2), 1:stride:h / (partition / 2)) = img(1:w / partition, 1:h / partition);
    % upSampledimg(1:stride:w / (partition / 2), 2:stride:h / (partition / 2)) = img(1:w / partition, 1:h / partition);
    % upSampledimg(2:stride:w / (partition / 2), 1:stride:h / (partition / 2)) = img(1:w / partition, 1:h / partition);

end
```

PSNR

```m
function [psnr, difference] = PSNR(img, filtered_img)
    [h, w] = size(img);
    MAXI = 255;
    difference = (img - filtered_img) .^ 2;
    MSE = sum(difference, "all") / (h * w);

    psnr = 10 * log10(MAXI ^ 2 / MSE);
end
```

Main driver

```m
%=================================================================
%  RD image
%=================================================================
img = imread('image.bmp');
img = double(img);
[h, w] = size(img);
n = 1;
stride = 2;

filtered_img = img;

figure(1);
imshow(img, []);
title('Original image');
```

```m
%=================================================================
%  DWT
%=================================================================
% Octave 1
gn_filtered_img = gn_HPF(img);
w1n = downSampler(gn_filtered_img, stride, 0);

hn_filtered_img = hn_LPF(img);
s1n = downSampler(hn_filtered_img, stride, 1);

% Octave 2
gn_filtered_img = gn_HPF(s1n);
w2n = downSampler(gn_filtered_img, stride, 0);

hn_filtered_img = hn_LPF(s1n);
s2n = downSampler(hn_filtered_img, stride, 1);

% Octave 3
gn_filtered_img = gn_HPF(s2n);
w3n = downSampler(gn_filtered_img, stride, 0);

hn_filtered_img = hn_LPF(s2n);
s3n = downSampler(hn_filtered_img, stride, 1);
```

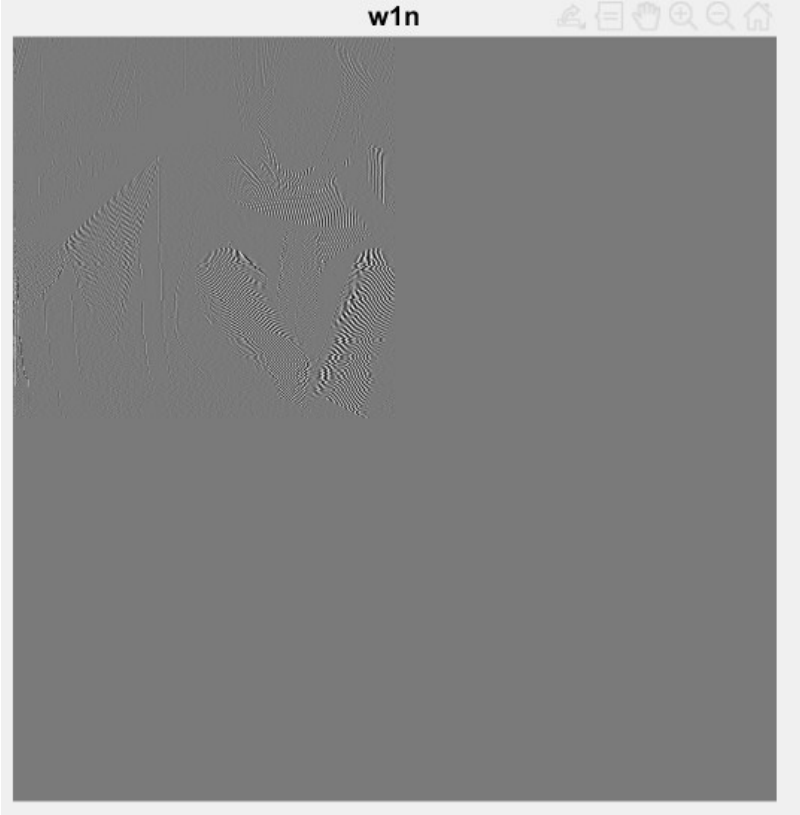
```
%=================================================================
%  IDWT
%=================================================================
% Octave 1
w3n_ = upSampler(w3n, stride, 3);
pn_filtered_img = pn_HPF(w3n_);

s3n_ = upSampler(s3n, stride, 3);
qn_filtered_img = qn_LPF(s3n_);

s2_hatn = pn_filtered_img + qn_filtered_img;

% Octave 2
w2n_ = upSampler(w2n, stride, 2);
pn_filtered_img = pn_HPF(w2n_);

s2_hatn_ = upSampler(s2_hatn, stride, 2);
qn_filtered_img = qn_LPF(s2_hatn_);

s1_hatn = pn_filtered_img + qn_filtered_img;

% Octave 3
w1n_ = upSampler(w1n, stride, 1);
pn_filtered_img = pn_HPF(w1n_);

s1_hatn_ = upSampler(s1_hatn, stride, 1);
qn_filtered_img = qn_LPF(s1_hatn_);

s0_hatn = pn_filtered_img + qn_filtered_img;

% Quantization
s0_hatn_ = ceil(s0_hatn);
s0_hatn_(s0_hatn_ > 255) = 255;
s0_hatn_(s0_hatn_ < -255) = -255;
```

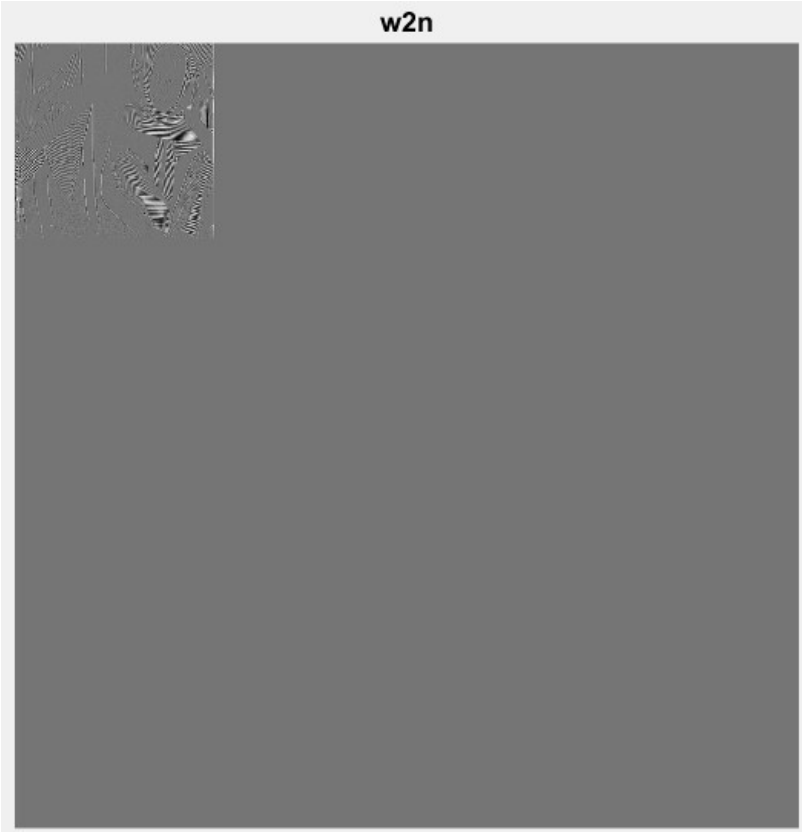
```
%=================================================================
%  Plots
%=================================================================
%=====================================
%  DWT
%=====================================
%==================
%  HPF
%==================
figure(2);
title('High pass filtered image');
imshow(w1n, []);
title('w1n');

figure(3);
imshow(w2n, []);
title('w2n');

figure(6);
imshow(w3n, []);
title('w3n');

%==================
%  LPF
%==================
figure(7);
imshow(s1n, []);
title('s1n');

figure(6);
imshow(s2n, []);
title('s2n');

figure(7);
imshow(s3n, []);
title('s3n');


%=====================================
%  IDWT
%=====================================

figure(8);
imshow(s0_hatn_, []);
title('s0 hatn');
```

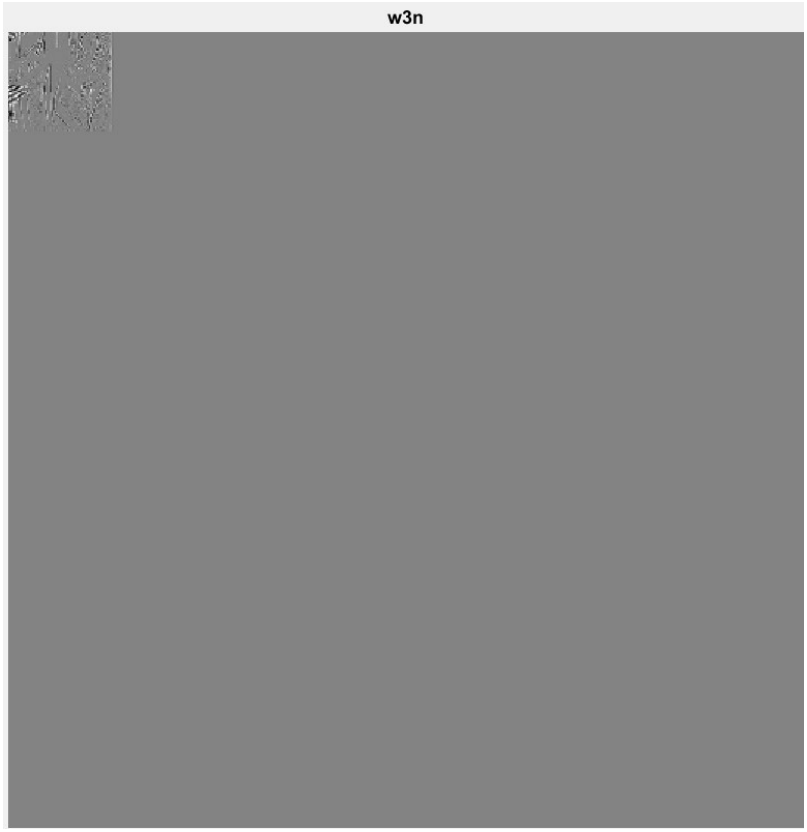
```
%=================================================================
%  PSNR
%=================================================================
disp("PSNR:");
[psnr, difference] = PSNR(img, s0_hatn_);
fprintf('%.2f db\n', psnr);
```

Result

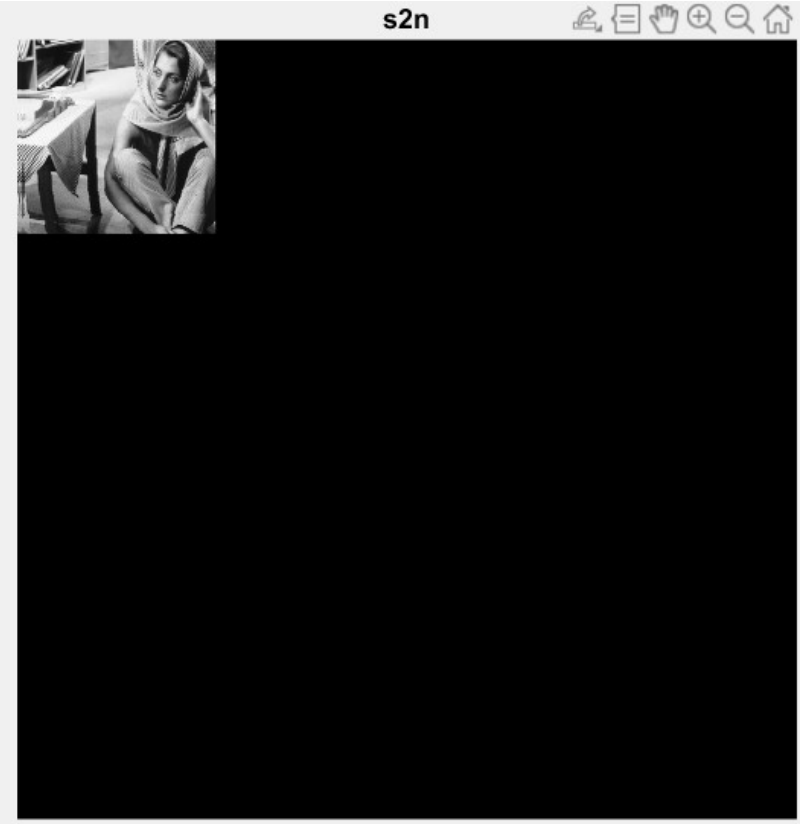
DWT w1n,w2n,w3n

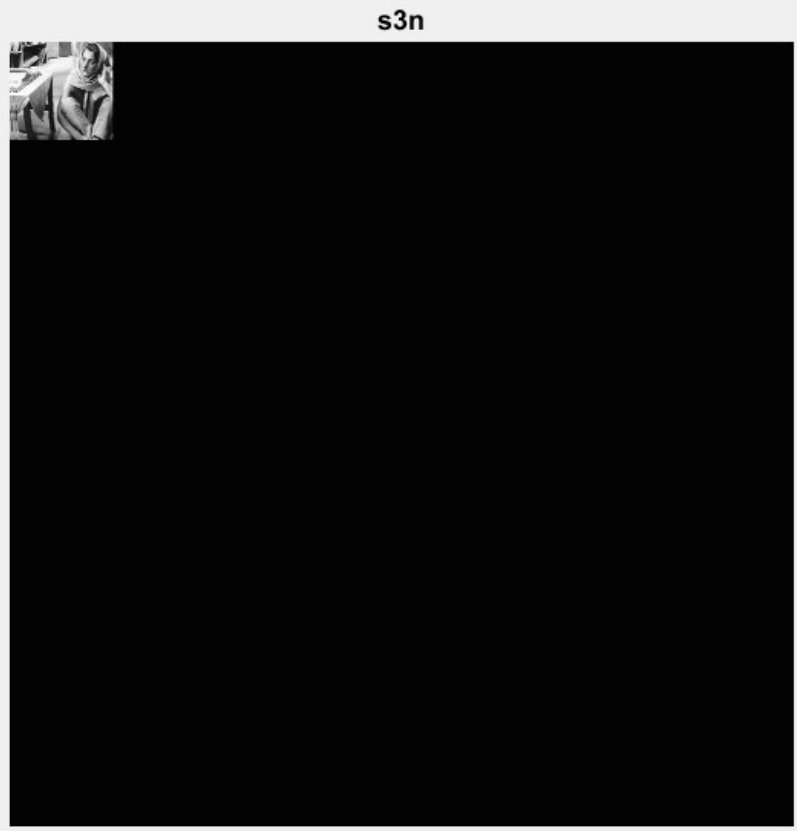






DWT s1n,s2n,s3n


s2n


s3n

s0 hatn Restored image

PSNR

- Note the PSNR is terrible, yet we can still see the image being briefly restored, I suppose more filtering should be performed on the image to fully restore the whole image.

Note

1. The result of the Wavelet transform can briefly restore the image, but I am not able to fully restore the whole resolution of the image. Perhaps some action should be taken to combat the loss of data during the upSampling process.
2. PSNR rate needs to be higher than 50db to achieve the Lossless filtering, I suppose the problem occurs during the upSampling process, where better algorithm should be adopted s.t. the image can restore its resolution.

References

[1] Advanced Digital Signal Processing, Adaptive Filters by Prof.Vaibhav Pandit

[2] Advanced Digital Signal Processing, LMS Algorithm by Prof.Vaibhav Pandit

[3] MIT RES.6-008 Digital Signal Processing,Lec 17, 1975 by Alan Oppenheim

[4] EE123 Digital Signal Processing, SP'16 L12 - Discrete Wavelet Transform

[5] Easy Introduction to Wavelets, by Simon Xu

[6] VLSI Digital Signal processing systems Design and Implementation, p25~28 by Parhi