Complex Wavelet Transform: application to denoising

Ioana ADAM

Thesis Advisors : Jean-Marc BOUCHER
                    Alexandru ISAR
Introduction

- Wavelet theory is a relatively new and powerful tool in signal and image processing

- Applications:
  - compression
  - classification/segmentation
  - denoising

Complex Wavelet Transform: application to denoising
Continuous Wavelet Transform (CWT) vs. Fourier Transform (FT)

CWT:

\[ Wx(u, s) = \langle x, \psi_{u,s} \rangle = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi^*(\frac{t-u}{s}) \, dt \] (1)

FT:

\[ X(\omega) = \int_{-\infty}^{\infty} x(t) \cdot e^{-j\omega t} \, dt, \] (2)
Discrete Wavelet Transform (DWT)

Decomposition tree:

Reconstruction tree:

Advantages:
- Sparsity of coefficients

Disadvantages:
- Shift-sensitivity
- Poor directional selectivity in 2D
Two-dimensional DWT (2D DWT)

original image

One-level 2D DWT decomposition scheme

2-level 2D DWT’s coefficients
Wavelet Packet Transform (WPT)

One-dimensional Wavelet Packet Decomposition

Advantages:
- Increased directional selectivity

Disadvantages:
- Increased computation time
Undecimated Discrete Wavelet Transform (UDWT)

Three-Level UDWT Decomposition Scheme

Relation between the filters corresponding to two consecutive levels of UDWT decomposition

Advantages:
- Shift-invariant

Disadvantages:
- High redundancy
- Reduced directional selectivity
Dual-Tree Complex Wavelet Transform (DT CWT)

Advantages:
- Quasi shift-invariant
- Good directional selectivity

Disadvantages:
- Filters from the 2nd branch can be only approximated
Analytic signal:

\[ x_a(t) = x + jH(x) \]  

where \( H(x) \) denotes the Hilbert transform of \( x \).

\[
\begin{align*}
DWT \{x(t)\} &= \langle x(t), \psi(t) \rangle, \\
ADWT \{x(t)\} &= \langle x(t), \psi_a(t) \rangle \\
&= \langle x(t), \psi(t) + jH \{\psi(t)\} \rangle \\
&= \langle x(t) + jH \{x(t)\}, \psi(t) \rangle \\
&= \langle x_a(t), \psi(t) \rangle = DWT \{x_a(t)\}.
\end{align*}
\]
Degree of shift-invariance:

\[ \text{Grad} = 1 - \frac{d}{m}, \]  

(5)

where \( d \) - standard deviation and \( m \) - mean of the sequences of energies of a certain type of coefficients corresponding to 16 shifts.
Shift-invariance of ADWT - qualitative results

(a) ADWT - new version
(b) Dual Tree CWT
(c) Real DWT
Shift-invariance of ADWT - quantitative results

<table>
<thead>
<tr>
<th></th>
<th>ADWT</th>
<th>CS (1,64)</th>
<th>CS (1,512)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daubechies, 10</td>
<td>0.8594</td>
<td>0.7551</td>
<td>0.7551</td>
</tr>
<tr>
<td>Scaling fn. level 3</td>
<td>0.9981</td>
<td>0.9965</td>
<td>0.9996</td>
</tr>
<tr>
<td>Wavelets level 3</td>
<td>0.9982</td>
<td>0.9968</td>
<td>0.9996</td>
</tr>
<tr>
<td>Wavelets level 1</td>
<td>0.9992</td>
<td>0.9985</td>
<td>0.9998</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WT</th>
<th>Degree of invariance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W. lev. 1</td>
</tr>
<tr>
<td>ADWT</td>
<td>0.9967</td>
</tr>
<tr>
<td>DT CWT</td>
<td>1.0000</td>
</tr>
<tr>
<td>DWT</td>
<td>0.9236</td>
</tr>
</tbody>
</table>
Hyperanalytic Wavelet Transform (HWT) 

Hypercomplex mother wavelet:

\[ \psi_h(x,y) = \psi(x,y) + iH_x\{\psi(x,y)\} + jH_y\{\psi(x,y)\} + \]
\[ + kH_x\{H_y\{\psi(x,y)\}\} \] (6)

where \(i^2 = j^2 = -k^2 = -1, ij = ji = k, jk = kj = -i, ki = ik = -j\)
and \(ijk = 1\).

The HWT of an image \(f(x,y)\) is:

\[ HWT\{f(x,y)\} = \langle f(x,y), \psi_h(x,y) \rangle = DWT\{f_h(x,y)\} \]
\[ = DWT\{f(x,y)\} + iDWT\{H_x\{f(x,y)\}\} + \]
\[ + jDWT\{H_y\{f(x,y)\}\} + \]
\[ + kDWT\{H_y\{H_x\{f(x,y)\}\}\} \]
\[ = \langle f_h(x,y), \psi(x,y) \rangle = DWT\{f_h(x,y)\}. \] (7)
HWT implementation with directional selectivity enhancement

Initial computations

Directional selectivity enhancement
**HWT directional selectivity - example**

<table>
<thead>
<tr>
<th>Input</th>
<th>Horizontal details</th>
<th>Diagonal details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DWT</th>
<th>HWT Real part</th>
<th>Imaginary part</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Complex Wavelet Transform: application to denoising
HWT’s shift invariance

Components of reconstructed 'disc' images

Input (256 x 256)

HWT

DT CWT

DWT

wavelets: level 1 level 2 level 3 level 4 level 4 scaling fn.

Complex Wavelet Transform: application to denoising
Hyperanalytic Wavelet Packet Transform (HWPT)
Complex Wavelet Transform: application to denoising

**Shift-invariance of HWPT**

<table>
<thead>
<tr>
<th></th>
<th>Best basis</th>
<th>E. DWPT</th>
<th>E. HWPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.6916</td>
<td>1.2390 · 10^5</td>
<td>1.0469 · 10^6</td>
</tr>
<tr>
<td>2</td>
<td>3.9403</td>
<td>5.9904 · 10^5</td>
<td>1.5056 · 10^6</td>
</tr>
<tr>
<td>3</td>
<td>3.9403</td>
<td>5.9904 · 10^5</td>
<td>1.5056 · 10^6</td>
</tr>
<tr>
<td>4</td>
<td>3.6916</td>
<td>1.2390 · 10^5</td>
<td>1.0469 · 10^6</td>
</tr>
<tr>
<td>5</td>
<td>3.6916</td>
<td>1.2390 · 10^5</td>
<td>1.0469 · 10^6</td>
</tr>
<tr>
<td>6</td>
<td>3.9403</td>
<td>5.9904 · 10^5</td>
<td>1.5056 · 10^6</td>
</tr>
<tr>
<td>7</td>
<td>3.9403</td>
<td>5.9904 · 10^5</td>
<td>1.5056 · 10^6</td>
</tr>
<tr>
<td>8</td>
<td>3.6916</td>
<td>1.2390 · 10^5</td>
<td>1.0469 · 10^6</td>
</tr>
</tbody>
</table>

\[ \text{Deg}_{2D-DWPT} = 0.3, \text{Deg}_{HWPT} = 0.81. \]
Complex Wavelet Transform: application to denoising
2D DWPT versus HWPT

Complex Wavelet Transform: application to denoising
Denoising

Additive noise: \( x = s + n \).

Can be performed:
- in the spatial domain,
- in the wavelet domain:
  - computation of the forward WT,
  - filtering of coefficients,
  - computation of the inverse WT.

Filters can be
- non-parametric
- parametric

Performance evaluation:

\[
PSNR = 10 \log_{10} \left( \frac{R^2}{MSE} \right)
\]  

(8)

where \( MSE = \frac{\sum_{m=1}^{M,N} [I_o(m,n) - I_i(m,n)]^2}{MN} \) and \( R = 255 \).
Parametric Denoising - Bayesian Approach

**Spatial domain:** \( x = s + n. \)

Bayes rule:

\[
p_{S|X}(s \mid x) = \frac{p_{X|S}(x \mid s) \cdot p_S(s)}{p_X(x)} = \frac{p_N(x - s) \cdot p_S(s)}{p_X(x)}. \tag{9}\]

MAP filter: the estimated value is the value that maximizes the a posteriori probability, \( p_{S|X}(s \mid x) \):\[
\hat{s} = \operatorname{argmax}_s (p_N(x - s) \cdot p_S(s)) \tag{10}\]

**Wavelet domain:** \( W_x^j = W_s^j + W_n^j. \)

\[
\hat{W}_s = \operatorname{argmax}_{W_s} (p_{W_N}(W_x - W_s) \cdot p_{W_S}(W_s)) \tag{11}\]

Complex Wavelet Transform: application to denoising
Zero-order Wiener Filter

Assumptions:

1. the signal and the noise are not correlated,
2. noise - Additive White Gaussian (AWG), \( N(0, \sigma_n^2) \),
3. signal - AWG, \( N(0, \sigma_s^2) \).

Spatial domain

\[
\hat{s}(i,j) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \cdot x(i,j).
\]

Wavelet domain

\[
\hat{W}_s(i,j) = \frac{\sigma_{W_s}^2}{\sigma_{W_s}^2 + \sigma_{W_n}^2} \cdot W_x(i,j).
\]
Zero-order Wiener Filter - simulation results

<table>
<thead>
<tr>
<th>$\sigma_n$</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNRi</td>
<td>28.16</td>
<td>24.64</td>
<td>22.15</td>
<td>20.20</td>
<td>18.62</td>
<td>17.28</td>
<td>14.17</td>
</tr>
<tr>
<td>PSNRo spatial</td>
<td>31.38</td>
<td>30.31</td>
<td>29.35</td>
<td>28.48</td>
<td>27.75</td>
<td>27.10</td>
<td>25.37</td>
</tr>
<tr>
<td>PSNRo hwt</td>
<td>34.81</td>
<td>32.80</td>
<td>31.37</td>
<td>30.21</td>
<td>29.29</td>
<td>28.47</td>
<td>26.61</td>
</tr>
</tbody>
</table>
Adaptive Soft-Thresholding Filter

Assumptions:

1. noise - AWG, \( p_{W_n}(W_n) = \frac{1}{\sqrt{2\pi\sigma_{W_n}}} \cdot e^{-\frac{w_n^2}{2\sigma_{W_n}^2}} \),

2. signal - Laplacian, \( p_{W_s}(W_s) = \frac{1}{\sqrt{2\sigma_{W_s}}} \cdot e^{-\frac{\sqrt{2}|W_s|}{\sigma_{W_s}}} \).

Input-output relation:

\[
\hat{W}_s = \text{sgn} (W_x) \left( |W_x| - \sqrt{2} \frac{\sigma_{W_n}^2}{\sigma_{W_s}} \right)_+. \tag{12}
\]

Standard deviation estimation

\[
\hat{\sigma}_{W_n}^2 = \frac{\text{median} (|W_x|)}{0.6745}, \ W_x \in \text{subband } HH^1.
\]

\[
\hat{\sigma}_{W_x}^2 = \frac{1}{M} \sum_{W_x \in N(k)} (W_x)^2, \ \hat{\sigma}_{W_s} = \sqrt{\left( \hat{\sigma}_{W_x}^2 - \hat{\sigma}_{W_n}^2 \right)_+}
\]
Adaptive Stf - simulation results

<table>
<thead>
<tr>
<th>$\sigma_n$</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNRI</td>
<td>28.15</td>
<td>24.65</td>
<td>22.15</td>
<td>20.21</td>
<td>18.62</td>
<td>17.28</td>
</tr>
<tr>
<td>PSNRo Adaptive Stf</td>
<td>34.81</td>
<td>32.97</td>
<td>31.64</td>
<td>30.55</td>
<td>29.67</td>
<td>28.98</td>
</tr>
<tr>
<td>PSNRo Wiener 0</td>
<td>34.81</td>
<td>32.80</td>
<td>31.37</td>
<td>30.21</td>
<td>29.29</td>
<td>28.47</td>
</tr>
</tbody>
</table>
Bishrink Filter

Assumptions:

1. **noise:** Gaussian,
   \[ p_{W_n}(W_n) = \frac{1}{2\pi\sigma^2_{W_n}} \cdot e^{-\frac{(W_{1n})^2+(W_{2n})^2}{2\sigma^2_{W_n}}}, \]

2. **signal:** 2-dimensional spherically contoured multivariate pdf,
   \[ p_{W_s}(W_s) = \frac{3}{2\pi\sigma^2_{W_s}} \cdot e^{-\frac{\sqrt{3}}{\sigma_{W_s}} \cdot \sqrt{(W_{1s})^2+(W_{2s})^2}}. \]

Input-output relation:

\[
\hat{W}_s^1 = \frac{\sqrt{(W_x^1)^2+(W_x^2)^2} - \frac{\sqrt{3}\sigma^2_{W_n}}{\sigma_{W_s}}}{\sqrt{(W_x^1)^2+(W_x^2)^2}} + \cdot W_x^1. \tag{13}
\]
The sensitivity of the estimation $\hat{W}_s^1$ with $\hat{\sigma}_W$ is:

$$S_{\hat{\sigma}_W} = \begin{cases} \frac{\sqrt{3}\sigma_W^2}{\sigma_W \sqrt{(W_1^1)^2 + (W_2^1)^2} - \sqrt{3}\sigma_W^2}, & \text{if } \sqrt{(W_1^1)^2 + (W_2^1)^2} > \sqrt{3}\sigma_W^2 \\ 0, & \text{otherwise}. \end{cases}$$

(14)
Bishrink - simulation results

<table>
<thead>
<tr>
<th>$\sigma_n$</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNRi</td>
<td>28.17</td>
<td>24.64</td>
<td>22.15</td>
<td>20.20</td>
<td>18.62</td>
<td>17.29</td>
</tr>
<tr>
<td>PSNRO</td>
<td>35.06</td>
<td>33.31</td>
<td>32.04</td>
<td>31.07</td>
<td>30.25</td>
<td>29.58</td>
</tr>
</tbody>
</table>
Speckle Filtering - general considerations

Speckle

- affects SAR images (Radar and Sonar);
- acts as a multiplicative noise: \( x = s \cdot n \).

Speckle reduction techniques

- Homomorphic filtering: \( \ln x = \ln s + \ln n \).

- Pixel-ratioing - based filtering:
  \[
  x = s \cdot n = s + s \cdot (n - 1) = s + n'.
  \]
Adaptive Soft-thresholding Filtering

Complex Wavelet Transform: application to denoising

<table>
<thead>
<tr>
<th>Input</th>
<th>Raw</th>
<th>HWT + Astf + cor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B9/7</td>
</tr>
<tr>
<td>1-look</td>
<td>12.1</td>
<td>25.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25.4</td>
</tr>
<tr>
<td>4-look</td>
<td>17.8</td>
<td>30.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29.9</td>
</tr>
<tr>
<td>16-look</td>
<td>23.7</td>
<td>32.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>33.2</td>
</tr>
</tbody>
</table>

(PSNR = 21.4 dB) (PSNR = 31.94 dB)
Bishrink - synthetic noise

(without correction)  (with correction)
**HWT + Bishrink vs. UDWT + GGPDF-based MAP**

<table>
<thead>
<tr>
<th>Input</th>
<th>Raw</th>
<th>HWT + Bish + cor</th>
<th>GGPDF-MAP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D4</td>
<td>B9/7</td>
</tr>
<tr>
<td>1-look</td>
<td>12.1</td>
<td>25.9</td>
<td>25.9</td>
</tr>
<tr>
<td>4-look</td>
<td>17.8</td>
<td>30.2</td>
<td>30.2</td>
</tr>
<tr>
<td>16-look</td>
<td>23.7</td>
<td>33.0</td>
<td>33.1</td>
</tr>
</tbody>
</table>

(HWT + bishrink) vs. (UDWT + GGPDF-based MAP)
Bishrink - test image

(Wiener 0)

(HWT + bishrink)

(Lee)
Bishrink - SONAR image

(ENL = 3.32)  (ENL = 155.04)
Complex Wavelet Transform: application to denoising

Bishrink - SAR image
Contributions I

1. A new one-dimensional wavelet transform called the Analytic Discrete Wavelet Transform (ADWT), introduced to overcome the shift-sensitivity of the Discrete Wavelet Transform (DWT).


3. The quasi shift-invariance of ADWT was measured in simulations both visually and through the values of the newly introduced measure, namely the degree of invariance.

4. The advantage of choosing ADWT over the DT CWT is the possibility to freely choose the mother wavelet from the wide range classically associated with the DWT.

5. A new two-dimensional wavelet transform called the HWT

6. The association of the HWT with the zero-order local Wiener filter for denoising purposes.
Contributions II

7 The association of the HWT with the adaptive soft-thresholding filter.

8 I have taken into account the inter-scale dependency of the wavelet coefficients and associated HWT with the bishrink filter in denoising.

9 Taking into account the drawbacks of the bishrink filter, it results that regions with different homogeneity degrees must be treated using different strategies.

10 For speckle reduction purposes, I have used a homomorphic filtering, replacing the additive denoising kernel with algorithms using the association HWT - adaptive soft-thresholding and HWT - bishrink, the results obtained being comparable with those in the literature.
in what concerns the HWT, it can be improved by finding a better implementation of the Hilbert transform, thus reducing the approximation errors the present implementation introduces;

a thorough research regarding the criteria to be used in choosing a mother wavelet suitable for the denoising of a particular image.

associating a wider range of estimators with the HWT, such as the BLS-GSM algorithm and others.
Publications I


Publications IV
