COMPARISON OF DISCRETE WAVELET TRANSFORM AND COMPLEX WAVELET TRANSFORM IN HYBRID SKELETONIZATION BASED ON CVANN

Murat CEYLAN¹
Osman Nuri UCAN²
Yüksel ÖZBAY³
Rachid JENNANE⁴
Gabriel AUFORT⁵
Claude Laurent BENHAMOU⁶

ABSTRACT

Curve and surface thinning are widely-used skeletonization techniques for modeling objects in 3 dimensions. In the case of disordered porous media analysis, however, neither is really efficient since the internal geometry of the object is usually composed of both rod and plate shapes. This paper concludes an application of discrete wavelet transform (WT) and complex wavelet transform (CWT) in image processing problem such as hybrid skeletonization of trabecular bone images. Hybrid skeleton combines 2D surfaces and 1D curve to represent respectively the plate-shaped and rod-shaped parts of the object. For hybrid skeletonization, two cascade structures are proposed. In these structures, features of images were extracted with discrete wavelet transform and complex wavelet transform. After that, obtained features were used as

¹ Selçuk University, Department of Electrical and Electronics Engineering.
² Istanbul Aydin University, Department of Electrical and Electronics Eng.
³ Konya Selçuk University, Eng. Faculty.Konya
⁴ Instiut PRISME / LESI, University of Orleans.
⁵ Equipe INSERM U658, Hospital of Orleans.
⁶ Orleans, France
inputs of complex-valued artificial neural network (CVANN) which is multilayered artificial neural networks with two dimensions (real and imaginary parts). Effects of the feature extraction methods are compared for ability of the hybrid skeletonization on a trabecular bone sample. Results show that the CWT succeeded to hybrid skeletonization with lower error rate than WT.

**Key Words:** Wavelet transform, complex wavelet transform, complex-valued artificial neural network, hybrid skeletonization.

1. Introduction

Skeletonizing an object often requires a compromise between minimum thinness, homotopic equivalence, geometry preservation and reversibility, which are usually incompatible. For example, algorithms based on distance transforms [1,2] are usually reversible but lose the connectivity information. 3D thinning [3,4] is another method based on iterative erosions of an object until only its skeleton is left. The most widely-used thinning variants are curve and surface thinning. Curve thinning techniques [5,6] generate skeletons composed of 1D curves. On the other hand, surface thinning techniques [7-11] generate 2D surfaces which better preserve the plates geometry, but do not thin rod shapes enough. Curve thinning and surface thinning are efficient respectively for one or the other, but not for both. For example, in the case of trabecular bone, it is clearly established that the structure is composed of both rod and plate parts. A lot of work has been carried out on the characterization of this porous medium, especially using skeletons.

In this paper, effects of WT and CWT in the skeletonizing hybrid-shaped media composed of rods and plates using CVANN were studied. It is expected that complex-valued artificial neural networks (CVANN) whose parameters (weights, threshold values, inputs and outputs) are all complex numbers, will have applications in fields dealing with signal processing [12-14] and image processing [15,16]. When using the existing method for real numbers, we must apply the method individually to their real and imaginary parts. On the other hand, complex-valued neural networks allow us to directly process data.
2. Curve Thinning Algorithm (CTA) and Surface Thinning Algorithm (STA)

The geometrical characterization of a porous medium using a thinning technique basically depends on the type of skeleton needed to model the material structure. Figure 1 shows the limitations of the curve and surface skeletons on a trabecular bone sample. In the curve case (b), plate-shaped zones are not efficiently described. In the surface case (c), the object shape is accurate but rod zones are not sufficiently thinned. Furthermore, the curve skeleton is sensitive to object surface irregularities, and generates many unexpected small segments.

![Figure 1: A trabecular bone sample (a), its curve skeleton (b) and its surface skeleton (c).](image)

2.1 Curve Thinning Algorithm (CTA)

The curve thinning algorithm used in this paper is based on the well established work of Morgenthaler [17]. This iterative process relies on the evaluation of the Betti numbers in each voxel’s neighbourhood. A voxel is said
to be removable if its deletion does not change these numbers. Voxels are removed sequentially if they match the deletion criterions, which ensures a perfect topology preservation. The process is repeated until only the curve skeleton remains. Figure 1.b shows the result of the curve thinning algorithm on the bone volume of figure 1.a. In this example, plate zones have fully disappeared, which confirms that curve thinning is not sufficient to describe the entire porous medium. However, it efficiently models the rod-shaped parts of the object.

2.2 Surface Thinning Algorithm (STA)

There are different types of surface skeletons, depending on the erosion conditions used in the case of thinning algorithms. Several surface thinning algorithms have been compared [9,11,18-20]. They all preserve the connectivity of the object. The important constraint in this work was to obtain a simple and relevant surface geometry in order to use the new skeleton as a basis for structural models. The MESPTA (Modified Extended Safe Point Thinning Algorithm) algorithm, [8,19,21] appeared to give the simplest surface geometry. The MESPTA consists in an iterative parallel topology-preserving process that transforms an object into its medial surface. The main advantage of this technique is that it generates a simple surface geometry, which is convenient to model the plate-shaped parts of an object. In the hybrid thinning technique, all plate voxels that match the safe point condition as explained in [19] are deleted at the end of an iteration. The process is iterated until no voxel is deleted. Figure 1.c shows the result of the MESPTA on the bone volume of figure 1.a, which has been thinned to simple 2D subsets.
2.3 Discrete Wavelet Transform

Wavelet is a powerful tool for representing nonlinearity [22]. A function \( f(x) \) can be represented by the superposition of daughters \( \psi_{a,b}(x) \) of a mother wavelet \( \psi(x) \). Where \( \psi_{a,b}(x) \) can be expressed as

\[
\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi \left( \frac{x-b}{a} \right)
\]  

(1)

\( a \in \mathbb{R} \) and \( b \in \mathbb{R} \) are, respectively, called dilation and translation parameters.

The continuous wavelet transform of \( f(x) \) is defined as

\[
w(a,b) = \int_{-\infty}^{\infty} f(x) \overline{\psi_{a,b}(x)} \, dx
\]  

(2)

and the function \( f(x) \) can be reconstructed by the inverse wavelet transform

\[
f(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(a,b) \psi_{a,b}(x) \frac{da \, db}{a^2}.
\]  

(3)

The continuous wavelet transform and its inverse transform are not directly implementable on digital computers. When the inverse wavelet transform is discretized, \( f(x) \) has the following approximate wavelet-based representation form:
\[ \hat{f}(x) \approx \sum_{k=1}^{K} w_k \psi \left( \frac{x-b_k}{a_k} \right), \]

(4)

where the \( w_k \), \( b_k \) and \( a_k \) are weight coefficients, translations and dilations for each daughter wavelet [22].

### 2.4 Complex discrete wavelet transform (CWT)

Wavelet techniques are successfully applied to various problems in signal and image processing. Data compression [23], motion estimation [24], segmentation and classification [25, 26] and denoising [27] are only some examples. It is perceived that the wavelet transform is an important tool for analysis and processing of signals and images. In spite of its efficient computational algorithm, the wavelet transform suffers from three main disadvantages.

**Limitations of wavelet transform**

Although the standard DWT is a powerful tool, it has three major disadvantages that undermines its application for certain signal and image processing tasks [28, 29].

**a) Shift sensitivity**

A transform is shift sensitive, if the shifting in time, for input-signal causes an unpredictable change in transform coefficients. It has been observed that the Standard DWT is seriously disadvantaged by the shift sensitivity that arises from down samplers in the DWT implementation [28, 30]. Shift sensitivity is an undesirable property because it implies that DWT coefficients fail to distinguish between input-signal shifts.

**b) Poor directionality**
An m-dimensional transform (m>1) suffers poor directionality when the transform coefficients reveal only a few feature orientations in the spatial domain. Wavelet transform has been poor directional selectivity for diagonal features. Because the wavelet filters are separable and real.

c) Absence of phase information

For a complex-valued signal or vector, its phase can be computed by its real and imaginary projections. Phase information is valuable in many signal and image processing applications [31] such as e.g. in image compression and power measurement [32, 33].

Most DWT implementations use separable filtering with real coefficient filters associated with real wavelets resulting in real-valued approximations and details. Such DWT implementations cannot provide the local phase information. All natural signals are basically real-valued, hence to avoid the local phase information, complex-valued filtering is required [34, 35].

Recent research in the development of CWTs can be broadly classified in two groups; RCWT (Redundant CWTs) and NRCWT (Non-redundant CWTs). Standard DWT is critically decimated and gives N samples in transform domain for the same N samples of a given signal. While the redundant transform gives M samples in transform domain for N samples of given input signal (where M>N) and hence it is expensive by the factor M/N. The NRCWT follows the design aim to approach towards N samples in transform domain for a given N input samples [28, 29].

The RCWT include two almost similar CWTs. They are denoted as DT-DWT (Dual-Tree DWT based CWT, see Figure 2) with two almost similar versions namely Kingsbury’s and Selesnick’s [36]. In this paper, we used Kingsbury’s CWT [27, 35] for feature extraction of image to be segmented.
Recently, there has been an increased interest in applications of the CVANN to process complex signals [37-39]. In this study, a complex back-propagation (CBP) algorithm has been used for image segmentation. We will first give the theory of the CBP algorithm as applied to a multi layer CVANN. Figure 3 shows a CVANN model used in the study.
Figure 3: A model neuron used in the complex-BP algorithm.

The input signals, weights, thresholds, and output signals are all complex numbers. The activity $Y_n$ of neuron $n$ is defined as:

$$Y_n = \sum_m W_{nm} X_m + V_n \quad (5)$$

where $W_{nm}$ is the complex-valued (CV) weight connecting neuron $n$ and $m$, $X_m$ is the CV input signal from neuron $m$, and $V_n$ is the CV threshold value of
neuron \( n \). To obtain the CV output signal, the activity value \( Y_n \) is converted into its real and imaginary parts as follows:

\[
Y_n = x + iy = z
\]  

(6)

where \( i \) denotes \( \sqrt{-1} \). Although various output functions of each neuron can be considered, the output function used in this study is defined by the following equation:

\[
f_c(z) = f_R(x) + i.f_R(y)
\]

(7)

where \( f_R(u) \) is called the activation function of neural network. One of the difficulties encountered in applying the CBP algorithm to the complex domain involves the appropriate choice of activation function. For a practical implementation of the complex multilayer perceptron, it is necessary that the activation function be bounded. Several researchers developed a set of properties that a complex activation function must satisfy in order to be useful in a multilayer perceptron trained with the back-propagation algorithm [40]. Complex activation function that used in this study is a superposition of real and imaginary logarithmic sigmoids, as shown by

\[
f_R(u) = \frac{1}{1 + \exp(-u_R)} + j \frac{1}{1 + \exp(-u_i)}
\]

(8)

Summary of CBP algorithm:

1. Initialization

Set all the weights and thresholds to small complex random values.
2. Presentation of input and desired (target) outputs

Present the input vector X(1), X(2),... ,X(N) and corresponding desired (target) response T(1), T(2),... ,T(N), one pair at a time, where N is the total number of training patterns.

3. Calculation of actual outputs

To obtain the complex-valued output signal, the activity value $Y_n$ is converted into its real and imaginary parts as follows:

$$Y_n = x + iy = z$$  \hspace{1cm} (9)

where $i$ denotes $\sqrt{-1}$.

4. Calculation of the stopping criteria with respect to Eq (10) [38].

If this condition is satisfied, algorithm is stopped and weights and biases are frozen.

$$\sqrt{\sum_{p} \sum_{n=1}^{N} |T_n^{(p)} - O_n^{(p)}|^2} = 10^{-1}$$ (10)

where $T_n^{(p)}$ and $O_n^{(p)}$ are complex numbers and denote the desired and output value, respectively. The actual output value of the neuron $n$ for the pattern $p$, i.e the left side of (Eq 10) denotes the error between the desired output pattern and the actual output pattern. $N$ denotes the number of neurons in the output layer.

5. Adaptation of weights and thresholds
We will use $W_{ml}$ for the weight between the input neuron $l$ and the hidden neuron $m$, $V_{nm}$ for the weight between the hidden neuron $m$ and the output neuron $n$, $\theta_m$ for the threshold of the hidden neuron $m$, and $\gamma_n$ for the threshold of the output neuron $n$. Let $I_l$, $H_m$, $O_n$ denote the output values of the input neuron $l$, the hidden neuron $m$, and the output neuron $n$, respectively. Let also $U_m$ and $S_n$ denote the internal potentials of the hidden neuron $m$ and the output neuron $n$, respectively. $U_m$, $S_n$, $H_m$, and $O_n$ can be defined respectively as

$$U_m = \sum_l W_{ml} I_l + \theta_m,$$

$$S_n = \sum_m V_{nm} H_m + \gamma_n,$$

$$H_m = f_c(U_m),$$

$$O_n = f_c(S_n).$$

Let $\delta_n = T_n - O_n$ denote the error between the actual pattern $O_n$ and the target pattern $T_n$ of output neuron $n$. We will define the square error for the pattern $p$ as

$$E_p = \frac{1}{2} \sum_{n=1}^{N} |T_n - O_n|^2,$$

where $N$ is the number of output neurons.

We can show that the weights and the thresholds should be modified according to the following equations [38].

$$\Delta V_{nm} = -\varepsilon \left( \frac{\partial E_p}{\partial \text{Re}[V_{nm}]} - i \varepsilon \frac{\partial E_p}{\partial \text{Im}[V_{nm}]} \right)$$

$$\Delta \gamma_n = -\varepsilon \left( \frac{\partial E_p}{\partial \text{Re}[\gamma_n]} - i \varepsilon \frac{\partial E_p}{\partial \text{Im}[\gamma_n]} \right)$$

$$\Delta W_{ml} = -\varepsilon \left( \frac{\partial E_p}{\partial \text{Re}[W_{ml}]} - i \varepsilon \frac{\partial E_p}{\partial \text{Im}[W_{ml}]} \right)$$
\[
\Delta \theta_n = -\varepsilon \frac{\partial E_p}{\partial \text{Re}[\theta_n]} - i\varepsilon \frac{\partial E_p}{\partial \text{Im}[\theta_n]} \tag{14}
\]

Equations (11)-(14) can be expressed as:

\[
\Delta V_{nm} = H_n \Delta \gamma_n \tag{15}
\]

\[
\Delta \gamma_n = \varepsilon \left( \text{Re}[\delta^n] \bigg[ 1 - \text{Re}[O_n] \bigg] \text{Re}[O_n] \right)
+ i \varepsilon \left( \text{Im}[\delta^n] \bigg[ 1 - \text{Im}[O_n] \bigg] \text{Im}[O_n] \right) \tag{16}
\]

\[
\Delta W_{nl} = \bar{I}_l \Delta \theta_n \tag{17}
\]

\[
\Delta \theta_n = \varepsilon \sum_n \text{Re}[\delta^n] \left( \begin{array}{c}
\left[ (1 - \text{Re}[H_m]) \text{Re}[H_m] \right] \\
\text{Re}[O_n] \text{Re}[V_{nm}] \\
+ \text{Im}[\delta^n] \left( 1 - \text{Im}[O_n] \right) \\
\text{Im}[O_n] \text{Im}[V_{nm}] 
\end{array} \right)
\]
where $\overline{z}$ denotes the complex conjugate of a complex number $z$.

**Calculation of Training and Test Errors:**

The training and test errors given in tables were found according to Eq. (19).

$$\text{Error(\%)} = \left( \frac{\sum_{i=1}^{k} |t(i) - a(i)|}{m* n} \right) * 100$$  \hspace{1cm} (19)$$

where $t(i)$ is desired outputs, $a(i)$ is outputs of neural network, $k$ is the number of samples in training or test data, $m$ is the number of segments in training or test data and $n$ is the number of outputs of neural network for training and test procedures [12].
3. Results

In this paper, 3D thinning of hybrid shaped porous media is realized using CWT-CVANN. Firstly, 2D slices were extracted from the 3D objects using curve thinning (CTA) and surface thinning (STA). Then, CTA and STA parts of image are determined as real and imaginary parts of input image, respectively. Features of the obtained complex-valued images (CTA + i.STA) are extracted using CWT with two levels. For the second level of CWT, input matrix size of images was reduced to 32x32 from 128x128 (original image size). These new feature vectors are used to CVANN as inputs. The outputs of CVANN are separated to real (CTA) and imaginary (STA) components as illustrated in Figure 4.

For comparison, WT was used for feature extraction of CTA and STA, separately. Daubechies 2 function is selected as wavelet function and applied with two levels. Thus, size of CTA and STA was reduced to 32x32. Obtained new images of CTA and STA were coded as complex-valued (CTA + i.STA). Figure 5 shows the hybrid skeleton process using WT-CVANN.
The complex-valued backpropagation algorithm is used for training of the proposed networks. In training phase, the weights and biases of CVANNs are initialized with small random complex numbers. An error goal (stopping criteria of training of $10^{-1}$) is specified as in equation (10). The training of WT-CVANN and CWT-CVANN is stopped when the error goal is achieved. Then, performance of networks is tested by presenting test images. The optimum learning rate and numbers of hidden nodes are determined as 0.7 and 60, respectively, via experimentation. Number of iterations is determined as 1000 for all structures.

Test results are presented in Table-1. As seen in Table-1, the second level CWT-CVANN structure achieves good results than WT-CVANN. Error rate is 0.1317 % for CWT-CVANN. For the same images, higher error than CWT-CVANN is obtained using WT-CVANN. Error of WT-CVANN is calculated as 0.7225 %. The resulting output images using two structures for random selected input images are given in Figure 6, Figure 7 and Figure 8. These images show that, using WT as feature extractor for hybrid
skeletonization task generates artifacts in the images, especially imaginary part of image (STA). These artifacts are occurred high error in the image analysis.
Table 1. CWT-CVANN and WT-CVANN results for 10 different images

<table>
<thead>
<tr>
<th>Patient File</th>
<th>WT-CVANN</th>
<th>CWT-CVANN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Error (%)</td>
<td>Error (%)</td>
</tr>
<tr>
<td>1</td>
<td>0.5053</td>
<td>0.0873</td>
</tr>
<tr>
<td>2</td>
<td>0.6667</td>
<td>0.1272</td>
</tr>
<tr>
<td>3</td>
<td>0.6892</td>
<td>0.0480</td>
</tr>
<tr>
<td>4</td>
<td>0.7527</td>
<td>0.1896</td>
</tr>
<tr>
<td>5</td>
<td>0.6991</td>
<td>0.1056</td>
</tr>
<tr>
<td>6</td>
<td>0.6982</td>
<td>0.1181</td>
</tr>
<tr>
<td>7</td>
<td>0.6307</td>
<td>0.1474</td>
</tr>
<tr>
<td>8</td>
<td>1.0897</td>
<td>0.2808</td>
</tr>
<tr>
<td>9</td>
<td>0.7798</td>
<td>0.1421</td>
</tr>
<tr>
<td>10</td>
<td>0.7140</td>
<td>0.0710</td>
</tr>
<tr>
<td>AVERAGED</td>
<td>0.7225</td>
<td>0.1317</td>
</tr>
</tbody>
</table>
Figure 6. Results of hybrid skeletonization for random selected image (file no. 1)

(a) Target for CTA  (b) Outputs of CWT-CVANN for CTA  (c) Outputs of WT-CVANN for CTA  (d) Target for STA  (e) Outputs of CWT-CVANN for STA  (f) Outputs of WT-CVANN for STA
Figure 7. Results of hybrid skeletonization for random selected image (file no. 4)

(a) Target for CTA (b) Outputs of CWT-CVANN for CTA (c) Outputs of WT-CVANN for CTA (d) Target for STA (e) Outputs of CWT-CVANN for STA (f) Outputs of WT-CVANN for STA
4. Conclusions

This paper compares two feature extraction methods for CVANN based hybrid skeletonization. A two complex-valued neural network model for segmentation is studied: WT-CVANN and CWT-CVANN. The following conclusions may be drawn based on the results presented:
1- Feature extraction with CWT is found to be more satisfactory (error with 0.1317%) for skeletonizing hybrid shaped porous media.
2- Using WT for feature extraction generates artifacts in the images because of poor directionality of WT.

REFERENCES


