Normalized Embedded Zero-tree Wavelet Coding
Applied in Tracking DC Arc Furnace Characteristics
Considering Data Compression

SHU-CHEN WANG
Department of Computer and Communication Engineering
Taipei College of Maritime Technology
Taipei, Taiwan
scwang@mail.tcmt.edu.tw

CHENG-PING HUANG
Department of Mechatronic Engineering, Dahan Institute of Technology,
Hualien, Taiwan

CHI-JUI WU,
Dep. of Electrical Engineering, National Taiwan University of Science and Technology
Taipei, Taiwan
cjwu@mail.ntust.edu.tw

Abstract: - The application of data compression technique using the normalized embedded zero-tree wavelet (NEZW) coding is presented for the long-duration monitoring of a DC electric arc furnace, where voltage fluctuation and loading fluctuation are critical and stochastic load characteristics. While keep enough stochastic load information, it is desired to reduce data size in long duration recording of voltage and current waveforms. The effects of multi-resolution analysis levels and threshold values in the NEZW coding are investigated. From the calculation results of field measurement data, the NEZW coding almost preserves the values of voltage fluctuation and power quantities. For storage of the field measurement voltage and current waveforms, the NEZW coding can not only greatly reduce data size, but it also can preserve sufficient load information.

Key-Words: - electric power quality, data compression, discrete wavelet transform, multi-resolution analysis, normalized embedded zero-tree wavelet coding.

1 Introduction
Stochastic fluctuating loads, such as electric arc furnaces (EAFs), may cause the disturbances of voltage fluctuation (voltage fluctuation) [1-4]. However, it is sometimes desired to record the voltage and current waveforms for a long duration to track the disturbance levels. The instantaneous waveforms are helpful in determining mitigation methods and devices. But the recording of instantaneous three-phase voltage and current waveforms, a few seconds period (hundreds of power cycles) would need memory high to several mega bytes. Therefore, in consideration of memory space and cost down, the recorded data length of an instrument will be limited. How to effectively reduce the data size becomes a very important issue.

There are basically many approaches in data compression. It generally uses a suitable transform to analyze the data and therefore a coding method is to extract the characteristic information and remove redundancy. Many transform methods have been used, such as discrete Fourier transform, short-term discrete Fourier transform, discrete cosine transform (DCT), and discrete wavelet transform (DWT) [5-6]. There are several representatives of state-of-art coders, such as embedded zero-tree wavelet coding (EZW) [7], embedded zero-tree DCT coding, zero-tree entropy coding, set partition in hierarchical trees coding, embedded block coding with optimized truncation [8], and vector quantization (VQ) coding [9]. For example, the EZW coding is the data compression algorithm of JPEG2000 [5-6].

In this paper, the DWT with the normalized embedded zero-tree wavelet (NEZW) coding, is presented and used as a data compression method for long-duration recording purposes. A DC EAF and a railway electrification substation are examined. The compression results are compared with the well-known threshold (TH) coding and vector quantization (VQ) coding. The errors in integral...
values of arithmetic apparent power, fundamental active power, and fundamental reactive power are also used to reveal that the data compression approach can keep the information of load fluctuating.

2 Wavelet Transform

The continuous wavelet transform of signal \( x(t) \) is

\[
CWT(a, b) = \langle x(t), \psi_{a,b}(t) \rangle = \int_{-\infty}^{\infty} x(t) \psi_{a,b}(t) \, dt
\]

where \( a \) is the scale parameter, and \( b \) is the time shift parameter. The wavelet basic function \( \psi_{a,b}(t) \) is defined on the right half plane of \( (a, b) \), that is, \( b \in \mathbb{R}, a > 0 \) . The original signal \( x(t) \) can be reconstructed by the parameter \( (a, b) \). The inverse continuous wavelet transform is given by

\[
x(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int \psi_{a,b}(t) \frac{d\psi_{a,b}(t)}{a^2}
\]

where \( C_\psi \) is the normalized constant.

For a digital recorder with a sampling period \( T \), the discrete wavelet transform (DWT) is needed. Selecting \( a=2^m \), \( b=n2^m b_0 \), \( t=kT \), and \( k, m, n \in \mathbb{Z} \), the DWT is

\[
DWT(m, n) = \langle x[k], \psi_{m,n}(k) \rangle = \sum_k x[k] \psi^*(k-na^m b_0) / a^m
\]

The inverse discrete wavelet transform (IDWT) is given by

\[
x[k] = \left( \sum_{m,n} DWT(m, n) \psi^*(k-na^m b_0) / a^m \right) / C_\psi
\]

So we can use multi-resolution filters to achieve DWT or IDWT.

3 Normalized Embedded Zero-Tree Wavelet (NEZW) Coding

If the DWT is applied to data compressing, it will relate to the address characteristics of wavelet set in multi-resolution analysis (MRA). The original signal is decomposed to several levels. The NEZW coding is used to extract the characteristic information and remove the redundancy as shown in Fig. 1, where a 3-level MRA decomposition and reconstruction method is given. In the \( s^{th} \)-level, \( s=1, 2, \) and \( 3 \), \( C_{s-1}[n] \) is resolved into \( C_s[n] \) and \( D_s[n] \), where the high frequency band coefficients \( D_s[n] \) will be processed through the NEZW coding. The coefficients in wavelet decomposition form an inverted hierarchical pyramid, with the coarsest scale components at the bottom and the finest scale components at the top. In the interior of the pyramid, each coefficient of \( D_s[n] \) has two children at the next finer scale, \( D_{s-1}[2n] \) and \( D_{s-1}[2n+1] \), which correspond to the same spatial locations as \( D_s[n] \). If the original signal \( C_0[n] \) is normalized, it can let magnitudes of all \( D_s[n] \) be less than 2 in each level.

Therefore, in contrast to EZW, the NEZW algorithm can adopt 1 (i.e. \( 2^0 \)) to be the initial threshold value instead of searching the initial threshold value through all \( D_s[n] \) for any original signal \( C_0[n] \). Figure 2 shows an example of the high frequency band coefficients for a sampled signal in a 3-level MRA. The significance of a coefficient is determined by comparing its magnitude with a set of gradually decreasing thresholds. The forward wavelet transform concentrates its information into a relatively small number of coefficients with large magnitude. That is, the NEZW prioritization scheme transmits the large (significant) coefficients before transmitting the smaller (insignificant) coefficients.

The choices of thresholds in this paper would be \( 2^0, 2^{-1}, 2^{-2}, 2^{-3}, 2^{-4} \), and so on. The effect is like as

\[
\text{Fig. 1. Decomposition and reconstruction using 3-level multi-resolution analysis and NEZW coding.}
\]

\[
\text{Fig. 2. The inverted hierarchical pyramid parent-child relationships for one tree of wavelet coefficients.}
\]
factorization in mathematics. Suppose the minimum threshold value of NEZW coding is limited to $4^2$, the coding algorithms are as follows. Figure 3 demonstrates the synopsis for the example in Fig. 2.

3.1 Encoding Algorithm
The original signal is split into non-overlapping blocks of samples. Each block is transformed and encoded separately. The blocking approach allows compression to be done in real time. Let RL (the refinement list) be the list of coefficients which have previously been found to be significant, and let SL (the search list) be the list of coefficients which have not yet been found to be significant. The compression algorithm is as follows.

1. Set $k=0$ and set RL to empty.
2. For each $D_s[n]$ in RL, output the most significant bit of $D_s[n]$.
3. Set SL to be all the coefficients except those in RL, ordered so that parents are listed before their children. For each $D_s[n]$ in SL:
   - If $D_s[n] \geq 2^{-k}$, then output POS to the symbol S and add $D_s[n]$ to RL; else if $D_s[n] \leq -2^{-k}$ then output NEG and add $D_s[n]$ to RL; else if no descendant of $D_s[n]$ is significant, i.e., $2^{-k} > D_s[n] > -2^{-k}$, then output ROOT and remove all the descendants of $D_s[n]$ from SL; else if $D_s[n]=0$ then do nothing.
4. If $k \leq 4$, then increase $k$ by one and go to Step (2).

3.2 Decoding Algorithm
The decoding algorithm is as follows:

1. Let $k=0$, set RL to empty, and set all $D_s[n]$ to zero.
2. For each $D_s[n]$ in RL, input the most significant bit of $D_s[n]$.
3. Set SL to be all the coefficients except those in RL, ordered so that parents are listed before their children. For each $D_s[n]$ in SL:
   - Input a symbol S. If S=POS then set $D_s[n]$ to $2^{-k}$ and add to RL; else if S=NEG then set $D_s[n]$ to $-2^{-k}$ and add $D_s[n]$ to RL; else if S=ROOT then remove all the descendants of $D_s[n]$ from SL; else if S=ZERO then do nothing.
4. If $k \leq 4$, then increase $k$ by one and go to Step (2).

4 Characteristics of Fluctuating Load
4.1 Voltage fluctuation
In a short duration, a voltage fluctuation waveform can be described as

$$v(t) = \sqrt{2} V_{rms} [1 + \frac{1}{2} \sum_{n} \Delta V_{f_n} \sin(2\pi f_n t + \phi_n)] \sin(2\pi f_{sys} t)$$

(6)

where $f_{sys}$ is the fundamental frequency (power frequency), $V_{rms}$ is the RMS value, and $\Delta V_{f_n}$ is fluctuation component of the amplitude modulation frequency $f_n$. For the voltage fluctuation limitation, we only need to consider $f_n$ in the range of 0.1Hz~30Hz. The definitions of voltage deviation $\Delta V$ is

$$\Delta V = \sqrt{\sum_{n}(\Delta V_{f_n})^2}$$

(7)

4.2 Power Quantities [10]
For a three-phase three-wire load under non-sinusoidal and unbalanced conditions, the arithmetic apparent power is

$$S_A = S_R + S_S + S_T = V_R I_R + V_S I_S + V_T I_T$$

(8)

If only fundamental components are considered, the fundamental active power and reactive power are, respectively, as follows.

$$P_1 = P_{R1} + P_{S1} + P_{T1}$$

(9)

$$Q_1 = Q_{R1} + Q_{S1} + Q_{T1}$$

(10)

Therefore, the corresponding fundamental apparent power is given by

$$S_1 = P_1^2 + Q_1^2$$

(11)

To reveal harmonic condition, the non-fundamental arithmetic apparent power can be defined as
Additionally, the effective representation of three-phase fundamental voltages and currents can be given by

\[
V_{el} = \sqrt{\frac{V_{11}^2 + V_{11}'^2 + V_{11}''^2}{3}}
\]

\[
I_{el} = \sqrt{\frac{I_{11}^2 + I_{11}'^2 + I_{11}''^2}{3}}
\]

Then the fundamental effective apparent power is defined as

\[
S_{el} = 3V_{el}I_{el}
\]

It is noted that \( S_{el} \) is different with \( S_1 \). When there is an unbalanced situation, the fundamental positive-sequence apparent power is defined as

\[
S_i^+ = 3V_i^+I_i^+
\]

Where \( V_i^+ \) and \( I_i^+ \) are the fundamental positive-sequence components. Therefore the unbalanced condition can be represented by the fundamental unbalanced apparent power as follows.

\[
S_i^u = \sqrt{S_{el}^2 - S_i^+^2}
\]

5 Test of Voltage Fluctuation and Harmonic Current Waveforms

In order to reveal the efficiency of data compression, the compression ratio (CR) is defined as

\[
CR = \frac{\text{original file size}}{\text{compressed file size}}
\]

The evaluation of the signal quality after reconstruction can be achieved by the normalized mean-square error (NMSE). It is given by

\[
NMSE = \frac{\|C[n] - \hat{C}[n]\|^2}{\|C[n]\|^2}
\]

TABLE I shows the compression results of NEZW coding, TH coding, and VQ coding that are applied to a given voltage fluctuation waveform. The denotation MRAx_NEZW2\(^y\) implies the x-level MRA with NEZW coding and its threshold value down to 1/(2\(^y\)). Similarly, MRAx TH and MRAx VQ mean the x-level MRA with TH coding and VQ coding, respectively. Bioorthogonal-3.3 is chosen through an evaluation for the following study. Some observations can be obtained.

1. With a given threshold in the NEZW, the CR values and the errors increase with increasing decomposition levels.

2. With a fixed decomposition level, the CR values and the errors decrease with decreasing threshold.

(3) With a same resolution level (MRA3), the NEZW coding is the best and the VQ coding is the worst in data compression to effectively keep accuracy of voltage fluctuation values.

![Fig. 4. One-line diagram of the DC electric arc furnace steel plant.](image1)

![Fig. 5. 60-cycle voltage waveforms measured at DC EAF feed line.](image2)

![Fig. 6. 60-cycle current waveforms measured at DC EAF feed line.](image3)
the one-line diagram of the plant. Only for illustration, the 60-cycle three-phase voltage and current waveforms (64 samples per cycle, total 3840 samples) measured at the EAF feeder during the melting-down

![Image of voltage and current waveforms](image)

Fig. 7. The \( \Delta V \) values of the original measured voltage waveform and their errors versus to reconstructed waveforms of three coding methods in the 15-minute period of DC EAF feed line.

period are shown in Fig. 5 and Fig. 6, respectively. They reveal that the fluctuations of voltage and current waveforms are obvious. In total, 3,456,000 samples were recorded in the 15-minute measurement and used to reveal the performance of data compression. Considering a lower NMSE and an acceptable CR, MRA3_NEZW2-4, MRA3_TH, and MRA3_VQ are chosen for comparison.

### 6.1 Voltage fluctuation values

Figure 7 shows the \( \Delta V \) values of the original voltage waveform and errors of reconstructed waveforms in the 15-minute period.

### 6.2 Power quantities

The second column of TABLE II shows the integral true values of \( P_1, Q_1, S_1, S_A, S_{AN}, \) and \( S_{IU} \) of the 15-minute measurement data, respectively. The other figures within Figs. 8-10 give deviations of absolute errors of those powers between the original and reconstructed waveforms of the 15-minute measurement data. The third, the fourth, and the fifth column of TABLE II also give the integral absolute errors and percentage errors of those power quantities between the original and reconstructed waveforms of the 15-minute measurement data. It can be found that the MRA3_NEZW2-4 coding and MRA3_TH coding have lower errors than MRA3_VQ one. Especially, the MRA3_NEZW2-4 coding almost reflects the fundamental power quantities truly, and the integral errors are even lower than 0.1%. While the integral error of \( S_{AN} \) is a little higher, that of \( S_{IU} \) is not. However, the MRA3_VQ coding would cause larger errors, where the integral error of \( S_{IU} \) is even up to 356%.

### 7 Conclusion

Since the voltage and current waveforms of fluctuating load may be different cycle by cycle, the compression ratios of traditional data compression methods may be not large to keep enough data information. The data compression method using the normalized embedded zero-tree wavelet (NEZW) coding has been presented to be effective in reducing the data size in recording waveforms of DC EAF. This approach is better than conventional methods. It gives excellent ability to preserve information of voltage fluctuation and power quantities. From the computation results of the field measurement voltage and current waveforms, this method can reduce data memory size near to 15% while almost keep power quality characteristics. It is very suitable to be applied to power quality instruments in long duration load monitoring.

### References:


### TABLE I  COMPRESSION RESULTS OF THE GIVEN VOLTAGE FLUCTUATION WAVEFORM

<table>
<thead>
<tr>
<th></th>
<th>Vrms</th>
<th>$\Delta V$</th>
<th>$\Delta V_0$</th>
<th>$\Delta V_5$</th>
<th>$\Delta V_{10}$</th>
<th>$\Delta V_{15}$</th>
<th>$\Delta V_{30}$</th>
<th>CR</th>
<th>NMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Given Value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRA3_NEZW2^-5</td>
<td>1</td>
<td>0.031</td>
<td>0.052</td>
<td>0.08</td>
<td>0.047</td>
<td>0.003</td>
<td>11.08</td>
<td>6.92</td>
<td>5.57E-06</td>
</tr>
<tr>
<td>MRA3_NEZW2^-4</td>
<td>1</td>
<td>0.031</td>
<td>0.052</td>
<td>0.08</td>
<td>0.047</td>
<td>0.003</td>
<td>11.08</td>
<td>7.89</td>
<td>1.37E-05</td>
</tr>
<tr>
<td>MRA3_NEZW2^-3</td>
<td>1</td>
<td>0.031</td>
<td>0.051</td>
<td>0.079</td>
<td>0.046</td>
<td>0.029</td>
<td>11.06</td>
<td>7.95</td>
<td>1.40E-05</td>
</tr>
<tr>
<td>MRA4_NEZW2^-5</td>
<td>1</td>
<td>0.031</td>
<td>0.052</td>
<td>0.08</td>
<td>0.047</td>
<td>0.003</td>
<td>11.07</td>
<td>10.69</td>
<td>2.99E-05</td>
</tr>
<tr>
<td>MRA4_NEZW2^-4</td>
<td>1</td>
<td>0.031</td>
<td>0.051</td>
<td>0.079</td>
<td>0.047</td>
<td>0.003</td>
<td>11.05</td>
<td>11.54</td>
<td>5.10E-05</td>
</tr>
<tr>
<td>MRA4_NEZW2^-3</td>
<td>1</td>
<td>0.031</td>
<td>0.051</td>
<td>0.079</td>
<td>0.046</td>
<td>0.029</td>
<td>10.99</td>
<td>12.80</td>
<td>1.91E-04</td>
</tr>
<tr>
<td>MRA3_TH</td>
<td>1</td>
<td>0.031</td>
<td>0.052</td>
<td>0.08</td>
<td>0.047</td>
<td>0.003</td>
<td>11.08</td>
<td>6.87</td>
<td>7.12E-05</td>
</tr>
<tr>
<td>MRA4_TH</td>
<td>1</td>
<td>0.031</td>
<td>0.050</td>
<td>0.077</td>
<td>0.045</td>
<td>0.003</td>
<td>10.66</td>
<td>11.9</td>
<td>2.11E-03</td>
</tr>
<tr>
<td>MRA3_VQ</td>
<td>1</td>
<td>0.016</td>
<td>0.041</td>
<td>0.061</td>
<td>0.031</td>
<td>0.007</td>
<td>8.17</td>
<td>4.23</td>
<td>2.10E-03</td>
</tr>
<tr>
<td>MRA4_VQ</td>
<td>1</td>
<td>0.019</td>
<td>0.028</td>
<td>0.044</td>
<td>0.020</td>
<td>0.010</td>
<td>5.98</td>
<td>4.70</td>
<td>3.55E-03</td>
</tr>
</tbody>
</table>

### TABLE II  COMPRESSION RESULTS OF INTEGRAL POWER QUANTITY OF THE 15-MINUTE MEASUREMENT DATA (USING MRA3 ANALYSIS)

<table>
<thead>
<tr>
<th>Integral Power Quantity</th>
<th>True Value</th>
<th>NEZW2^-4 Absolute Error (% Error)</th>
<th>TH Absolute Error (% Error)</th>
<th>VQ Absolute Error (% Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sum P_1 \Delta t$ (kWh)</td>
<td>449.175</td>
<td>0.029 (0.006)</td>
<td>0.631 (0.141)</td>
<td>60.215 (13.4)</td>
</tr>
<tr>
<td>$\sum Q_1 \Delta t$ (kvarh)</td>
<td>-707.415</td>
<td>0.326 (0.046)</td>
<td>0.662 (0.094)</td>
<td>84.851 (12.0)</td>
</tr>
<tr>
<td>$\sum S_1 \Delta t$ (kVAh)</td>
<td>841.201</td>
<td>0.304 (0.036)</td>
<td>0.898 (0.107)</td>
<td>103.245 (12.3)</td>
</tr>
<tr>
<td>$\sum S_4 \Delta t$ (kVAh)</td>
<td>897.240</td>
<td>2.468 (0.275)</td>
<td>5.048 (0.563)</td>
<td>101.852 (11.4)</td>
</tr>
<tr>
<td>$\sum S_{4X} \Delta t$ (kVAh)</td>
<td>304.227</td>
<td>7.126 (2.342)</td>
<td>13.371 (4.395)</td>
<td>18.401 (6.05)</td>
</tr>
<tr>
<td>$\sum S_{1u} \Delta t$ (kVAh)</td>
<td>389.895</td>
<td>0.661 (0.170)</td>
<td>0.799 (0.205)</td>
<td>69.071 (17.7)</td>
</tr>
</tbody>
</table>