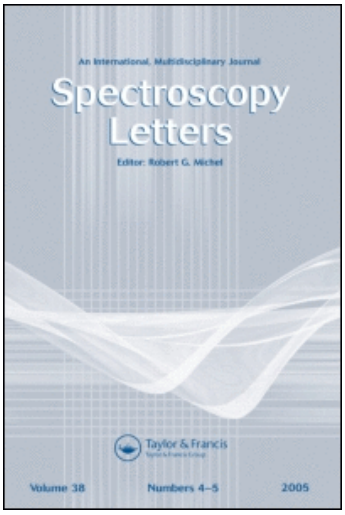


Spectroscopy Letters
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597299

Method for Infrared Spectral Compression Based on the Embedded Zerotree Wavelet

Menglong Li[a]; Fusheng Nie[a]; Huayi Qi[a]; Zhining Wen[a]; Bin Kang[a]
[a] College of Chemistry, Sichuan University, Chengdu, Sichuan, People's Republic of China

To cite this Article Li, Menglong , Nie, Fusheng , Qi, Huayi , Wen, Zhining and Kang, Bin(2005) 'Method for Infrared Spectral Compression Based on the Embedded Zerotree Wavelet', Spectroscopy Letters, 38: 2, 171 — 184
To link to this Article: DOI: 10.1081/SL-200049602
URL: http://dx.doi.org/10.1081/SL-200049602

Spectroscopy Letters, 38: 171–184, 2005

ISSN 0038-7010 print/1532-2289 online
DOI: 10.1081/SL-200049602

Method for Infrared Spectral Compression Based on the Embedded Zerotree Wavelet

Menglong Li, Fusheng Nie, Huayi Qi, Zhining Wen, and Bin Kang
College of Chemistry, Sichuan University, Chengdu, Sichuan,
People's Republic of China

Abstract: In this paper, the embedded zerotree wavelet (EZW) method and Huffman coding is proposed for the compression of infrared (IR) spectra. This technique is described based on the wavelet transform (WT) and tested with IR spectra of some compounds and compared with another compression technique. The results showed that this technique is better than others in terms of efficiently coding wavelet coefficients, because the zerotree quantization is an effective way of exploiting the self-similarities of wavelet coefficients at various resolutions.

Keywords: Compress, Embedded zerotree wavelet, Infrared spectra, Wavelet transform

INTRODUCTION

The wavelet transform (WT) has been developed rapidly in recent years and has been found to play a significant role in signal processing. From 1989 onwards, the WT has been applied in chemical studies owing to its efficiency, the large number of basis functions available, and the high speed in data treatment.[1,2]

Received August 19, 2003, Accepted September 25, 2004

This paper was invited as a contribution to a special issue of the journal concerning Chemical Spectroscopy in China. It was presented at Sichuan University, Chengdu, China, August 1–7, 2004: the 4th Conference for Worldwide Young Chinese Chemists.

Address correspondence to Menglong Li, College of Chemistry, Sichuan University, Chengdu, Sichuan 610064, People's Republic of China. Fax: +86-28-8541-2907; E-mail: liml@scu.edu.cn

Infrared (IR) spectroscopy has found widespread use in the identification and characterization of chemicals. IR spectral libraries such as the Aldrich Library have been established for this purpose. In general, the more information stored within the spectrum, the better for searching the database. However, as the size of a library increases, more space and a longer time for searching the library are required. To tackle the problem, researchers may use several approaches. An important one is to compress the spectra to a smaller data set. The WT has been proposed for the purpose, but the question is how to achieve the most efficient compression. For example, how to process these wavelet coefficients efficiently is a problem remaining to be solved. In previous work, most researchers often discarded small-amplitude values directly but without considering the self-similarities of wavelet coefficients at various resolutions.[3–10] In this work, we applied a different WT procedure. The embedded zerotree wavelet (EZW) algorithm, which was initially proposed by Shapiro to process images,[11] was first used in analytical chemistry in this work.

THEORY

The EZW algorithm is a simple yet remarkably effective compression algorithm, having the property that the bits in the bitstream are generated in order of importance, yielding a fully embedded code. The EZW encoder is based on progressive encoding to compress a spectrum into a bitstream with increasing accuracy. This means that when more bits are added to the stream, the well-encoded spectrum will contain more detail. Every added digit increases the accuracy somehow, and it can be stopped at any accuracy we like. Furthermore, this algorithm requires no training, no prestored table or codebooks, and requires no prior knowledge of spectra.

Wavelet Transform

Wavelets have shown great applicability in many diverse fields of science. Anything that involves the analysis of a time series is a good candidate for the wavelet treatment. The basic idea of wavelet analysis is that of multi-resolution. All of the basis functions are self-similar because they are derived from one prototype. A number of "self-similar" wavelets can be obtained from the mother wavelet by two process: (1) shifts in time variable that are needed to cover the whole signal range, and (2) dilations (or scaling), which allows a multiresolution analysis (MRA) of signals. A specific introduction to the theory on wavelets and MRA has been described in Refs.[12–15]

A signal function or vector (e.g., an IR spectrum) can be represented in WT at different scales and positions. Let the experimental data be expressed in digital form as

$$C^{(0)} = \{c_1^{(0)}, c_2^{(0)}, \ldots, c_N^{(0)}\} \tag{1}$$

at the lowest resolution level 0. Applying MRA to $C^{(0)}$, we get the scale $C^{(1)}$ and wavelet $D^{(1)}$ coefficients at the first level $J = 1$:

$$C^{(1)} = \left\{c_1^{(1)}, c_2^{(1)}, \ldots, c_{N/2}^{(1)}\right\} \tag{2}$$

$$D^{(1)} = \left\{d_1^{(1)}, d_2^{(1)}, \ldots, d_{N/2}^{(1)}\right\} \tag{3}$$

Then, the decomposition process is applied to $C^{(1)}$ again to obtain $C^{(2)}$, $D^{(2)}$. The process is repeated until the preset resolution level J is reached (Fig. 1). Finally, the signal $C^{(0)}$ is transformed into detailed components at different levels as

$$WT\{C^{(0)}\} = \{C^{(J)}, D^{(J)}, D^{(J-1)}, \ldots, D^{(1)}\} \tag{4}$$

which are called the wavelet representation of $C^{(0)}$, with the data number equal to that of the original signal.

Because of the self-similarities among $D^{(1)}$, $D^{(2)}$, ..., $D^{(J-1)}$ and $D^{(J)}$, EZW is thought to be the most efficient compression method. In fact, EZW reorders wavelet coefficients in such a way that they can be compressed very efficiently.

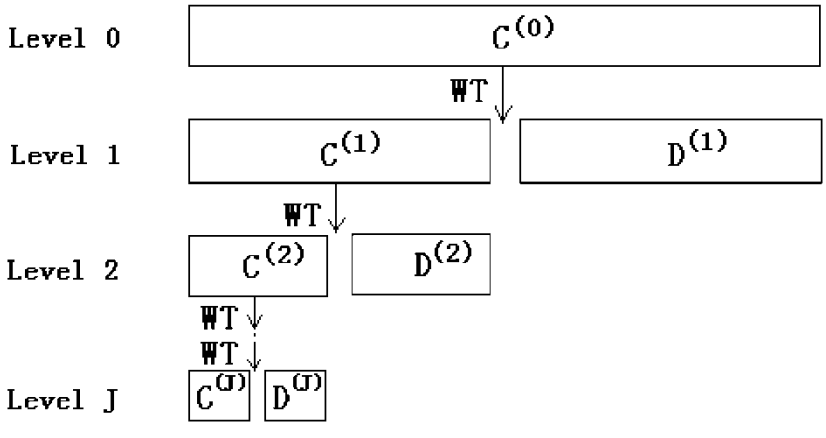


Figure 1. A diagram to show data decomposition at different resolution levels with the use of MRA.

Zerotree

The hierarchical structure at various scales and the self-similarity allow for a new quantization method called zerotree quantization.[11–15] The zerotree is a simple tree-structured significance scheme. We use it for quantizing an n-vector $Y = (Y_1, Y_2, \ldots, Y_n)$. The basic data structure in the scheme is a tree or a set of trees in which each leaf corresponds to one of the scalars Y_k. Given a threshold TH, a wavelet transform coefficient x is defined as significant with respect to a threshold TH if $|x| \geq TH$, otherwise x is said to be insignificant. The main trust of this quantization strategy is in the prediction of corresponding wavelet coefficients at the finer scales by exploiting the hierarchical structure and reducing the cost of encoding the significance scheme by exploiting the self-similarity. More traditional techniques are employing transform coding via some form of run-length encoding. It requires a symbol for each run-length, which must be encoded. Zerotree coding, however, will only encode a "terminating" symbol, which indicates that all remaining wavelet coefficients approximately equal to zero. In fact, this approach works well in terms of efficiently coding the wavelet coefficients. The zerotree is based on the hypothesis that if a coefficient at a coarse scale is insignificant with respect to a threshold, then all of its descendants, as defined below, are also insignificant and can be predicted. Empirical evidence suggests that this hypothesis is often true. The coefficient at the coarse scale is called the parent, and the coefficients corresponding to the same spatial location at the next finer scale of similar orientation are called children. For a given parent, the set of all coefficients at all finer scales of similar orientation corresponding to the same location are called descendants. Similarly, for a given child, the set of coefficients at all coarser scales of similar orientation corresponding to the same location are called ancestors (Fig. 2). For example, the children of $d_3^{(2)}$ are $d_5^{(1)}$ and $d_6^{(1)}$. The descendants of $d_2^{(3)}$are $d_3^{(2)}$, $d_4^{(2)}$, $d_5^{(1)}$, $d_6^{(1)}$, $d_7^{(1)}$, and $d_8^{(1)}$.

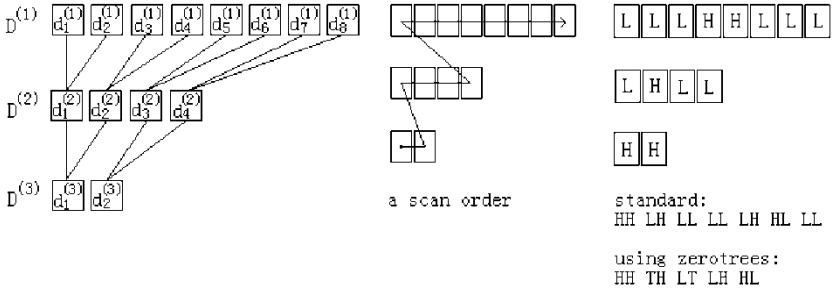


Figure 2. The relations between wavelet coefficients at different scales (left), how to scan them (middle), and the results of using zerotree symbols (T) in the coding process (right). An H means that the coefficient is higher than the threshold and an L means that it is below the threshold. The zerotree symbol (T) replaces some Ls.

Now let's see how to encode coefficient values. After taking the discrete wavelet transform of a set of spectra data, all wavelet transform coefficients except the approximate coefficients C will be ordered in such a way that the coefficients at coarser scale will come before the coefficients at finer scale. Given a threshold TH to determine whether or not a coefficient is significant, a coefficient x is said to be an element of a zerotree for threshold TH if itself and all of its descendants are insignificant with respect to TH. An element of a zerotree for threshold TH is a zerotree root if it is not the descendant of a previously found zerotree root for threshold TH (i.e., it cannot be predicted and is insignificant from the discovery of a zerotree root at a coarser scale with the same threshold). A zerotree root is encoded with a special symbol indicating that the insignificance of the coefficients at finer scales is completely predictable. Thus, in practice, four symbols are used: (1) (T) zerotree root; (2) (Z) isolated zero, which means that the coefficient is insignificant and is not an element of the zerotree; (3) (P) positive and significant coefficient; (4) (N) negative and significant coefficient.

A WT transforms a signal from the time domain to the joint time-scale domain. This means that the wavelet coefficients are two-dimensional. If we want to compress the transformed signal, we have to code not only the coefficient values, but also their positions in time. Indeed, without the information of the positions, the decoder will not be able to reconstruct the encoded signal. It is the encoding of positions that makes some encoders efficient and the others inefficient. Now if a scan order is predefined to scan the spectrum, going from the coarsest scale to the finest, implicitly many positions are coded through the use of zerotree symbols (Fig. 2).

Embedded Coding

Zerotree coding can perform so well on encoding the wavelet coefficients that it can be expected to encode the spectra more efficiently. So successive-approximation quantization (SAQ) is applied to perform the embedded coding.[11] In fact, embedded coding is similar in spirit to binary finite-precision representations of real numbers. All real numbers can be represented by a string of binary digits. Every digit we add increases the accuracy of the number, but we can stop at any accuracy we like. The SAQ sequentially applies a sequence of thresholds $TH_0, \ldots, TH_{N-1}$ to determine significance, where the thresholds are chosen so that $TH_i = TH_{i-1}/2$. If the threshold sequence is a sequence of power of two, it is called bit plane coding, as the thresholds in this case correspond to the bits in the binary representation of the coefficients. EZW encoding uses this type of coefficient value encoding.

First, all wavelet coefficients are measured against the initial threshold with a predefined scanning order, and a symbol is assigned to every coefficient. If an absolute value of the wavelet coefficient is larger than the threshold, it is

encoded with a "P" (for a positive) or an "N" (for a negative) and subtracts the threshold from the coefficient. If the wavelet coefficient is smaller and not a root of a zerotree, it is encoded with a "Z" (isolated zero) and left for the next scanning process. If the wavelet coefficient is smaller and is a root of a zerotree, it is encoded with a "T" (zerotree) and ignores all its descendants. Then, when all the wavelet coefficients are visited, the threshold is lowered, and the spectrum is scanned again to add more detail to the already encoded spectrum. This process is repeated until all the wavelet coefficients are encoded completely or some criterion is satisfied.

The initial threshold T_0 is chosen so that $|x_j| < 2TH_0$ for all transform coefficients x_j. If we adopt bit plane coding then it will be

$$TH_0 = M2^E \tag{5}$$

where E is an integer, and M is a constant.

The width of quantizer step size defines an uncertainty interval for the true magnitude of the coefficient. The reconstruction value can be anywhere in that uncertainty interval. But a practical approach used in the experiments, MINMAX optimal, is simply to use the center of the uncertainty interval as the reconstruction value. So in step 0, when x_j is coded as "P," the reconstruction value x_j' is $TH_0 + TH_0/2 = 3M2^{E-1}$ and the remainder r_j is $x_j - 3M2^{E-1}$, which satisfies:

$$|r_j| = |x_j - 3M2^{E-1}| \leq TH_0/2 = TH_1 \tag{6}$$

To perform this coding process better, in practice, two symbols P0 and P1 are used for positive significant coefficients; N1 and N0 are used for negative significant coefficients. If $x_j < TH_0 + TH_0/2$, x_j is coded as P0 and its reconstruction value x_j' is $TH_0 + TH_0/2 - TH_0/4 = 5M2^{E-2}$. If $x_j \geq TH_0 + TH_0/2$, x_j is coded as P1 and x_j' is $TH_0 + TH_0/2 + TH_0/4 = 7\,M2^{E-2}$. Then

$$|r_j| \leq TH_0/2^2 = TH_2 \tag{7}$$

Negative significant coefficients are coded as N0 and N1 similarly.

When all the coefficients have been scanned, the remainders are scanned with the threshold that is halved. The process is similar to the fore step. In step $N - 1$, the scanning process will be ended if all the remainders can be considered as zeros. This is SAQ. Of course, the process can be ended at any step according to actual needs.

Huffman Coding

EZW encoding does not really compress anything, it only reorders wavelet coefficients in such a way that they can be compressed very efficiently. Therefore, a symbol encoder should always follow an EZW encoder. The

Huffman coding method, one of the most popular entropy coding techniques, may be adopted.

In Huffman coding, symbols with higher probabilities are encoded with shorter codewords than the others. The generated codewords are of variable lengths. In this work, the output symbols produced by EZW can be considered as source symbols. The codeword assigned for the element of Z or T with the highest probability consists of two bits, which occupies the least storage space. Then four-bit codewords are designated for P1, P0, N1, and N0 that occur with the lowest probability. The Huffman code thus is a compact one with an average word length less than or equal to the average length of all other uniquely decodable codes for the same set of input probabilities, that is, it is a minimum-length code.

EXPERIMENTAL

In this work, we used 10 IR spectra to test the performance of compression. The IR spectra were exported from the Aldrich Condensed Phase Library by using the Nicolet OMNIC FT-IR software package. Each IR spectrum contains 460 data points.

All computations were carried out in Microsoft Windows 98 environment on a PC compatible with a Pentium processor. The programs WTRAN.M and IWTRAN.M were coded in MATLAB for carrying out the WT and inverse WT, respectively. The embedded zerotree procedure and Huffman coding and decoding process were performed with the use of the READ.CPP and UNREAD.CPP programs, which were developed in the C++ language.

RESULTS AND DISCUSSION

The proposed EZW method and Huffman coding algorithm were applied to compress the IR spectra of 10 compounds. The WTRAN.M program was used with the use of the biorthogonal wavelet series of bior1.1, bior2.8, bior3.7, bior4.4, and bior6.8.

In this work, the compression ratio, CR, measured the compression efficiency of the method mentioned above.

$$\mathrm{CR} = \text{No. of bytes of the original data/No. of bytes of the compressed data} \tag{8}$$

Ideally, the spectrum reconstructed from the compressed data is identical to the original one. In practice, this is not true, even if no data process is applied in the WT treatment, because of computational errors. The root-mean-square difference (RMSD) between the original spectrum $\{C_K^{(O)}\}$ and

the reconstructed spectrum $\{C_K^{(R)}\}$ was used to judge the validity of the proposed methods and is defined by

$$\mathrm{RMSD} = \left[\frac{1}{N}\sum_{K=1}^{N}(c_K^{(O)} - c_K^{(R)})^2\right]^{1/2} \tag{9}$$

Tables 1 and 2 shows the results obtained using different orders of biorthogonal wavelet. The resolution level J was set at a value of 6 and the final threshold TH_{N-1} was select to 0.015. In the last two rows of these tables, the average and standard deviation were calculated. It was found that these CRs are nearly same. From the viewpoint of the theory of function approximation, the principle of compression is to use as few terms of the sum of the basis functions as possible to approximate the original signal. The more similar the original signal to the basis function, the fewer the number of needed terms and therefore the higher the compression rate. Therefore, the CR of a simple spectrum must be large when compressed by a simple wavelet function, while the CR must be small by a complicated wavelet function. For example, the waveform of bior1.1 is simpler than that of

Table 1. Results of applying the EZW methods to compress IR spectra of the 10 compounds by using the biorthogonal wavelet functions bior1.1, bior2.8, and bior3.7

	bior1.1		bior2.8		bior3.7	
Spectra	CR	RMSD	CR	RMSD	CR	RMSD
Tridecane, 99+%	8.58	0.0050	8.27	0.0037	8.27	0.0047
(1*S*)-(−)-*b*-Pinene, 98%	6.32	0.0063	6.32	0.0081	6.07	0.0080
1-Chlorohexadecane, 99%	7.84	0.0054	7.34	0.0045	7.84	0.0052
1,5-Diiodopentane, 97%	7.58	0.0053	6.89	0.0047	7.00	0.0051
Pentachloroethane, 96%	11.1	0.0051	8.58	0.0045	8.27	0.0049
Tetrachloroethylene, 99+%	12.3	0.0048	10.83	0.0047	10.34	0.0051
1,3,5-Triphenylbenzene, 97%	5.83	0.0055	5.62	0.0062	5.42	0.0071
3-Chlorobenzyl chloride, 98%	6.89	0.0073	6.23	0.0067	6.07	0.0075
Decyl alcohol, 99%	5.83	0.0071	6.89	0.0053	7.71	0.0066
4-*t*-Butylcyclohexanol, mixture of isomers	5.91	0.0098	5.76	0.0055	5.62	0.0050
Average	7.818	0.00616	7.273	0.00539	7.261	0.00592
Standard deviation	2.264	0.00155	1.587	0.00130	1.529	0.00124

EZW, embedded zerotree wavelet; IR, infrared; CR, compression ratio; RMSD, root-mean-square difference.

Table 2. Results of applying the EZW methods to compress IR spectra of the 10 compounds by using the biorthogonal wavelet functions bior4.4 and bior6.8

	bior4.4		bior6.8	
Spectra	CR	RMSD	CR	RMSD
Tridecane, 99+%	8.43	0.004	8.92	0.0061
(1*S*)-(−)-*b*-Pinene, 98%	6.23	0.0068	5.83	0.0064
1-Chlorohexadecane, 99%	7.71	0.0044	8.13	0.007
1,5-Diiodopentane, 97%	6.79	0.0043	7.58	0.0065
Pentachloroethane, 96%	8.43	0.0037	7.84	0.0042
Tetrachloroethylene, 99+%	11.1	0.0055	10.58	0.0038
1,3,5-Triphenylbenzene, 97%	5.35	0.0058	6.32	0.009
3-Chlorobenzyl chloride, 98%	6.15	0.006	5.99	0.0058
Decyl alcohol, 99%	7.71	0.0056	7.84	0.0063
4-*t*-Butylcyclohexanol, mixture of isomers	5.91	0.0052	6.41	0.0067
Average	7.381	0.00513	7.544	0.00618
Standard deviation	1.691	0.00099	1.482	0.00144

EZW, embedded zerotree wavelet; IR, infrared; CR, compression ratio; RMSD, root-mean-square difference.

bior6.8. So the CR of simple spectrum (pentachloroethane) is 11.1 when using bior1.1, and it is 7.84 when using bior6.8. Similarly, the CR of a complicated spectrum must be small when compressed by a simple wavelet function, while the CR must be large by a complicated wavelet function. For instance, the CR of complicated spectrum (1,3,5-triphenylbenzene) is 5.83 if using bior1.1, and it is 6.32 if using bior6.8. Now some of the 10 spectra are simple and some of them are complicated. So the mean CRs with using different orders of biorthogonal wavelet are close to each other.

Then, we calculated the average CR and RMSD of each spectrum by using different wavelets (Table 3). One may conclude from the experimental data that a more complicated spectrum would result in a smaller CR. This observation can be explained by the fact that more peaks in the spectrum would give more high-frequency components in the wavelet domain. So this means that there are more significant coefficients in finer scales. This will lead to less zerotrees. Hence, more bits are needed to code them. Plots of the spectra with the highest CR and the lowest CR are given in Fig. 3 and Fig. 4, respectively. It is very clear that the simplest spectrum (tetrachloroethylene) holds the largest CR = 11.03 (Fig. 3) while the most complicated spectrum (1,3,5-triphenylbenzene) holds the smallest CR = 5.71 (Fig. 4).

In order to show the performance of the proposed compression method, we made a simple comparison. CR and RMSD were chosen as the key parameters for comparison. Table 4 shows the results in which method I

Table 3. The average CR and RMSD of each spectrum by using different wavelet functions

	Average		Standard deviation	
Spectra	CR	RMSD	CR	RMSD
Tridecane, 99+%	8.494	0.00470	0.2708	0.000941
(1*S*)-(−)-*b*-Pinene, 98%	6.154	0.00712	0.2079	0.000870
1-Chlorohexadecane, 99%	7.772	0.0053	0.2863	0.00104
1,5-Diiodopentane, 97%	7.168	0.00518	0.3834	0.000832
Pentachloroethane, 96%	8.844	0.00448	1.291	0.000559
Tetrachloroethylene, 99+%	11.03	0.00478	0.7643	0.000630
1,3,5-Triphenylbenzene, 97%	5.708	0.00672	0.3900	0.00141
3-Chlorobenzyl chloride, 98%	6.266	0.00666	0.3601	0.000757
Decyl alcohol, 99%	7.196	0.00618	0.8519	0.000733
4-*t*-Butylcyclohexanol, mixture of isomers	5.922	0.00644	0.2983	0.00199

CR, compression ratio; RMSD, root-mean-square derivative.

represents the proposed algorithm in this work, whereas method II represents the direct thresholding method as most researchers did in the previous similar work. It was found that our method gives superior performance. CR improves 36.1% on average and the biggest improvement by 68.1%. By the way, it is commonsense that the original data size will influence compression ratio. Under the other same conditions, the larger the original data number is, the higher the compression ratio is. In our experiment, the data number was 460, whereas in the previous similar work it was often larger. So, it is

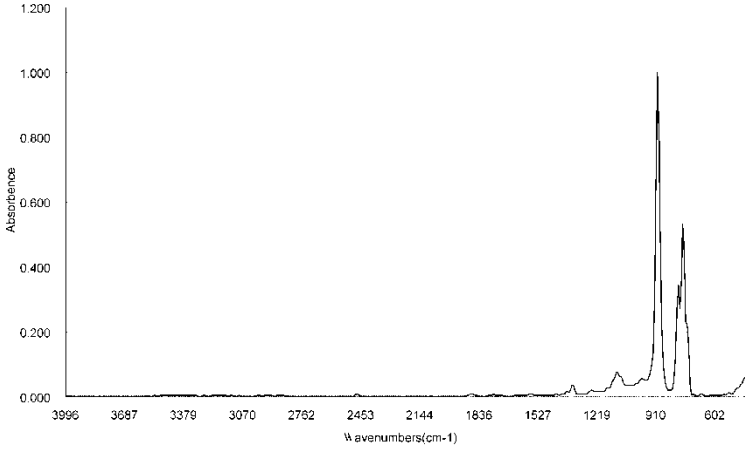


Figure 3. The IR spectra of tetrachloroethylene.

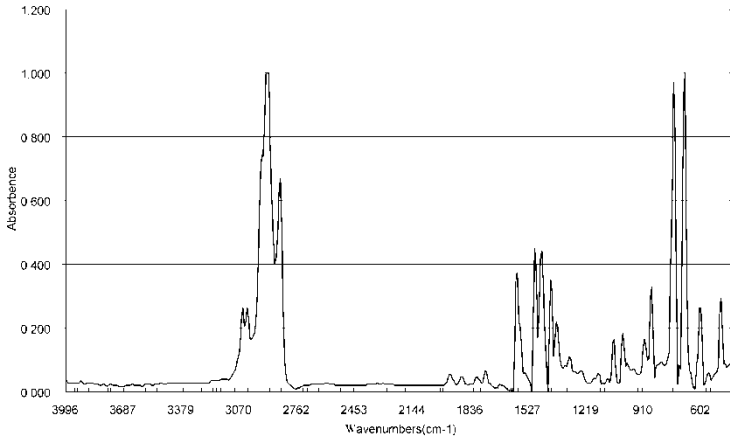


Figure 4. The IR spectra of 1,3,5-triphenylbenzene.

expected that a higher CR can be achieved if we compress the IR spectrum with data length longer than 460.

CONCLUSIONS

The algorithm with the use of EZW and the Huffman coding method was proposed in this work to compress IR spectra of a few compounds. This

Table 4. Results of comparison

	Method I		Method II		
Spectra	CR	RMSD	CR	RMSD	Percentage
Tridecane, 99+%	8.43	0.0040	6.39	0.0042	31.9
(1*S*)-(−)-*b*-Pinene, 98%	6.23	0.0068	3.93	0.0068	58.5
1-Chlorohexadecane, 99%	7.71	0.0044	5.54	0.0044	39.2
1,5-Diiodopentane, 97%	6.79	0.0043	4.04	0.0043	68.1
Pentachloroethane, 96%	8.43	0.0037	6.67	0.0037	26.4
Tetrachloroethylene, 99+%	11.1	0.0055	11	0.0055	0.909
1,3,5-Triphenylbenzene, 97%	5.35	0.0058	3.24	0.0058	65.1
3-Chlorobenzyl chloride, 98%	6.15	0.0060	3.74	0.0060	64.4
Decyl alcohol, 99%	7.71	0.0056	5.97	0.0056	29.1
4-*t*-Butylcyclohexanol, mixture of isomers	5.91	0.0052	3.71	0.0052	59.3
Average	7.381	0.00513	5.423	0.00515	36.1
Standard deviation	1.691	0.000994	2.323	0.000971	

CR, compression ratio; RMSD, root-mean-square derivative.

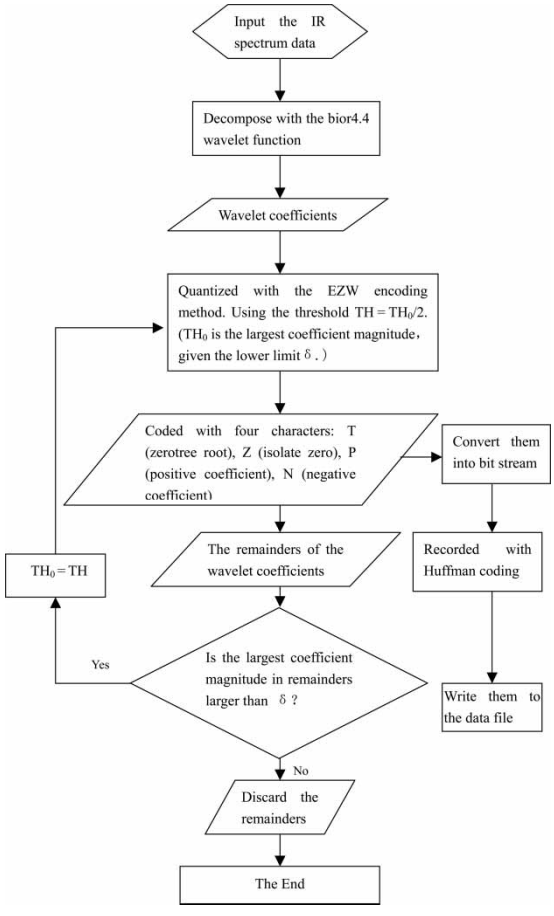
kind of technique sufficiently takes advantage of not only the property that there are many small-amplitude values in wavelet coefficients but also the self-similarities of wavelet coefficients at various resolutions. The results showed that the performance of the proposed compression method is much better than those suggested in previous studies.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 29877016).

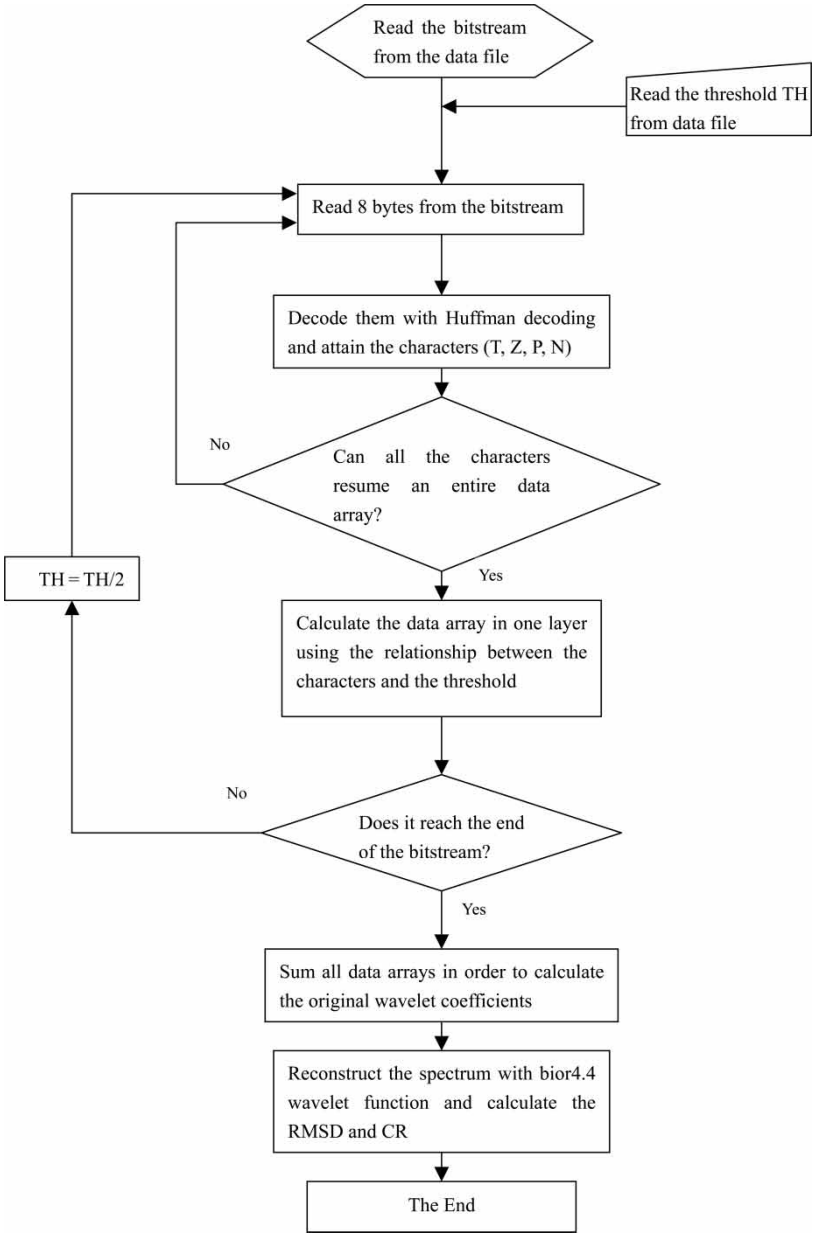
APPENDIX 1

A Diagram Showing the Compression Procedure of IR



APPENDIX 2

A Diagram Showing the Decompression Procedure of IR



REFERENCES

1. Shao, X. G.; Leung, A. K. M.; Chau, F. T. Wavelet: A new trend in chemistry. *Acc. Chem. Res.* **2003**, *36*, 276–283.
2. Leung, A. K.M.; Chau, F. T.; Gao, J. B. A review on applications of wavelet transform techniques in chemical analysis. *Chemom. Intell. Lab. Syst.* **1998**, *43* (1-2), 165–184.
3. Shao, X. G.; Yu, Z. L.; Sun, L. Resolution of multicomponent NMR signals using wavelet compression and immune algorithm. *Spectrochim. Acta A* **2003**, *59*, 1075–1082.
4. Yoon, J.; Lee, B.; Han, C. Calibration transfer of near-infrared spectra based on compression of wavelet coefficients. *Chemom. Intell. Lab. Syst.* **2002**, *64*, 1–14.
5. Urbas, A. A.; Harrington, P. B. Two-dimensional wavelet compression of ion mobility spectra. *Anal. Chim. Acta* **2001**, *446*, 393–412.
6. Depczynski, U.; Jetter, K.; Molt, K.; Niemöller, A. The fast wavelet transform on compact intervals as a tool in chemometrics: II. Boundary effects, denoising and compressing. *Chemom. Intell. Lab. Syst.* **1999**, *49*, 151–161.
7. Ho, H. L.; Cham, W. K.; Chau, F. T.; Wu, J. Y. Application of biorthogonal wavelet transform to the compression of ultraviolet-visible spectra. *Comput. Chem.* **1999**, *23*, 85–96.
8. Barclay, V. J.; Bonner, R. F. Application of wavelet transforms to experimental spectra: smoothing, denoising, and data set compression. *Anal. Chem.* **1997**, *69* (1), 78–90.
9. Leung, A. K.M.; Chau, F. T.; Gao, J. B.; Shih, T. M. Application of wavelet transform in infrared spectrometry: spectral compression and library search. *Chemom. Intell. Lab. Syst.* **1998**, *43* (1–2), 69–88.
10. Chau, F. T.; Shih, T. M.; Gao, J. B.; Chan, C. K. Application of the fast wavelet transform method to compress ultraviolet-visible spectra. *Appl. Spectrosc.* **1996**, *50* (3), 339–349.
11. Shapiro, J. M. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. Signal Processing* **1993**, *41* (12), 3445–3462.
12. Chen, G. X.; Harrington, P. B. Real-time two-dimensional wavelet compression and its application to real-time modeling of ion mobility data. *Anal. Chim. Acta* **2003**, *490*, 59–69.
13. Abo-Zahhad, M.; Rajoub, B. A. An effective coding technique for the compression of one-dimensional signals using wavelet transforms. *Med. Eng. Phys.* **2002**, *24*, 185–199.
14. Daubechies, I. Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* **1988**, *41*, 909–996.
15. Mallat, S. A theory of multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693.