

This item was submitted to Loughborough's Institutional Repository by the author and is made available under the following Creative Commons Licence conditions.



creative commons  
COMMONS DEED

**Attribution-NonCommercial-NoDerivs 2.5**

**You are free:**

- to copy, distribute, display, and perform the work

**Under the following conditions:**

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.

 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

**Your fair use and other rights are in no way affected by the above.**

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

# WAVELET COMPRESSION TECHNIQUES FOR COMPUTER NETWORK MEASUREMENTS

Konstantinos G. Kyriakopoulos, Prof. David J. Parish  
Loughborough University, Electronic and Electrical Engineering, High Speed Networks  
Loughborough, Leicestershire, LE11 3TU  
United Kingdom  
elkk@lboro.ac.uk

## ABSTRACT

Wavelet transform is a recent signal analysis tool that is already been successfully used in image, video and speech compression applications. This paper looks at the Wavelet transform as a method of compressing computer network measurements produced from high-speed networks. Such networks produce a large amount of information over a long period of time, requiring compression for archiving. An important aspect of the compression is to maintain the quality in important features of signals. In this paper two known wavelet coefficient threshold selection techniques are examined and utilized separately along with an efficient method for storing wavelet coefficients. Experimental results are obtained to compare the behaviour of the two threshold selection schemes on delay and data rate signals, by using the mean square error (MSE) statistic, PSNR and the file size of the compressed output.

## KEY WORDS

Wavelets, thresholds, compression, computer networks, measurements

## 1. Introduction

This paper is motivated by the need to monitor the performance of high-speed communication networks of the future and particularly of the UKLight experimental network. The UKLight initiative is a 10 Gb/s, high capacity research network facility that interconnects JANET, the UK's research and educational network, with several other continental research networks [1].

Monitoring for a long period of time such high-speed networks produces a high volume of data, making the storage of this information practically inefficient. To this end, there is a need to derive an efficient method of data analysis and reduction in order to archive and store the enormous amount of monitored traffic [2].

Satisfying this need is useful for researchers who run their experiments on the monitored network. The researchers would like to know how their experiments affect the network's behavior in terms of utilization, delay, packet loss, data rate etc. Such signals are derived from packet information and can be represented as a time series

process. Thus, these signals can be analyzed and compressed with a signal analysis technique. By compressing and storing such measurements, researchers have the ability to know how their algorithms behaved at a particular moment in the past or at present.

In this paper, Wavelet analysis is applied to network delay and data rate measurements in order to compress the size of the information without reducing the quality in important features of the signal. The data rate signals are from a real network and the delay signals are generated in HSN's test bed. Experimental results are obtained to compare two different thresholding techniques by using the mean square error (MSE), PSNR and the compression ratio (C.R) between the original and reconstructed signals.

The rest of the paper is structured as follows. In section 2 the concepts of Wavelet analysis, Wavelet compression and thresholding are introduced. In Section 3 the methodology of this work is presented with particular emphasis on the two examined threshold selection techniques. Section 4 discusses the results of the two proposed methods after being applied to thirty delay and thirty data rate measurement signals. Finally, the conclusions and the future work are given in Section 5.

## 2. Wavelets, Thresholds and Denoising

### 2.1 Advantages of wavelets

The Heisenberg uncertainty principle suggests that it is impossible to know the exact frequency and the exact time of occurrence of this frequency in a signal but it is possible to obtain the frequency bands that exist in a time interval. In other words, there is a trade off between the resolution of the time and the frequency domain [3, 4].

In contrast with other techniques that use a constant window size to analyze a section of a signal (for example DCT, STFT), wavelet analysis has the benefit of varying the window size. This means that wavelets can efficiently trade frequency resolution for time resolution or vice versa. For that reason, wavelets can adapt to various time-scales and perform *local analysis*. With simple words, wavelets can reveal both the forest and the trees [5, 6].

The local analysis feature of the wavelets provides the additional benefit of approximating an examined signal compactly, i.e. with few coefficients, making wavelets an appropriate choice for compression applications [6].

The finite nature of the wavelet can describe local features of the signal better than the infinite length of a sinusoid. Thus, another attribute of wavelet analysis is the ability to detect characteristics of non-stationary signals, i.e. stochastic (random) signals whose statistical properties change with time. Most interesting signals are non-stationary signals [7,8].

## 2.2 Wavelet analysis

Wavelet analysis is not a compression tool but a transformation to a domain that provides a different view of the data that is more eligible to compression than the original data itself. Following the analysis stage, classical compression techniques can be applied on the produced wavelet coefficients [7].

The level of wavelet decomposition determines the scale of detail that is subtracted from the analysed signal. As the depth of decomposition increases, a cruder approximation of the signal is analysed and the detail wavelet coefficients correspond to a larger scale of detail. If the scale of detail exceeds a limit, then the analysed signal becomes distorted [7].

## 2.3 Achieving compression

What really matters in the compression scheme is the sequence of zero coefficients. Zero coefficients should be gathered sequentially rather than spread and being separated by non-zero values. If all zero coefficients are gathered together then by applying Run Length Encoding (RLE) we can take advantage of the repetitive values and achieve compression.

Many of the detail coefficients produced from the wavelet analysis have an absolute value close to zero. These small coefficients are likely attributed to the noise of the signal, while large coefficients represent important characteristics of the signal. The small coefficients contain a small percentage of the signal's total energy and can be discarded without a significant loss in the quality of the signal and more importantly of the interesting features [7,8,9].

The detail coefficients can be discarded by applying a threshold that sets to zero all coefficients that are less than this threshold. Thus the sequence of zeros is increased while an insignificant amount of the signal's energy is lost. A higher threshold would yield a better compression but a greater loss in the signal's quality [7].

## 3. Methodology

The connection between lossy compression and denoising has been discussed in several papers, like [9]

and references therein. In this work, denoising techniques are used in order to achieve compression.

The compression - decompression algorithm is given below:

1. Perform Wavelet analysis at multiple levels on the examined signal
2. For each decomposition level select the detail coefficient threshold as discussed in section 3.1
3. Apply the threshold on the detail coefficients
4. Normalize the coefficients as discussed in section 3.2
5. Apply Run Length Encoding (RLE)
6. Apply inverse RLE
7. De-normalize coefficients
8. Perform Inverse Wavelet transform to reconstruct the signal

### 3.1 Threshold selection

Even though a lot of research has been done in selecting a threshold [5, 7-13], most of it is focused on recovering signals that have been corrupted by additive Gaussian noise. In this paper, the examined signals, network delay and data rate measurements, are not affected by noise, thus a threshold selection scheme that depends on the values of the wavelet coefficients has to be deployed.

The first scheme that is examined is proposed by Birge and Massart and has become a popular threshold selection technique used widely in image and speech compression [5].

The scheme depends on three parameters:

1. The level of decomposition  $J$
2. A positive constant  $M$
3. A sparsity parameter  $a$

This scheme keeps all approximation coefficients at the level of decomposition  $J$ . At each level  $i$  only the  $n_i$  largest coefficients are kept.  $n_i$  is estimated by this formula:

$$n_i = \frac{M}{(J + 2 - i)^a} \quad (1)$$

Usually,  $a$  takes the value 1.5 for compression and  $M$  depends on how scarcely the wavelet coefficients are spread and on the number  $L$  of approximation coefficients in the coarsest level (Table 1). For highly scarcely spread coefficients,  $M$  becomes equal to the number  $L$ . For low scarcely spread coefficients  $M$  becomes twice the number  $L$ .

Scarce	M value
High	$L$
Medium	$1.5 * L$
Low	$2 * L$

Table 1:  $M$  depends on how scarce the coefficients are spread.

For delay signals the high scarcity option of the Birge Massart algorithm was used. For data rate the low scarcity was more appropriate as more coefficients were required for more precise reconstruction.

The second of the examined threshold selection techniques was recently proposed by Gupta and Kaur [14]. They proposed an adaptive thresholding technique that is calculated from the value of the wavelet coefficients. Specifically, the standard deviation ( $\sigma$ ) and mean ( $\mu$ ) of the absolute value of non-zero detail coefficients is first calculated. If the standard deviation is larger than the mean, then the threshold is set to two times the mean ( $2*\mu$ ), otherwise it is equal to the mean minus the standard deviation ( $\mu-\sigma$ ).

In the following paragraphs the first technique will be referred to as BM technique and the second as GK technique from the initials of the names of their developers.

### 3.2 Normalization

In order to improve the way that data is stored, normalization of the coefficients takes place. The aim is to use just 8 bits to store one coefficient. But with an 8 bit variable only 256 values can be stored (0...255) or 127 values ( $2^7$  bits) saving one bit for the sign of the wavelet coefficient. So, first the coefficient values have to be normalised using the following formula:

$$\text{norm}(x) = \text{round}\left(\frac{x - \min}{\max - \min} * \text{scaling factor}\right) \quad (2)$$

where  $x$  is a coefficient,  $\min$  is the minimum value that appears in the coefficients and  $\max$  the maximum. The scaling factor in this case is 127, which is the maximum value of a signed number that can be stored in 1 byte. In order to avoid the detail coefficients to be skewed by the larger values of the approximation coefficients, the normalization process is applied separately for the detail and for the approximation coefficients.

### 3.3 Run Length Encoding

The simplest version of the RLE algorithm replaces a sequentially repetitive symbol with the symbol itself followed by a number that indicates how many times the symbol should be repeated. However, this simple version of RLE expands single symbols into a pair of symbol-run length.

In order to avoid this shortcoming, a more sophisticated RLE implementation utilises a run length that is used only for symbols that appear more than 2 times. This method has beneficial effect only for symbols that appear 3 or more times [15]. However, the RLE limitation persists for symbols that appear sequentially for just two instances and expand into symbol-symbol-run length triple.

## 4. Results and Discussion

For the purpose of examining the behaviour of the applied threshold selection techniques 30 delay and 30 data rate measurement signals are examined. Each signal has 1024 measurement points and is decomposed at all possible levels, level 1 through level 10, using the Haar wavelet. The following PSNR, MSE and C.R. values refer to the average over those 30 signals (delay or data rate) except when explicitly noting for which signal they refer.

In order to examine if the normalization and RLE steps of the coefficient values in the compression scheme have any negative effect on the error of the reconstructed signal, the same experiments as above were repeated but without applying the normalization and RLE steps. The MSE is calculated for relative comparison between the cases of applying the normalization and RLE steps and of omitting them.

The results show that for most cases the average error of compression performs similarly either if normalization and RLE are applied or not. The only exception happens with the data rate signals examined with the GK technique. The reason is discussed later in section 4.2.

MSE is also used to indicate how the error increases with the increase of the decomposition level. MSE is calculated by

$$MSE = \frac{1}{N} \sum_{i=0}^{N-1} |x_i - \bar{x}_i|^2$$

Its value itself is not of concern, as it does not reveal much about the quality of the reconstructed signal. The quality of the reconstructed signal is compared with the original by using the PSNR value calculated by

$$PSNR = 10 * \log\left(\frac{MAX^2}{MSE}\right)$$

where MAX is the maximum value of the signal.

In addition to PSNR values there are also figures of some signals with the original and the reconstructed signal. The error is also provided for easier understanding of the magnitude of error.

### 4.1 Results for Birge Massart (BM) technique

Fig. 1 and 2 show the compression performance and the error of compression for the delay signals with respect to the level of decomposition. As the level of decomposition increases, the number of approximation coefficients decreases and so does the number of kept detail coefficients as can be inferred from equation (1).

This has as a result the MSE and C.R. to become very high above level 4, giving an average PSNR of 34.2 dB at level 4. PSNR values less than 35 dB loose some of the important signal characteristics while PSNR values less than 30 dB are not acceptable for such signals. Table 2 includes PSNR and C.R. values for the first four levels where PSNR remains above 30 dB.

The first two levels of decomposition give good PSNR values (above 40 dB) and perform very well for

almost all signals. Fig. 3 shows signal 24, its reconstruction and the error after analysis at level 2.

Signal 10 of the experiments is the only signal from the delay measurements that is much more bursty than the rest. BM technique fails to retain the quality of that signal even in the lowest level of decomposition giving a PSNR of 25.7 db at level 1 (Fig. 4). Such bursty signals require more coefficients to be kept in order to preserve their quality and thus a more appropriate algorithm.

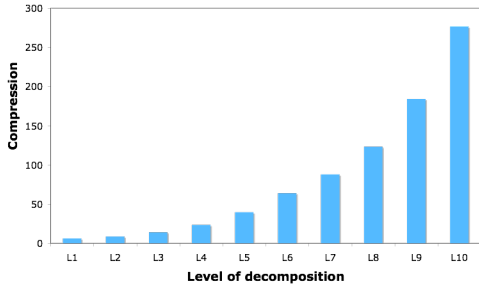


Fig. 1: Performance of compression for delay signals with BM technique

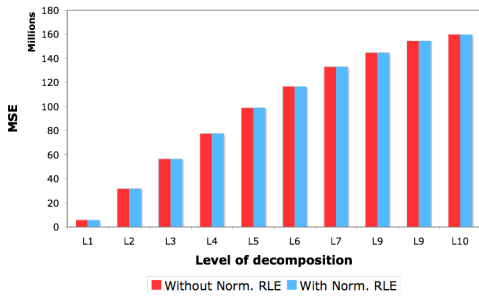


Fig. 2: Error of compression for delay signals with BM technique

Level	L1	L2	L3	L4
PSNR	51.2	43	38.5	34.2
C.R.	5.3	7.9	13.1	22.6

Table 2: PSNR and C.R. values for levels 1-4 of decomposition for delay signals with the BM technique

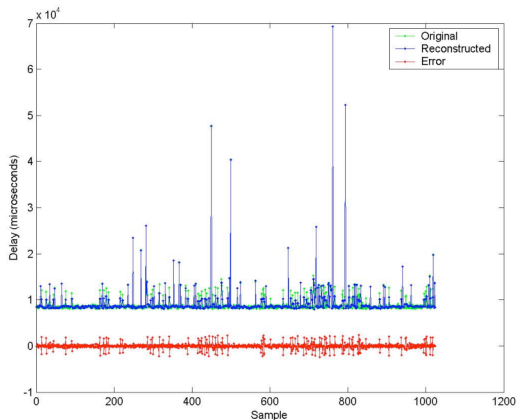


Fig. 3: Delay signal 24 decomposed at level 2 with BM method, PSNR = 40.3 dB

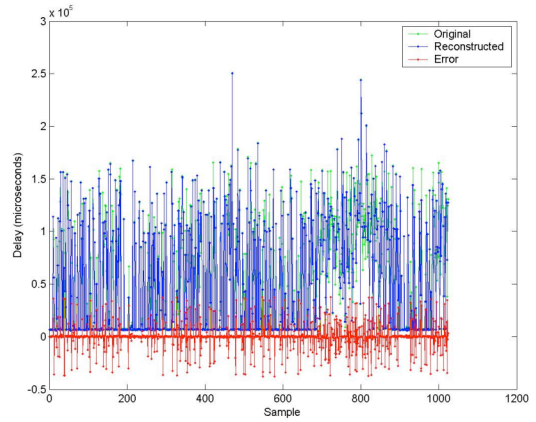


Fig. 4: Delay signal 10 decomposed at level 1 with BM method, PSNR = 25.7 dB

Regarding the Data Rate signals, the performance of compression behaves as in delay signals. The significant increase in error in level 10 (Fig. 6) occurs because the nominator in equation (1) is very small which makes the number of kept detail coefficients in each level to be very small. Compression ratios and PSNR values for the first 4 levels are given in Table 3. From level 5 and above the average PSNR values become less than 30 dB and are not included in Table 3

BM technique behaves the same in all cases of Data rate signals keeping close values of PSNR for all signals ranging from 49 dB to 52.8 dB in level 1 and 35.7 db to 40.7 in level 2. Fig. 7 shows signal 16, its reconstruction with PSNR=40.7 dB and the error after analysis at level 2.

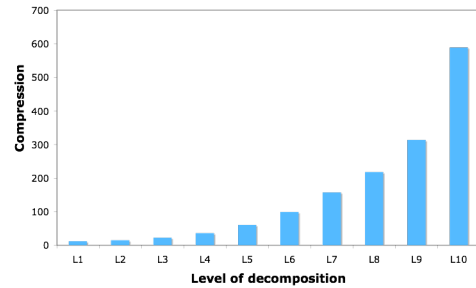


Fig. 5: Performance of compression for data rate signals with BM technique

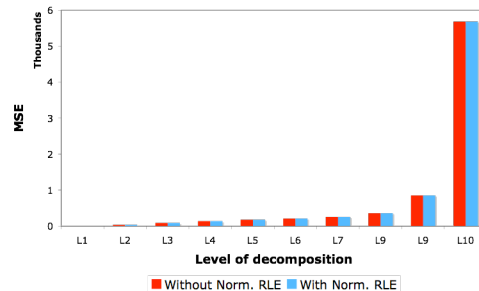


Fig. 6: Error of compression for data rate signals with BM technique

Level	L1	L2	L3	L4
PSNR	49	35.7	32.2	30.4
C.R.	10.3	13.2	20.5	34.2

Table 3: PSNR and C.R. values for levels 1-4 of decomposition for data rate signals with the BM technique

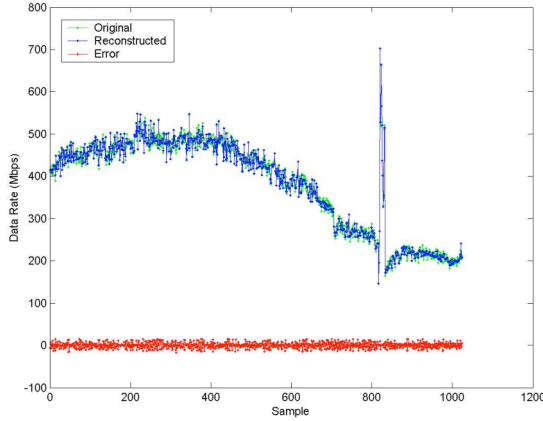


Fig. 7: Data rate signal 16 decomposed at level 2 with BM method, PSNR= 40.7 dB

## 4.2 Results for Gupta Kaur (GK) technique

First the results for delay signals using the GK technique are discussed and then the results of data rate signals.

Fig. 8 shows the performance of compression and Fig. 9 the average MSE of the reconstructed signals over the depth of decomposition. The C.R. begins at 7.5 in level 1 and stabilizes around 17 from level 6 and above. The MSE stabilizes after level 7. The PSNR and C.R. average values for all levels are in Table 4a and 4b.

In contradiction to the BM technique, GK performs very well in keeping the reconstructed quality of signal 10 even at the crudest level. This happens because GK is based on the statistical characteristics of the coefficients to determine the number of coefficients that should be kept. The results can be seen in Fig. 10.

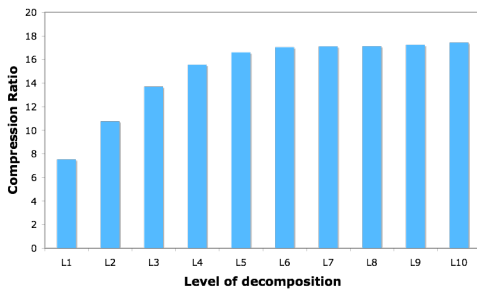


Fig. 8: Performance of compression for delay signals with GK technique

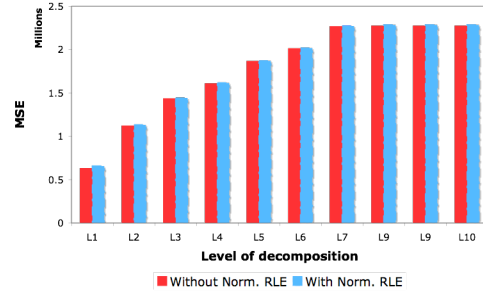


Fig. 9: Error of compression for delay signals with GK technique

Level	L1	L2	L3	L4	L5
PSNR	44.3	42	41	40.5	40
C.R.	7.5	10.7	13.6	15.5	16.5

Table 4 (a): PSNR and C.R. values for levels 1-5 of decomposition for delay signals with the GK technique

Level	L6	L7	L8	L9	L10
PSNR	39.6	39.2	39.2	39.2	39.2
C.R.	17	17	17	17.2	17.4

Table 4 (b): PSNR and C.R. values for levels 6-10 of decomposition for delay signals with the GK technique

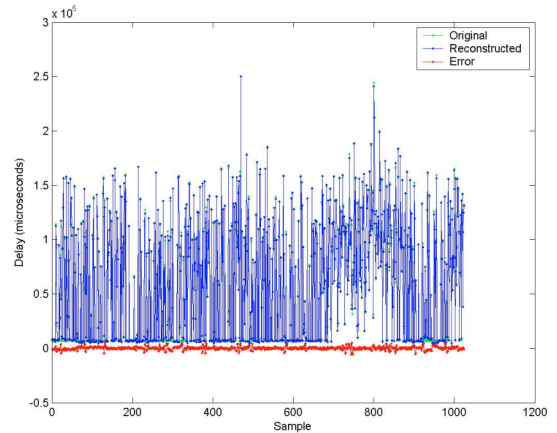


Fig. 10: Delay signal 10 decomposed at level 10 with GK method, PSNR= 44.3 dB

In contrast with the delay signals, for the data rate signals there is no significant increase in the compression ratio as the level of decomposition increases (Fig. 11). The C.R. range is between 10.5 and 11.2 in contrast with the wider range (7.5 – 17.4) for delay signals (see Table 4a b). This happens because data rate signals have a lot of high frequency components that make the GK algorithm to keep a lot of detail coefficients. The PSNR and C.R. average values for all levels of decomposition are in Table 5a and 5b.

An interesting implication of the normalization and RLE steps with the wavelet coefficients of data rate signals is that the compression performance does not stabilize as it happens with the delay signals. In particular,

there is a decrease in C.R. after level 5 and a sudden peak at level 10.

This is because by increasing the level of decomposition in data rate signals, some of the produced coefficients are much larger than the rest of coefficients. In other words the dynamic range of the detail coefficients is being increased. For this reason, after the normalization step, close values are assigned the same normalized value. This phenomenon happens occasionally across the coefficients taking advantage of the RLE limitation and producing file sizes that are larger than the files before the RLE step.

In the highest decomposition level, the maximum detail coefficient value is so large that normalizes many detail coefficients to the value of zero (because denominator in equation (2) is very high). In that case the RLE will take advantage of the repeating values but also the MSE will increase because of the loss of detail coefficients (Fig. 12).

This is the only case that the normalization and RLE steps increase significantly the error of the reconstructed signal. However, this is only limited in the last three levels of decomposition (Fig. 12).

Fig. 13 shows an example of a data rate signal decomposed at level 5 with the GK method. The reconstructed signal has a very good quality with PSNR=56.9 db and very low error.

Fig. 14 shows a more interesting case of a data rate signal. This signal includes a spike, which is kept intact after the compression. A characteristic of the GK algorithm is that it detects the spike as a more interesting feature than the rest of the signal.

As a result, the algorithm's first priority becomes to preserve this characteristic and then comes the rest of the signal. That is the reason why PSNR is around 35 and there is much higher error in comparison to Fig. 13.

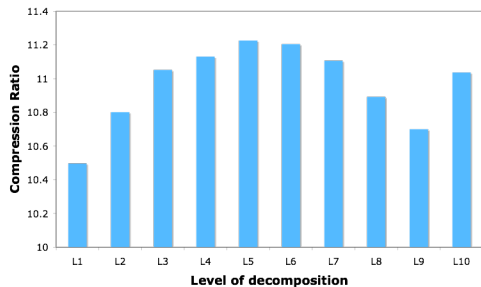


Fig. 11: Performance of compression for data rate signals with GK technique

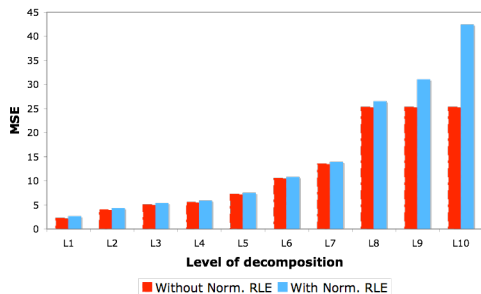


Fig. 12: Error of compression for data rate signals with GK technique

Level	L1	L2	L3	L4	L5
PSNR	56.3	55.6	55.4	55.2	54.9
C.R.	10.5	10.8	11	11.1	11.2

Table 5 (a): PSNR values for levels 1-5 of decomposition for data rate signals with the GK technique

Level	L6	L7	L8	L9	L10
PSNR	53.4	49	43.9	42.1	39.4
C.R.	11.2	11.1	10.9	10.7	11

Table 5 (b): PSNR values for levels 6-10 of decomposition for data rate signals with the GK technique

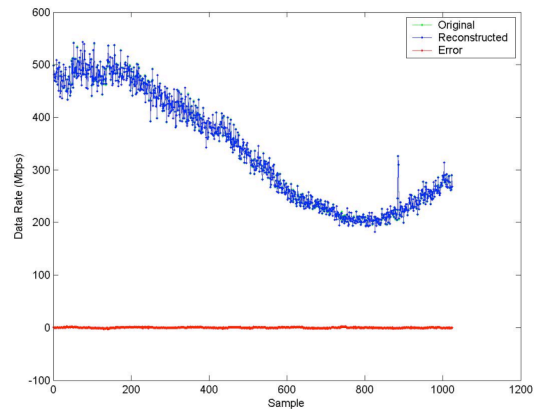


Fig. 13: Data rate signal 20 decomposed at level 5 with GK method, PSNR = 56.9

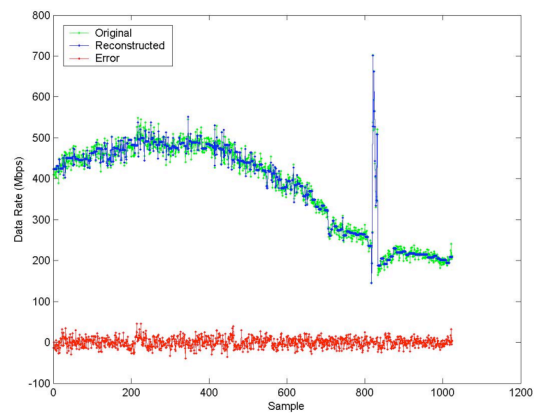


Fig. 14: Data rate signal 16 decomposed at level 5 with GK method, PSNR = 35.4 dB

## 5. Conclusions – Future work

This paper implements Wavelet based denoising in order to achieve lossy compression of network delay and data rate measurements while maintaining the characteristic features of the examined signals. Two techniques of coefficient threshold selection are utilized and their behaviour on these types of signals is examined.

In general, the BM technique increases the MSE and C.R. as the level of decomposition increases. On the other

hand, the GK method restricts C.R. in higher levels of decomposition while keeping the quality of the reconstructed signals at reasonable levels.

For the delay signals, the BM technique gives better PSNR than the GK only in the first two levels. However, the offered C.R. for both of those levels is lower than the one offered by GK (see Tables 2, 4a).

For data rate signals the BM technique is not giving better average PSNR, than GK even in the first level. The same applies for the C.R. However, it gives more consistent PSNR values for data rate signals (see Tables 3, 5a).

The GK method is more appropriate for both types of signals as it offers more reasonable C.R. and good PSNR values even when reaching high levels of decomposition. It can adapt to bursty signals like in the case of signal 10 (Fig. 10) and it does not require any parameter like BM. The reconstructed signals preserve quality on interesting features while smoothing out the detail information in non-significant parts.

However, some improvements should be done in how the algorithm deals with the threshold in cases that spikes occur in an already bursty signal like in signal 16 (Fig. 14). This would lead to more control over the quality of the reconstructed signal.

The GK algorithm is already implemented in CoMo. CoMo is a passive monitoring platform developed for the purpose of monitoring network links at high speeds and replying to real time queries [16, 17].

CoMo has various modules that each calculates one or more measurements. The proposed algorithm is imbedded in the modules and compresses these measurements. When CoMo receives a query, the information is first decomposed and then shown to the end user [16, 17].

## 6. References

[1] Press Release: JISC announces UKLight, a £6.5 million networking initiative, [http://www.jisc.ac.uk/index.cfm?name=uk\\_light\\_pr](http://www.jisc.ac.uk/index.cfm?name=uk_light_pr), Page last visited 30/04/06

[2] UKLight: Case for support, <http://www.ee.ucl.ac.uk/~lsacks/acse/masts/masts.pdf>, Page last visited 30/04/06

[3] C. Valens, A really friendly guide to wavelets, <http://perso.orange.fr/polyvalens/clemens/wavelets/wavelets.html>, Page last visited: 28/08/06

[4] Deepika Sripathi, "Efficient Implementations of Discrete Wavelet Transforms using FPGAs", MSc Thesis, Florida State University College of Engineering, 2003, [http://etd.lib.fsu.edu/ETD-db/ETD-browse/browse?first\\_letter=S](http://etd.lib.fsu.edu/ETD-db/ETD-browse/browse?first_letter=S), Page last visited 28/08/06

[5] M. Misiti, Y. Misiti, G. Oppenheim and J. Poggi, "Matlab Wavelet Toolbox",

[http://www.mathworks.com/access/helpdesk/help/pdf\\_doc/wavelet/wavelet\\_ug.pdf](http://www.mathworks.com/access/helpdesk/help/pdf_doc/wavelet/wavelet_ug.pdf), Page last visited 28/08/06

[6] Johnson Ihyeh Agbinya, Discrete wavelet transform techniques in speech processing. *IEEE TENCON - Digital Signal Processing Applications*, pages 514--519, 1996

[7] BaseGroup Lab, "Wavelet Analysis Applications", [http://www.basegroup.ru/filtration/wavelet\\_applications\\_en.htm](http://www.basegroup.ru/filtration/wavelet_applications_en.htm), Page last visited 30/04/06

[8] L. Kaur, S. Gupta, R.C.Chauhan, "Image denoising using wavelet thresholding", Indian Conference on computer Vision, Graphics and Image Processing, Ahmedabad, Dec. 2002.

[9] S. Grace Chang, Bin Yu and M. Vattereli, "Adaptive Wavelet Thresholding for Image Denoising and Compression", *IEEE Trans. Image Processing*, vol. 9, pp. 1532-1546, Sept. 2000.

[10] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, pp. 613-627, May 1995.

[11] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425-455, 1994.

[12] D. L. Donoho and I. M. Johnstone "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Assoc.*, vol. 90, no. 432, pp. 1200-1224, December 1995.

[13] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: Asymptopia?", *J. R. Stat. Soc. Ser. B*, vol. 57, pp. 301-369, 1995.

[14] Savita Gupta and Lakhwinder Kaur, "Wavelet Based Image Compression using Daubechies Filters", In proc. 8th National conference on communications, I.I.T. Bombay, NCC-2002

[15] Run Length Encoding Discussion and Implementation <http://michael.dipperstein.com/rle/index.html>, Page last visited on 28/08/06

[16] Intel Research Cambridge [http://www.intel.com/research/network/cambridge\\_collab\\_p2.htm](http://www.intel.com/research/network/cambridge_collab_p2.htm), Page last visited 28/08/06

[17] Gianluca Iannaccone, Christopher Diot, Derek McAulley, Andrew Moore, Ian Pratt, Luigi Rizzo, "The CoMo White Paper", INTEL research technical report