

# Training Methods for Image Noise Level Estimation on Wavelet Components

## A. De Stefano

*Institute of Sound and Vibration Research, University of Southampton, Highfield, Hants SO17 1BJ, UK*  
Email: [ads@isvr.soton.ac.uk](mailto:ads@isvr.soton.ac.uk)

## P. R. White

*Institute of Sound and Vibration Research, University of Southampton, Highfield, Hants SO17 1BJ, UK*  
Email: [prw@isvr.soton.ac.uk](mailto:prw@isvr.soton.ac.uk)

## W. B. Collis

*The Foundry, 35-36 Great Marlborough Street, London W1F 7JE, UK*  
Email: [bill@thefoundry.co.uk](mailto:bill@thefoundry.co.uk)

*Received 25 July 2003; Revised 14 January 2004*

The estimation of the standard deviation of noise contaminating an image is a fundamental step in wavelet-based noise reduction techniques. The method widely used is based on the mean absolute deviation (MAD). This model-based method assumes specific characteristics of the noise-contaminated image component. Three novel and alternative methods for estimating the noise standard deviation are proposed in this work and compared with the MAD method. Two of these methods rely on a preliminary training stage in order to extract parameters which are then used in the application stage. The sets used for training and testing, 13 and 5 images, respectively, are fully disjoint. The third method assumes specific statistical distributions for image and noise components. Results showed the prevalence of the training-based methods for the images and the range of noise levels considered.

**Keywords and phrases:** noise estimation, training methods, wavelet transform, image processing.

## 1. INTRODUCTION

Noise reduction plays a fundamental role in image processing, and wavelet analysis has been demonstrated to be a powerful method for performing image noise reduction [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. The procedure for noise reduction is applied on the wavelet coefficients achieved using the wavelet decomposition and representing the image at different scales. After noise reduction, the image is reconstructed using the inverse wavelet transform. Decomposition and reconstruction are accomplished using two banks of filters constrained by a perfect reconstruction condition [3, 13]. The structure of these filter banks is characterised by the frequency responses of two filters and by the presence or absence of sub/up-sampling, generating, respectively, decimated or undecimated wavelet transforms. Undecimated wavelet transforms have been considered for image noise reduction [3, 4, 5, 6, 10, 11, 12] as well as the decimated transforms [1, 2, 6, 7, 8, 9].

Whilst alternative techniques have been proposed [14, 15, 16, 17, 18], the technique most widely used to reduce the

noise on the wavelet coefficients is to use one of a parameterised family of nonlinear functions, also called scheme. Schemes that have been proposed include soft thresholding [1, 6, 7], hard thresholding [2, 9], and optimal schemes [3, 11, 12, 19, 20, 21]. The parameters identifying the nonlinear functions to be applied on each scale depend on the characteristics of the image component and on the noise. Several techniques have been presented to estimate these parameters based on the median operator or on the histogram of the wavelet transform [6, 22]. Other techniques, such as the min-max threshold, global universal threshold, Sure threshold, and James-Stein threshold, have been proposed in numerous works [1, 2, 7]. Experiments performed to compare the performances of these techniques [23, 24] demonstrated that it is not possible to say which is the best, even if the global universal threshold appears to be the worst. Usually these techniques assume the knowledge a priori of the noise standard deviation level; therefore its correct estimation dramatically affects the performances of the noise reduction technique. Donoho proposed a robust noise level estimator: the mean absolute deviation (MAD) of wavelet coefficients at the

highest resolution [1, 2, 6, 7]. Nevertheless, in our literature research we did not find alternative methods addressed to this problem.

In this paper, we include an investigation of the problem of estimating the variance of the noise contaminating an image and we compare three novel algorithms, two of which based on training over a set of test images, with the MAD technique.

The training process extracts a number of parameters using a set of noise-contaminated images generated by synthetically combining noise-free images with realisations of the noise process with a given level (standard deviation). During the application of the algorithm, only a noisy image is available for analysis, from which the level of the contaminating noise is estimated employing the parameters extracted during the training.

In Section 2, we present the three new techniques for noise level estimation over the wavelet components. The results are described and commented upon in Section 3, and conclusions are drawn in Section 4.

## 2. NOISE LEVEL ESTIMATION

The most widely used method for estimating the variance of the noise on a wavelet component is the mean absolute deviation (MAD). This scheme tends to overestimate the noise standard deviation in applications where the SNR in the wavelet components is high, leading to unnecessary distortion of the image. The tendency for MAD to overestimate the noise level is due to the fact that it assumes that the image contribution in the band of interest can be neglected. However, the fact that MAD is based on absolute deviations makes it more robust to outliers (arising through image contributions in the band) than, say, direct estimate of the standard deviation.

This section presents three alternative methods for estimating the standard deviation of the noise from a noisy image. These methods are based on the assumption that the noise is Gaussian and additive. The methods can be applied to any of the wavelet components of the image. However, their performance degrades when the signal-to-noise ratio (SNR) in the component increases, so in practice one usually finds that the noise variance is most accurately estimated on the smallest scale (highest frequency) component where, in most cases, the SNR is the lowest. If one is willing to make the assumption of spatially white noise, then knowledge of the noise variance at the smallest scale allows the one to infer the noise variance at all other scales.

### 2.1. Model-based estimation of the noise variance

One method of estimating the noise variance is to assume a model for both the noise and image components and to fit the data to this model. One model, consistent with the use of a soft thresholding scheme, is to assume that the noise is additive and Gaussian and the image has a Laplacian distribution. If the Laplacian distribution has zero mean and standard deviation  $\sigma_u$ , then its probability distribution function

(pdf) is

$$p_{\text{im}}(x) = \frac{1}{\sigma_u \sqrt{2}} e^{-(\sqrt{2}/\sigma_u)|x|}. \quad (1)$$

The pdf of the image plus Gaussian noise (standard deviation  $\sigma_v$ ) is [3]

$$p(x) = \frac{e^{(\sigma_v/\sigma_u)^2}}{\sigma_v \sqrt{2}} \left\{ e^{(\sqrt{2}/\sigma_u)x} \operatorname{erfc}\left(\frac{\sigma_v}{\sigma_u} + \frac{x}{\sigma_v \sqrt{2}}\right) + e^{-(\sqrt{2}/\sigma_u)x} \operatorname{erfc}\left(\frac{\sigma_v}{\sigma_u} - \frac{x}{\sigma_v \sqrt{2}}\right) \right\}, \quad (2)$$

where  $\operatorname{erfc}(x)$  is the complementary error function. The problem is then to estimate the parameters  $\sigma_u$  and  $\sigma_v$  from the observed pixel values. An optimal method to achieve this is to employ the method of maximum likelihood (ML). In this problem, ML leads to a solution with no closed analytic form and one is faced with an optimisation task. The absence of a sufficient statistic for this problem makes the computation of the ML solution burdensome. At every iteration of the optimisation, one is required to evaluate (2) for every pixel in the image. In an off-line environment, this load may not be too onerous, but for real-time implementation presents a significant challenge.

An efficient, but suboptimal, alternative is to employ the method of matching moments [25, 26]. The technique presented here is based on the 2nd and 4th moments of the data. Assuming a Laplacian model for the image and a Gaussian noise distribution, then the 2nd and 4th moments of the image plus noise are given by

$$\begin{aligned} E\{x^2\} &= m_2 = \sigma_u^2 + \sigma_v^2, \\ E\{x^4\} &= m_4 = 6\sigma_u^4 + 3\sigma_v^4 + 6\sigma_u^2\sigma_v^2. \end{aligned} \quad (3)$$

The moment matching method utilizes estimates of the moments,  $m_2$  and  $m_4$  obtained directly from the data using

$$\widehat{m}_k = \frac{1}{N} \sum_{m,n} x(m,n)^k. \quad (4)$$

Replacing the theoretical moments  $m_k$ , by their estimates  $\widehat{m}_k$ , and solving (3) for the unknown noise variance, one obtains

$$\sigma_v^2 = \widehat{m}_2 \left( 1 - \sqrt{\frac{1}{3} \frac{\widehat{m}_4}{\widehat{m}_2^2} - 1} \right). \quad (5)$$

The above has a pleasant intuitive interpretation. The estimate of the noise variance is obtained by scaling the sample mean square value,  $\widehat{m}_2$ . The factor  $\widehat{m}_4/\widehat{m}_2^2$  represents an estimate of the kurtosis of the noisy image. If the image is dominated by noise, then the kurtosis will be three and  $\widehat{m}_2$  is unscaled. In the presence of a Laplacian component, then the estimated noise variance is reduced by a factor that decreases as the kurtosis increases. There are conditions under which the above expression can yield unrealistic values. If  $\widehat{m}_4 < 3\widehat{m}_2^2$ , the process is sub-Gaussian (platykurtic), and as such cannot be constructed by adding a Gaussian to a Laplacian process. In practice, this will occur in instances where

the process is nearly Gaussian, so that it is reasonable to use  $\sigma_v^2 = \widehat{m}_2$ . Alternatively, if  $\widehat{m}_4 > 6\widehat{m}_2^2$ , then the process has longer tails than those associated with a Laplacian model. Again, summing Laplacian and Gaussian processes cannot form such an image. Under these circumstances then  $\sigma_v^2 = 0$  is appropriate. Such methods are well suited to real-time implementation since they only require two summations across all pixels, conducted when estimating the moments, in contrast to the ML algorithm, which require repeated evaluation of (2).

## 2.2. Estimation of noise variance using trained moments

The method of moment matching relies upon an assumed statistical model for the image and noise. This section describes how this method can be extended to avoid the need to assume a statistical model and instead employs training to form an estimate of the noise variance. The method used to achieve this is based on fitting a linear model based on a normalised set of moments. The algorithm is described in the context of the three moments, but it can be readily generalised to incorporate other moment information. The moments are used in a normalised form and are defined as

$$\begin{aligned} M_1 &= m_1, \\ M_2 &= \frac{m_2}{m_1}, \\ M_4 &= \frac{m_4}{m_1 m_2}. \end{aligned} \quad (6)$$

These are designed to ensure that the normalised moments have the same dimensions as the noise standard deviation. The above choice of normalisations is not unique and similar schemes can be constructed employing different normalised moments. The noise standard deviation is then assumed to be related to these normalised moments through a linear equation

$$\sigma_v = \alpha_1 M_1 + \alpha_2 M_2 + \alpha_4 M_4, \quad (7)$$

where  $\alpha_k$  are constant coefficients. Equation (7) can be regarded as a series approximation to (5), where the assumption of a Laplacian image and additive Gaussian noise has been removed. The normalisation is designed to guarantee the dimensional consistency of (7). The  $K$  images in the training set are then used to evaluate the unknown coefficients  $\alpha_k$ . This is achieved by creating a library of images at different SNRs by adding noise with  $P$  different variances to each of the images. The noise variances are chosen to cover the range of noise levels expected in practice. For each image and noise level, the normalised moments are estimated, which leads to  $K \times P$  realisations of (5). The coefficients  $\alpha_k$  that generate the best approximations, in the least squares sense, to the known noise variance across the training set can be computed using standard linear algebra techniques. These coefficients can then be used to approximate the noise variance on a new noisy image by first computing the normalised moments and then applying (7) with the trained coefficients.

## 2.3. Estimating the noise variance using cumulative distribution functions

This method is based on trying to exploit plane regions in the image. Consider a plane area of the image; the standard deviation of the image computed over that area is a direct estimate of the standard deviation of the noise. It should be noted at this stage that the method is to be applied to wavelet components that, by construction, have a global zero mean. This means that by forming the sum of squared pixel values in a neighbourhood one obtains a localised estimate of the image variance.

In regions where there is image detail (at the scale associated with the particular component) then the local variance will, on average, be the sum of the local image variance and the noise variance, assuming the noise and image are statistically independent. Hence in these regions the local variance will be greater than in plane areas. This implies that information about the noise variance can be obtained by examining areas with the smallest values of the local variance. To form an estimate based on this information, the cumulative distribution function (cdf),  $c(x)$ , of the local pixel variances is formed. The value of  $c(x)$  represents the number of pixels with a local variance less than  $x$ .

The character of the cdf depends upon both the image and noise statistics, but for small values of  $x$  the values of  $c(x)$  are dominated by the noise. Computation of cdf for all possible values of  $x$  is burdensome and the solution adopted herein is designed with computational efficiency in mind. Specifically, we will only measure the cdf for a particular value of  $x = x_0$ . Mean values of  $c(x_0)$  are computed across the training set of images and are stored for a range of noise variances. This forms a lookup table of values of  $c(x_0)$  against noise variance. When a new image is presented, the value of  $c(x_0)$  is computed and the lookup table is employed to infer the noise variance.

The effectiveness of the method depends upon the choice of  $x_0$ . This point is chosen as the value that maximises a discrimination metric evaluated across the training set. As an example, the optimal grey level discriminator,  $x_{0_{l_1 l_2}}$ , between two noise levels can be defined using the function

$$\begin{aligned} f_{l_1, l_2}(x) &= \frac{|m_{l_1}(x) - m_{l_2}(x)|}{\sqrt{\sigma_{l_1}(x) + \sigma_{l_2}(x)}}, \\ x_{0_{l_1 l_2}} &= \arg \max_x (f_{l_1, l_2}(x)), \end{aligned} \quad (8)$$

where  $l_1$  and  $l_2$  are the two noise levels,  $m_l(x)$  and  $\sigma_l(x)$  are the mean and the standard deviation of the cdf at a grey level  $x$  computed across the set of images. The optimal grey level between all the noise levels  $x_0$  is defined as

$$\begin{aligned} x_0 &= \arg \max_x (f_{\text{tot}}(x)), \\ f_{\text{tot}}(x) &= \sum_{i,j} f_{l_i, l_j}(x), \quad i < j, \end{aligned} \quad (9)$$

where the summation is taken across all the noise levels considered.



FIGURE 1: Images used for training.

### 3. RESULTS

To assess the performance of the noise estimation processes, a series of simulations was conducted. The three methods for noise estimation presented in Section 2 were implemented along with MAD. Those methods that needed training were trained on a set of 13 images (Figure 1). The performance of the methods was then evaluated using a selec-

tion of five images (Figure 2).<sup>1</sup> Note that the training and test sets contained no common images. Gaussian noise was added to each of the five images using six different noise levels. The noise was estimated using only the highest frequency

<sup>1</sup>Training and test sets of images are available at <http://www.soton.ac.uk/~anto/image/figures and tables.htm>.

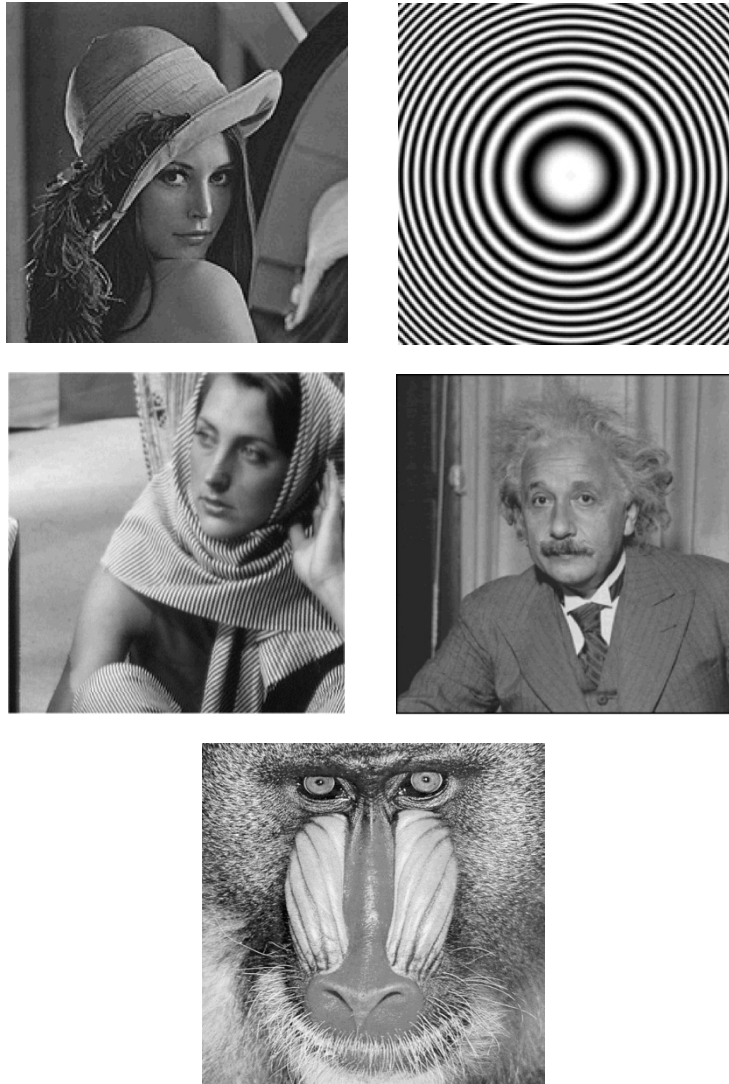


FIGURE 2: Images used for testing.

(smallest scale) wavelet component. The filter bank used for the wavelet decomposition is shown in Figure 3 and the coefficients described in (10),<sup>2</sup>

$$\begin{aligned}
 H_{11h} &= \frac{[1, 2, 1]}{4}, & H_{01h} &= \frac{[-1, 2, -1]}{4}, \\
 H_{11v} &= H_{11h}^T, & H_{01v} &= H_{01h}^T, \\
 H_{12h} &= \frac{[1, 0, 2, 0, 1]}{4}, & H_{02h} &= \frac{[-1, 0, 2, 0, -1]}{4}, \\
 H_{12v} &= h_{12h}^T, & h_{02v} &= h_{02h}^T, \\
 H_{13h} &= \frac{[1, 0, 0, 0, 2, 0, 0, 0, 1]}{4}, \\
 H_{03h} &= \frac{[-1, 0, 0, 0, 2, 0, 0, 0, -1]}{4}, \\
 H_{13v} &= H_{13h}^T, & H_{03v} &= H_{03h}^T.
 \end{aligned} \tag{10}$$

<sup>2</sup>The filter bank used for the wavelet decomposition is extensively described in [3].

The mean squared error between the estimated noise variance and the true variance of the added noise is computed; the results are presented in Tables 1 and 2.

Table 1 illustrates the improved performance of all three new methods relative to MAD with respect to the six noise levels. The mean squared error is computed over the five images. In general, the two training-based methods achieve a better performance than the moment matching technique or MAD. For four noise levels the cdf method performs best and for the other two levels the trained moment-based method achieves the best results. The grey level selected to estimate the noise level in the cdf method performs particularly well when the contaminating noise level is clearly higher than the standard deviation of the image component (last column in Table 2). All the methods tend to perform better as the noise level increases, as one would anticipate. This test provides some evidence of the utility of training-based schemes.

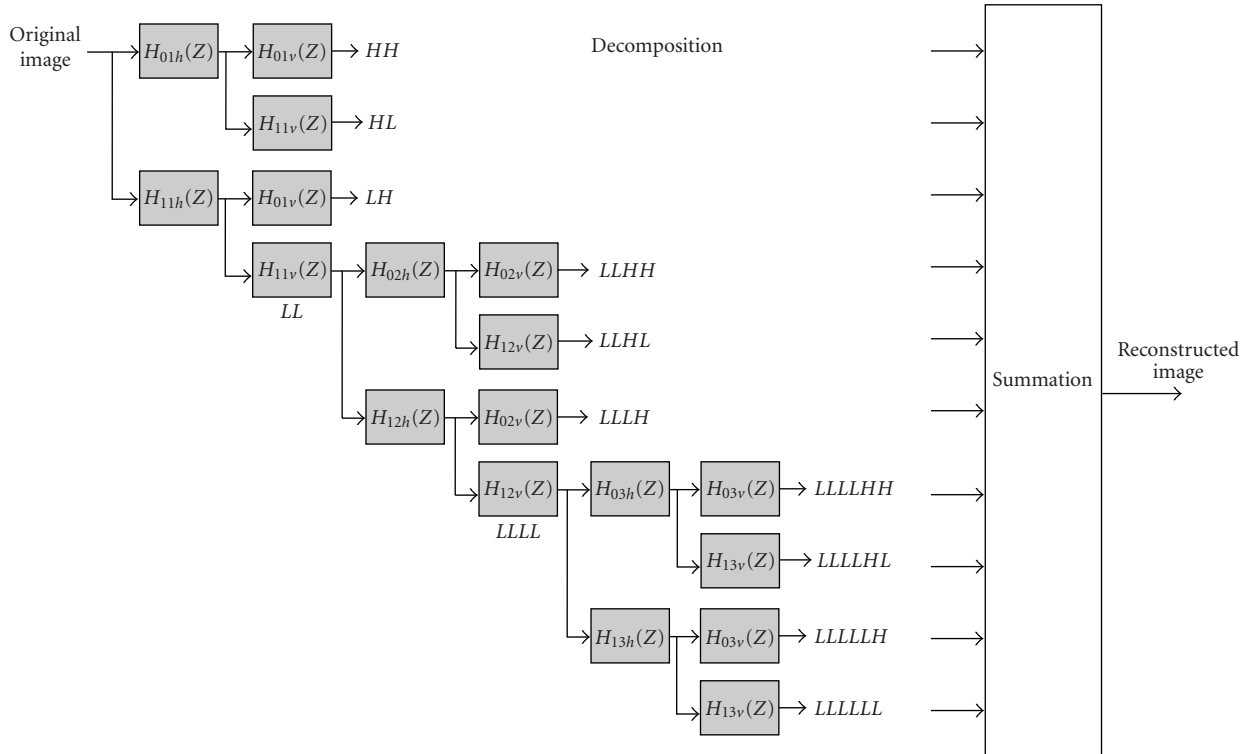


FIGURE 3: Filter bank used for the wavelet decomposition.  $H_{0mh}$  and  $H_{1mh}$  are the horizontal decomposition filters for decomposition level  $h$  and  $H_{0mv}$  and  $H_{1mv}$  are the vertical decomposition filters for decomposition level  $h$ . The reconstruction is achieved by summing of the components.

Table 2 compares the performance of the three new methods with those of the MAD method and with respect to the five images. The mean squared error is computed over the six noise levels. The cdf method achieves the best performances for three images (A, B, and C), the performances of the MAD and moment matching methods are superior, respectively, for images D and E. In general, again the two training-based methods achieve better performance than the moment matching technique or MAD.

We believe that the poor performance of the moment matching method (third column) can be attributed to the inadequacy of the Laplacian distribution for modelling the underlying image. This has been verified by comparing the mean squared error between synthetically generated images with optimal<sup>3</sup> Laplacian distribution and the image component distribution (sixth column). The ratio between the values in the third and sixth columns and in the same row is almost constant and this demonstrates that performances of the moment matching method are strongly related to the discrepancy between optimal (Laplacian distribution) and real image components.

<sup>3</sup>The parameters of the Laplacian distribution were selected to minimise the MSE between synthetic image and image component.

We also believe that the comparatively poor performance of MAD (second column) is due to the fact that it assumes that there is zero image contribution in the component being examined. The last column of Table 2 lists the standard deviation of the image component. Comparing the second and last columns, the relation between the MAD performance and the standard deviation of the image component contribution is clear.

#### 4. CONCLUSIONS

The problem of the noise standard deviation level estimation over the wavelet component is considered in this work. Three novel methods have been proposed and their performance was compared with that of those achieved using the classical MAD-based method. The techniques utilised to estimate the noise level are in general based on some type of assumption concerning image and noise characteristics. An alternative solution proposed here is to use training-based methods which do not rely on any prior assumption and utilise parameters extracted from a preliminary stage performed on a set of representative images. Among the methods proposed here, two are training-based, while the third is based on the assumption of specific statistical distributions for image and noise components. The set of images used for training is representative

TABLE 1: Mean squared errors for noise standard deviation estimates computed over 5 images.

Standard deviation of synthetic noise	MAD	Moment matching	Trained moments	cdf
5	3.49	2.49	2.35	2.73
7	3.01	2.36	1.91	1.86
9	2.63	2.13	1.59	1.55
11	2.33	1.97	1.38	1.40
13	2.08	1.76	1.37	0.46
15	1.84	1.62	1.31	0.43
Overall mean	2.56	2.06	1.65	1.41

TABLE 2: Mean squared errors for noise standard deviation estimates computed over 6 noise levels (columns 2–5); mean squared error between image component and synthetically generated image with optimal Laplacian distribution (column 6); and standard deviation of the image component (column 7).

Image	MAD	Moment matching	Trained moments	cdf	MSE ( $I_C - I_L$ )	STD $I_C$
Image A	2.75	2.89	0.84	0.34	1.18	8.14
Image B	1.27	1.32	1.25	1.00	0.62	2.57
Image C	1.95	3.25	0.98	0.71	1.45	3.49
Image D	0.83	1.22	1.52	1.73	0.47	1.67
Image E	5.98	1.62	3.68	3.26	0.83	11.11
Overall mean	2.56	2.06	1.65	1.41	—	—

of the class of the video images and completely disjoint from the set used for testing the methods and comparing the results. For the large majority of the images and noise levels considered, the training-based methods demonstrated their ability to offer superior performance. The advantages and disadvantages of the model-based techniques, such as the MAD and the novel model proposed here, are also discussed.

The results showed in this paper need to be generalised using larger sets of test images and different range noise levels. The techniques proposed seem also to be suitable for other classes of images and for non-spatially white Gaussian noise distributions. A desirable development of this work could focus on these aspects.

## ACKNOWLEDGMENT

We would like to gratefully acknowledge the financial support of Snell and Wilcox Ltd. in conducting this work and would particularly like to thank Martin Weston for his many comments and suggestions.

## REFERENCES

- [1] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.
- [2] A. Bruce, D. Donoho, and H. Gao, "Wavelet analysis [for signal processing]," *IEEE Spectrum*, vol. 33, no. 10, pp. 26–35, 1996.
- [3] A. De Stefano, *Wavelet based reduction of spatial video noise*, Ph.D. thesis, ISVR, University of Southampton, Southampton, UK, 2000.
- [4] M. R. Banham and A. K. Katsaggelos, "Spatially adaptive wavelet-based multiscale image restoration," *IEEE Trans. Image Processing*, vol. 5, no. 4, pp. 619–634, 1996.
- [5] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells Jr., "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.
- [6] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [7] D. L. Donoho and I. M. Johnstone, "Wavelet shrinkage: asymptopia?," *Journal Royal Statistics Society B*, vol. 57, no. 2, pp. 301–369, 1995.
- [8] E. P. Simoncelli, "Bayesian denoising of visual images in the wavelet domain," in *Bayesian Inference in Wavelet-Based Models*, P. Muller and B. Vidakovic, Eds., vol. 141 of *Lecture Notes in Statistics*, pp. 291–308, Springer-Verlag, New York, NY, USA, 1999.
- [9] H. Y. Gao and A. G. Bruce, "WaveShrink with firm shrinkage," *Statistica Sinica*, vol. 7, no. 4, pp. 855–874, 1997.
- [10] R. R. Coifman and D. L. Donoho, "Translation-invariant denoising," in *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, Eds., vol. 103 of *Lecture Notes in Statistics*, pp. 125–150, Springer-Verlag, New York, NY, USA, 1995.
- [11] A. De Stefano, P. R. White, and W. B. Collis, "An innovative approach for spatial video noise reduction using a wavelet based frequency decomposition," in *Proc. IEEE International Conference on Image Processing (ICIP '00)*, vol. 3, pp. 281–284, Vancouver, British Columbia, Canada, September 2000.
- [12] A. De Stefano, P. R. White, and W. B. Collis, "Selection of thresholding scheme for video noise reduction on wavelet

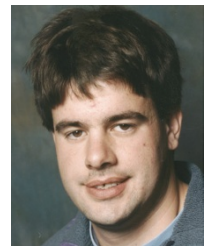
components using Bayesian estimation,” in *Proc. 5th IMA International Conference on Mathematics in Signal Processing*, Warwick, UK, December 2000.

- [13] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley Cambridge Press, Wellesley, Mass, USA, 1996.
- [14] M. Jansen and A. Bultheel, “Multiple wavelet threshold estimation by generalized cross validation for images with correlated noise,” *IEEE Trans. Image Processing*, vol. 8, no. 7, pp. 947–953, 1999.
- [15] Y. Xu, J. B. Weaver, D. M. Healy Jr., and J. Lu, “Wavelet transform domain filters: a spatially selective noise filtration technique,” *IEEE Trans. Image Processing*, vol. 3, no. 6, pp. 747–758, 1994.
- [16] S. Mallat and W. L. Hwang, “Singularity detection and processing with wavelets,” *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 617–643, 1992.
- [17] M. Malfait and D. Roose, “Wavelet-based image denoising using a Markov random field a priori model,” *IEEE Trans. Image Processing*, vol. 6, no. 4, pp. 549–565, 1997.
- [18] M. R. Banham and A. K. Katsaggelos, “Spatially adaptive wavelet-based multiscale image restoration,” *IEEE Trans. Image Processing*, vol. 5, no. 4, pp. 619–634, 1996.
- [19] P. Moulin and J. Liu, “Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors,” *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 909–919, 1999.
- [20] D. Wei and C. S. Burrus, “Optimal wavelet thresholding for various coding schemes,” in *Proc. IEEE International Conference on Image Processing (ICIP '95)*, vol. 1, pp. 610–613, Washington, DC, USA, October 1995.
- [21] T. R. Downie and B. W. Silverman, “The discrete multiple wavelet transform and thresholding methods,” *IEEE Trans. Signal Processing*, vol. 46, no. 9, pp. 2558–2561, 1998.
- [22] S. Mallat, *Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif, USA, 1998.
- [23] C. Taswell, “Experiments in wavelet shrinkage denoising,” *Journal of Computational Methods in Sciences and Engineering*, vol. 1, no. 2s-3s, pp. 315–326, 2001.
- [24] C. Taswell, “The what, how, and why of wavelet shrinkage denoising,” *IEEE Computing in Science and Engineering*, vol. 2, no. 3, pp. 12–19, 2000, invited paper.
- [25] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, USA, 1991.
- [26] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, John Wiley & Son, New York, NY, USA, 2001.

**P. R. White** is currently a Senior Lecturer in the Institute of Sound and Vibration Research, the University of Southampton. He attained his degree in applied mathematics from Portsmouth Polytechnic in 1985, whereupon he joined the ISVR to study for his Ph.D. In 1988, he was made a Lecturer in ISVR, finally completing his Ph.D. in 1992. In 1998, he was made a Senior Lecturer. The basic signal processing techniques that have formed the basis of this work include time-frequency analysis, non-linear systems, adaptive systems, detection and classification algorithms, higher-order statistics, and independent component analysis. The application areas which he has considered include image processing, underwater systems, condition monitoring, and biomedical applications. He has published more than 130 papers in the field, approximately 35 of which appearing in referred journals or as chapters in books. He is a Member of the Editorial Board of the *Journal of Condition Monitoring and Diagnostic Engineering Management (COMADEM)* and is a Member of the IEEE.



**W. B. Collis** has been Algorithms Engineer at the Foundry since December 2000. Prior to joining the Foundry, Collis led the algorithms team at Snell & Wilcox for 5 years working on standards conversion, motion estimation, archive restoration, and filter design. During this period he also developed the Flo-Mo retiming software used in the pioneering “bullet time” sequences in the film *The Matrix*. Collis graduated with high honours in electrical engineering from Southampton University, and went on to complete a Ph.D. in nonlinear signal processing at the Institute of Sound and Vibration Research, where he still holds the position of Associate Fellow. Collis is the author of a book, over 30 research papers, and five patents.



**A. De Stefano** is Distance Learning Coordinator for the Master Training Packages at the Centre of Biomedical Signal Processing, Institute of Sound and Vibration Research (ISVR), University of Southampton, UK. He received the Electronics Engineering degree in biomedical sciences, the CEng qualification from the University of Federico II in Naples, Italy, and the Ph.D. degree from the ISVR at the University of Southampton, UK. His research interests include wavelet-based noise reduction and image enhancement, biomedical signal processing, implementation of distance learning packages for biomedical subjects, techniques for EMG analysis during walking of pathological children, and mechanical models for speech design. He has over 20 publications in journals and international conferences.

