Dual-tree Complex Wavelet Transform based denoising for Random Spray image enahcement methods

Massimo Fierro, Wang-Jun Kyung, Yeong-Ho Ha; School of Electrical Engineering and Computer Science, Kyungpook National University; Taegu, Republic of Korea; fierro, kyungwj, yha @ee.knu.ac.kr

Abstract

This work introduces a novel way to reduce point-wise noise introduced or exacerbated by image enhancement methods leveraging the Random Spray sampling approach. Due to the nature of the spray, the sampling structure used, output images for such methods tend to exhibit noise with unknown distribution. The proposed noise reduction method is based on the assumption that the non-enhanced image is either free of noise or contaminated by non-perceivable levels of noise. The dual-tree complex wavelet transform is applied to the luma channel of both the non-enhanced and enhanced image. The standard deviation of the energy for the non-ehanced image across the six orientations is computed and normalized. The normalized map obtained is used to shrink the real coefficients of the enhanced image decomposition. A noise reduced version of the enhanced version can then be computed via the inverse transform. A thorough numerical analysis of the results has been performed in order to confirm the validity of the proposed approach.

Introduction

The field of image enhancement has been one of the most active even before digital imagery achieved a consumer status, but despite its age it has never stopped evolving. The present work introduces a novel denoising method, tailored to address a specific image quality problem expressed by Random Spray based image enhancement algorithms.

Random sprays are two-dimensional collection of points (coordinates) with a given spatial distribution around the origin. Such structures can be scaled and translated to sample an image's support in a way similar to that employed by the Human Visual System (HVS). Yet, the peaked nature of the sprays introduces unwanted noise in the output image. The amount and statistical characteristics of the noise so introduced depend on several factors, among which are image content and spray properties. Thus, common noise reduction methods tailored to deal only with one particular kind of noise (e.g. Gaussian noise) would not find the expected conditions.

The method here described approaches the problem via wavelet coefficient shrinkage. Algorithms based on wavelet shrinkage have a long history, nonetheless this work presents a novel view on the subject. This article was particularly inspired by the works on the Dual-tree Complex Wavelet Transform by Kingsbury [6], the work on the Steerable Pyramid Transform by Simoncelli et al. [15], and the work on Wavelet Coefficient Shrinkage by Donoho and Johnstone [3].

Dual-Tree Complex Wavelet Transform

Kingsbury developed the Complex Wavelet Transform (CWT) in order to solve certain problems that arise with the traditional Discrete Wavelet Transform (DWT), as well as other more advanced methods such as the Steerable Pyramid Transform (SPT) [5]. Similarly to the SPT, in order to retain the whole

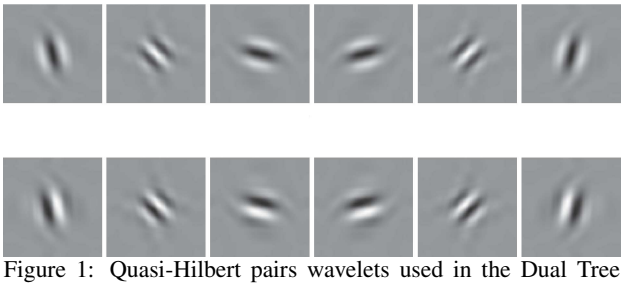


Figure 1: Quasi-Hilbert pairs wavelets used in the Dual Tree Complex Wavelet Transform

Fourier spectrum, the CWT needs to be overcomplete by a factor 4, i.e. there are 3 complex coefficients for each real one. The CWT is also efficient, as it can be computed through separable filters, yet it lacks the Perfect Reconstruction property.

Kingsbury also introduced the concept of Dual-tree Complex Wavelet Transform (DTCWT), which has the added characteristic of Perfect Reconstruction at the cost of only approximate shift-invariance [6].

Since the a full discussion on the Dual-Tree Complex Wavelet Transform would be too cumbersome, only a brief introduction the 2D variant of the DTCWT is given. The reader is referred to the the work by Selesnick et al. [14] for a very comprehensive coverage on the DTCWT and the relationship it shares with other transforms.

The 2D Dual Tree Complex Wavelet Transform can be implemented by using two distinct sets of separable 2D wavelet bases, as shown below.

$$\begin{aligned} \psi_{1,1}(x,y) &= \phi_h(x)\psi_h(y), & \psi_{2,1}(x,y) &= \phi_g(x)\psi_g(y), \\ \psi_{1,2}(x,y) &= \psi_h(x)\phi_h(y), & \psi_{2,2}(x,y) &= \psi_g(x)\phi_g(y), \\ \psi_{1,3}(x,y) &= \psi_h(x)\psi_h(y) & \psi_{2,3}(x,y) &= \psi_g(x)\psi_g(y) \end{aligned} \quad (1)$$

$$\begin{aligned} \psi_{3,1}(x,y) &= \phi_g(x)\psi_h(y), & \psi_{4,1}(x,y) &= \phi_h(x)\psi_g(y), \\ \psi_{3,2}(x,y) &= \psi_g(x)\phi_h(y), & \psi_{4,2}(x,y) &= \psi_h(x)\phi_g(y), \\ \psi_{3,3}(x,y) &= \psi_g(x)\psi_h(y) & \psi_{4,3}(x,y) &= \psi_h(x)\psi_g(y) \end{aligned} \quad (2)$$

The following equations shows the relationship between wavelet filters h and g

$$g_0(n) \approx h_0(n-1), \text{ for } j = 1 \quad (3)$$

$$g_0(n) \approx h_0(n-0.5), \text{ for } j > 1 \quad (4)$$

where j is the decomposition level.

When combined, the bases give rise to two sets of real, two dimensional, oriented wavelets (see Fig. 1).

$$\psi_i(x,y) = \frac{1}{\sqrt{2}}\left(\psi_{1,i}(x,y) - \psi_{2,i}(x,y)\right) \quad (5)$$

$$\psi_i + 3(x,y) = \frac{1}{\sqrt{2}}\left(\psi_{1,i}(x,y) + \psi_{2,i}(x,y)\right) \quad (6)$$

$$\psi_i(x,y) = \frac{1}{\sqrt{2}}\left(\psi_{3,i}(x,y) + \psi_{4,i}(x,y)\right) \tag{7}$$

$$\psi_i + 3(x,y) = \frac{1}{\sqrt{2}}\left(\psi_{3,i}(x,y) - \psi_{4,i}(x,y)\right) \tag{8}$$

The most interesting characteristic of such wavelets is that they are approximately Hilbert pairs. One can thus interpret the coefficients deriving from one tree as imaginary, and obtain the desired 2D DTCWT.

Random Spray Sampling

Random Spray sampling was introduced for the first time in [11], by Provenzi et al. Random sprays are an elaboration over the physical spatial scanning structure used by Land in the seminal work on Retinex [7]. In the experiments ran at the time, Land used a structure resembling a set of paths departing from a central point, on which he mounted a number of photo-detectors.

Land's model gave rise to the path-wise family of Retinex implementations [8, 13], which directly transposed Land's machinery into piece-wise linear paths used to scan the input image, and that, in turn, became the starting point of RSR.

A single point of a random spray may be generated using the following formulation, and the whole spray is obtained by reiterating the process

$$\mathbf{p} = [\rho\cos(\theta), \rho\sin(\theta)] \tag{9}$$

where $\rho = \text{rand}(0,R)$ and $\theta = \text{rand}(0,2\pi)$ and the function rand indicates the uniform random distribution. In particular, R is going to be set as the diagonal of the image, so that the spray may cover the whole of it.

Each spray can then be used to sample the image by transforming its points as follows

$$\tilde{\mathbf{p}} = \mathbf{p} + \mathbf{i} \tag{10}$$

where $\mathbf{i} = \left[i_x, i_y\right]$ are the coordinates of the pixel used as reference for sampling.

The problem of noise

The sharp selection imposed by the spray sampling leads, in certain images, to the unwanted effect of increasing the relevance of noise already present in the image. This is especially true for algorithms that employ differential operators in their computations, such as RSR and RACE.

The problem of noise has been partially addressed in [12] through the use of a form of attachment to the original data, strongly reducing the insurgence of unwanted speckles in uniform areas. Such process can be summarized in the following equation, computed for each chromatic channel

$$\tilde{O}(x,y) = \beta(x,y)\,O(x,y) + (1-\beta(x,y))\,I(x,y) \tag{11}$$

where I indicates the input image, O the output image of non-regularized RSR or RACE, and β is computed per-pixel. Given a specific spray k, the local parameter is automatically determined for each chromatic channel

$$\beta_k(x,y) = (2\sigma_k(x,y))^{\frac{\sigma_{\min}}{\sigma_{\max}}} \tag{12}$$

where the quantities $\sigma_{\min}$ and $\sigma_{\max}$ are the image-wise standard deviations, while $\sigma_k(x,y)$ is the (local) standard deviation for the spray.

Proposed method

The main idea behind this work can be subsumed in a simple sentence: highly directional content is what conveys the largest part of information to the HVS. This statement is backed up by past research, such as the Retinex theory as well as the high-order gray-world assumption (alias gray-edges) [16]. In particular, the local white patch effect described by Retinex comes into play when, for a given channel, the scanning structure samples a positive intensity change. For obvious geometrical reasons intensity changes of directional nature are are more easily crossed (or sampled) than point like structures such as noise.

Following such idea, the proposed method revolves around the shrinkage of the real wavelet coefficients generated by the Dual Tree Complex Wavelet Transform, according to data directionality.

Furthermore, since the human visual system is highly sensitive to changes in luminance [17], the presented approach first converts the image in a space where chroma is separated from luma (such as YCbCr), and it operates on the wavelet space of the luma channel. While this may seem counter intuitive, since spray-based image enhancement algorithms usually operate per channel, the results show vast improvements without visiblecolor artifacts.

A final, fundamental assumption is made: the input image is considered to be either free of noise, or noise is sopposed to be present but not perceivable. If such assumption holds, the input image contains the information needed to successfully perform noise reduction.

The algorithm for the proposed method is given in Alg. 1.

Algorithm 1 The algorithm for the proposed noise-reduction method.

$E_{RGB} \leftarrow \text{enhance}(I_{RGB})$
$I_{YCbCr} \leftarrow \text{rgb2ycbcr}(I_{RGB})$
$E_{YCbCr} \leftarrow \text{rgb2ycbcr}(E_{RGB})$
$b^I \leftarrow \text{dtcwt}(Y_I)$
repeat
  $b^E \leftarrow \text{dtcwt}(Y_E)$
  for $j = 1 \rightarrow J$ **do**
    for $k = 1 \rightarrow 6$ **do**
      $e_{j,k} \leftarrow \sum_k b^I_{j,k}$
    end for
    $w_j \leftarrow \text{mm}(\underset{k}{\text{stddev}}(e_{j,k}), \underset{k}{\text{median}}(e_{j,k}), \gamma_j)$
    for $k = 1 \rightarrow 6$ **do**
      $\tilde{b}^E_{j,k} \leftarrow w_j \cdot b^E_{j,k} + (1-w_j)\cdot b^I_{j,k}$
      if $\text{ord}(b^E_{j,k}) \in \{1,2\}$ **then**
        $b^O_{j,k} \leftarrow \tilde{b}^E_{j,k}$
      else
        $b^O_{j,k} \leftarrow b^I_{j,k}$
      end if
    end for
  end for
  $Y_E \leftarrow Y_O$
until $\text{ssim}(Y_I, Y_O) < 0.001$
$O_{YCbCr} = \text{concat}(Y_O, E_{CbCr})$
$O_{RGB} = \text{ycbcr2rgb}(O_{YCbCr})$

Wavelet coefficients shrinkage

Assuming level j of the wavelet pyramid, one can compute the energy for each direction $k \in \{1,2,...,6\}$ as

$$e_{j,k} = (b^I_{j,k})^2 + (i^I_{j,k})^2$$

Coefficients associated with non directional data will have similar energy for all directions. On the other hand, highly directional data will give rise to high energy for one or two directions (this is not entirely true, as more than two directions may have high energy for "L" ot "T" shaped features, but it does not compromise the efficacy of the method to large extents).

The standard deviation of energy across the six directions $k = 1,2,...,6$ is hence computed as a measure of directionality.

$$e_j = \operatorname*{stddev}_k(e_{j,k})$$

Standard deviation is, obviously, non-normalized and the range may vary widely. In order to make use of the computed data, we resort to the Michaelis-Menten function [9] for normalization of the data range. The Michaelis-Menten function is a sigmoid-like function that has been used to model many species' cones response, and thus looks to be a perfect candidate for the job. The equation is as follows:

$$\mathrm{mm}(x,\mu,\gamma) = \frac{x^\gamma}{x^\gamma + \mu^\gamma} \tag{13}$$

where x is the quantity to be compressed, γ a real-valued exponent and μ the data expected value (or its estimate).

A normalized map of directionally sensitive weights is obtained as

$$w_j = \mathrm{mm}(e_j, \operatorname*{median}_k(e_{j,k}), \gamma_j)$$

where the choice of γ depends on j as explained later on.

A shrinked version of the enhanced image's real coefficients, according to data directionality, is then computed as

$$\tilde{b}^E_{j,k} = w_j \cdot b^E_{j,k} + (1 - w_j) \cdot b^I_{j,k} \tag{14}$$

Since we are interested in retaining directional information, the output coefficients are computed according to Eq. 15

$$b^O_{j,k} = \begin{cases} \tilde{b}^E_{j,k}, \text{ if } \mathrm{ord}(b^E_{j,k}) \in \{1,2\} \\ b^I_{j,k}, \text{ if } \mathrm{ord}(b^E_{j,k}) \in \{3,4,5,6\} \end{cases} \tag{15}$$

where ord is a function that returns the index of a coefficient in $b^I_{k=1,2,...,6}$ when the set is sorted in descending order.

The meaning of the whole sequence can be roughly expressed as follows: where the non-enhanced image shows directional content replace its two most significant coefficients with a (possibly) shrinked version of those from the enhanced image.

Parameter tuning

When dealing with functions having free parameters, a fundamental problem is that of finding their optimum values. This can often be attempted with optimization techniques, but for this particular case it is unfeasible.

In order to at least provide a reasonable default value for γ_j, the parameter of the Michaelis-Menten function, as well as the depth of the complex wavelet decomposition J, three images from the USC-SIPI Image Database [1] were chosen. Such images provide a good mixture of mostly high-frequency detail (Mandrill), balanced high- and low-frequency content (Lenna), and mainly low frequency content (Splash). The chosen test images are shown in Fig. 2.

In different rounds, Gaussian, Poissonian and Speckle noise was added to the luma channel of said images and the proposed noise reduction method was run with 3 wavelet levels and values for γ_j varying from -5 to 10 in unary steps for the first and second levels of the decomposition. The value of γ_3 was fixed to 1, reducing the Michaelis-Menten function to the Naka-Rushton formulation [10].

Table 1: PSN ratios and SSIM scores for from the USC-SIPI database

		Noisy		Denoised	
Noise	Image	PSNR	SSIM	PSNR	SSIM
Gaussian	Lenna	27.43	0.58	35.78	0.95
	Splash	27.68	0.49	36.37	0.93
	Mandrill	27.45	0.73	34.78	0.98
Poissonian	Lenna	30.33	0.84	35.78	0.95
	Splash	30.92	0.82	36.37	0.93
	Mandrill	30.16	0.91	34.78	0.98
Speckle	Lenna	27.07	0.55	33.96	0.94
	Splash	27.55	0.53	33.84	0.94
	Mandrill	26.96	0.69	33.87	0.97

Since the decomposition depth is more dependent on the image than on γ_j, it was impossible to determine a single optimum value. Nonetheless, reasonable bounds for "normal" amounts of noise were found to be $J_{min} = 1, J_{max} = 3$. On the other hadn J was set to 4 when testing versus the Foveated NL-means approach. Please note that J is the only parameter that should be set explicitly, since the variation in image quality can be extremely evident.

Performance was tested employing the SSIM [18] measure, holding the unaltered luma channel as absolute reference. Iterations were stopped using a threshold of $t = 0.001$. Both SSIM scores as well as PSNRs are given in Table 1.

For all images the score improves for positive values of γ at all levels, yet, the change is less sudden for the first level. While non-optimal for all possible images, values of 3 and 1 for γ at the first and second level, respectively, represent a reasonable choice.

Experiments

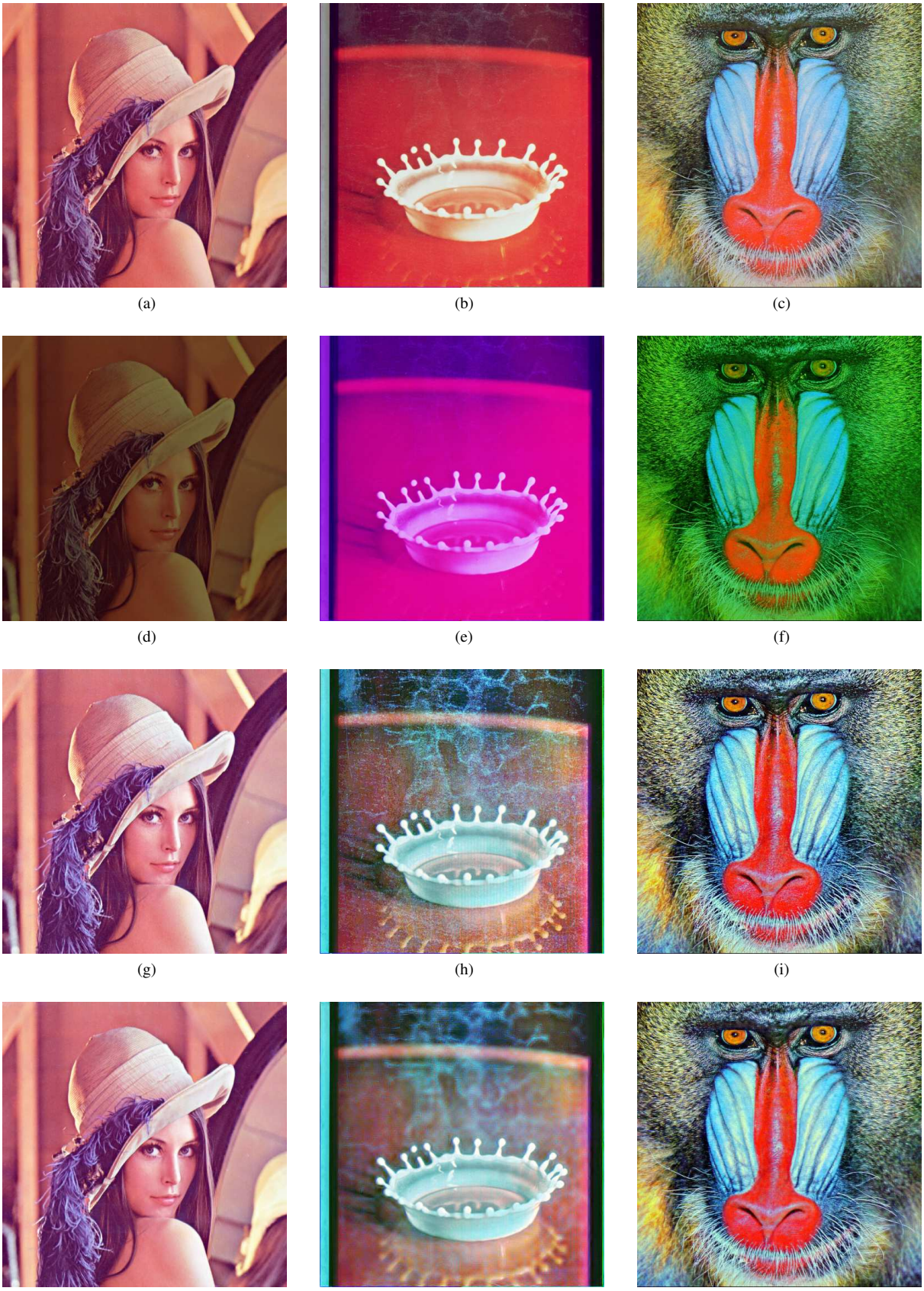
In order to test our approach, we modified the images shown in Figs. 2a - 2c so to reduce the available dynamic range and introduce a strong color cast. The resulting images have been used as test subjects, as illustrated in Fig. 2.

The proposed approach was also compared to a recent development on the Non Local means approach by Buades et al. [2], namely the Foveated NL-means by Foi and Boracchi [4]. Three images were chosen from said work, and the proposed noise reduction method was used after Gaussian noise of various standard deviation was added. Both the PSNR and SSIM scores were computed and they are reported in Tables 2 and 3: the numbers differ slightly from the original article since the images were generated anew. While the comparison is not entirely fair, as NL-means is a reference-less denoising method, it clearly shows the advantage of assuming a partial reference.

Conclusions

This work presented a novel approach to noise reduction which aims at image enhancement methods that make use of the Random Spray sampling technique.

In order to achieve noise reduction, the proposed method leverages the data orientation discriminating power of the Dual Tree Complex Wavelet Transform, as well as the information contained in the non-enhanced image. Wavelet coefficient shrinkage and selection are the basic mechanisms underlying the iterative processing. Unlike most of the state of the art, this ap-



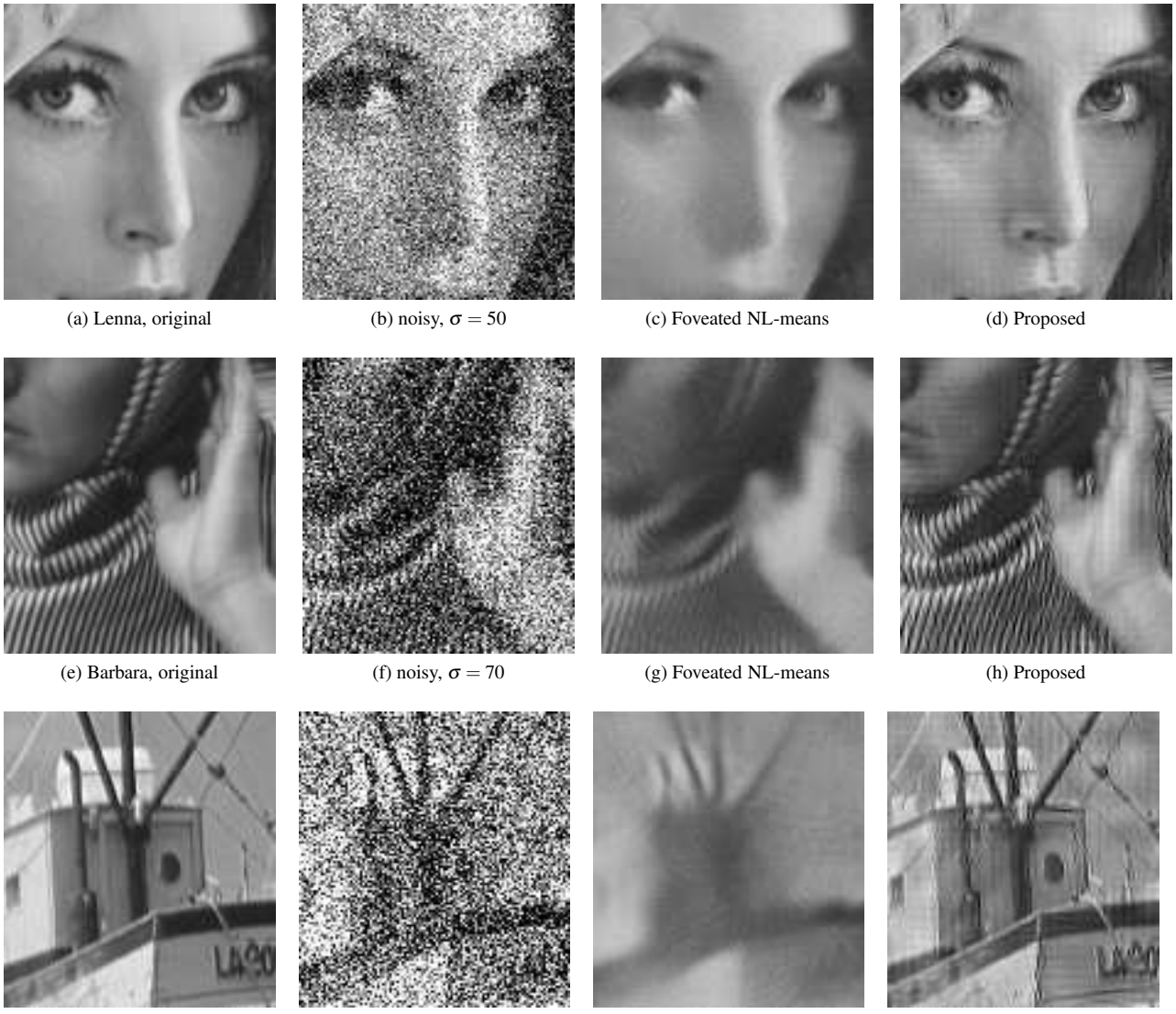
(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Figure 2: Test images from the UCS-SIPI image database, results. *Lenna* was enhanced with RSR, while the others with the spray formulation of ACE. The noise reduction algorithm was run with $J = 3$ for images *Lenna* and *Splash*, and $J = 2$ for *Mandrill*.



(a) Lenna, original (b) noisy, $\sigma = 50$ (c) Foveated NL-means (d) Proposed

(e) Barbara, original (f) noisy, $\sigma = 70$ (g) Foveated NL-means (h) Proposed

(i) Boats, original (j) noisy, $\sigma = 100$ (k) Foveated NL-means (l) Proposed

Figure 3: Comparison between the proposed method and Foveated NL-means. Noise is Gaussian with standard deviation as per captions.

Table 2: Comparison of PSNRs for three images taken from the work on Foveated NL-means.

	Barbara		Boats		Lenna	
σ	FNLM	Our	FNLM	Our	FNLM	Our
10	33.39	41.71	32.73	42.76	35.05	43.24
20	30.46	37.86	29.89	38.94	32.39	39.43
30	28.08	35.83	28.05	36.93	30.51	37.42
40	26.27	34.52	26.59	35.46	28.91	36.19
50	24.77	33.61	25.24	34.37	27.63	35.11
60	23.61	32.60	24.14	33.45	26.43	33.83
70	22.68	32.04	23.29	32.52	25.37	32.79
80	21.80	31.17	22.42	31.57	24.37	31.81
90	21.02	30.60	21.75	30.88	23.32	31.19
100	20.33	30.33	21.06	30.20	22.41	30.52

Table 3: Comparison of SSIM scores for three images taken from the work on Foveated NL-means.

	Barbara		Boats		Lenna	
σ	FNLM	Our	FNLM	Our	FNLM	Our
10	0.97	0.99	0.96	0.99	0.96	0.99
20	0.94	0.98	0.90	0.98	0.93	0.98
30	0.89	0.97	0.84	0.97	0.90	0.96
40	0.85	0.96	0.79	0.95	0.86	0.95
50	0.80	0.94	0.74	0.94	0.83	0.93
60	0.75	0.93	0.69	0.92	0.79	0.91
70	0.71	0.91	0.65	0.91	0.77	0.90
80	0.67	0.90	0.61	0.89	0.74	0.88
90	0.63	0.88	0.58	0.88	0.71	0.87
100	0.60	0.87	0.55	0.86	0.68	0.85

proach requires no prior knowledge of the statistical properties of noise. The only parameter that the user is required to choose explicitly is the depth of the DTCWT.

Prformance has been tested in two ways. First, noise with different statistical properties has been added to images with a well known reference. The proposed approach was able to achieve great improvements in both PSNRs and SSIM scores independently of the noise distribution.

The propsed approach has then been compared to a recent development of the NL-means denoising algorithm, using images with a well known prior contaminated by Gaussian noise of varying standard deviation. Our method shows consistent increases in PSNR of about 9 dBs on average, as well as higher SSIM scores, never dropping below 0.85.

The proposed noise reduction approach shows great ability in removing noise without altering the underlying structures, although its performance is naturally limited by the contrast of the non-enhanced image.

References

[1] USC-SIPI image database.

[2] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition CVPR 2005*, volume 2, pages 60–65, 2005.

[3] D. L. Donoho and I. M. Johnstone. Threshold selection for wavelet shrinkage of noisy data. In *Proc. 16th Annual Int. Conf. of the IEEE Engineering Advances: New Opportunities for Biomedical Engineers Engineering in Medicine and Biology Society*, 1994.

[4] A. Foi and G. Boracchi. Foveated self-similarity in nonlocal image filtering. In *Proc. IS&T / SPIE Electronic Imaging 2012*, 2012.

[5] N. G. Kingsbury. Image processing with complex wavelets. *Philos. Trans. R. Soc. London A, Math. Phys. Sci*, 3570(1760):2543–2560, 1999.

[6] N. G. Kingsbury. The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters. In *Proc. 8th IEEE DSP Workshop*, page Paper no. 86, Aug. 912, 1998.

[7] E. Land and J. McCann. Lightness and retinex theory. *J. Optical Soc. of America A*, 61:1–11, 1971.

[8] D. Marini and A. Rizzi. A computational approach to color adaptation effects. *Image and Vision Computing*, 18(13):1005–1014, 2000.

[9] L. Menten and M. I. Michaelis. Die kinetik der invertinwirkung. *Biochem*, 49:333369, 1913.

[10] K. I. Naka and W. A. H. Rushton. S-potential from colour units in the retina of sh (cyprinidae). *Journal of Physiology*, 185:536–555, 1966.

[11] E. Provenzi, M. Fierro, A. Rizzi, L. De Carli, D. Gadia, and D. Marini. Random spray retinex: A new retinex implementation to investigate the local properties of the model. 16(1):162–171, 2007.

[12] E. Provenzi, C. Gatta, M. Fierro, and A. Rizzi. A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast. 30(10):1757–1770, 2008.

[13] A. Rizzi, D. Marini, and L. D. Carli. Lut and multilevel brownian retinex colour correction. *Machine Graphics and Vision*, 11(2-3):153–168, 2002.

[14] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury. The dual-tree complex wavelet transform: a coherent framework for multiscale signal and image processing. 6:123151, 2005.

[15] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In *2nd Annual International Conference on Image Processing*, 1995.

[16] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *Image Processing, IEEE Transactions on*, 16(9):2207 – 2214, sept. 2007.

[17] C. J. van den Branden Lambrecht, editor. *Vision models and applications to image and video processing*. Kluwer Academic, 2001.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. 13:600–612, 2004.