A Simple Inter- and Intrascale Statistical Model for Video Denoising in 3-D Complex Wavelet Domain Using a Local Laplace Distribution

Hossein Rabbani[1,2], Mansur Vafadust[1] & Saeed Gazor[2]

1) Department of Bioelectric, Faculty of Biomedical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, IRAN

2) Electrical and Computer Engineering, Queen's University, Kingston, Ontario, CANADA

Abstract

This paper presents a new video denoising algorithm based on the modeling of wavelet coefficients in each subband with a Laplacian probability density function (pdf) with local variance. Since Laplacian pdf is leptokurtic, it is able to model the sparsity of wavelet coefficients. We estimate the local variance of this pdf employing adjacent coefficients at same scale and parent scale. This local variance models interscale dependency between adjacent scales and intrascale dependency between spatial adjacent. Within this framework, we design a maximum a posteriori (MAP) estimator for video denoising, which relies on the proposed local pdf. Because separate 3-D transforms, such as ordinary 3-D wavelet transforms, have artifacts that degrade the performance of this transform, we implement our algorithm in 3-D complex wavelet transform. This non-separable and oriented transform gives a motion-based multiscale decomposition for video that isolates in its subbands motion along different directions. In addition, we use our denoising algorithm in 2-D complex wavelet transform, where the 2-D transform is applied to each frame individually. Despite the simplicity of our method in implementation, it achieves better performance than several denoising methods both visually and in terms of peak signal-to-noise ratio (PSNR).

1. Introduction

Usually, noise reduction can significantly improve visual quality of a video. The main sources of noise are arising from the electronic hardware (shot noise) and from the channels during transmission (thermal noise) [1]. Most noise sources are modeled well by additive white Gaussian noise (AWGN) model, which is also treated in this paper. Early methods of video denoising are single-resolution solutions [1]. Although video is a 3-D data set, 3-D data transforms, that are separable implementations of 1-D transforms, are usually not used for its representation. Indeed, the standard 3-D data transforms do not provide useful representations that have a good energy compaction property, for most video data. For example, the multi-dimensional standard separable discrete wavelet transform (M-D DWT) mixes orientations in its subbands, which can lead to checkerboard artifacts. Thus, in contrast to still image denoising literature [2-3], relatively few publications have addressed so far *multi-resolution* video denoising. Roosmalen *et al* [4] proposed video denoising by thresholding the coefficients of a specific 3-D multiresolution representation, which combines 2-D steerable pyramid decomposition (of the spatial content) and a 1-D wavelet decomposition (in time). A motion adaptive filtering scheme, which uses 2-D wavelet transform of several successive video frames, is proposed in [5]. Selesnick and Li [6] investigated wavelet thresholding in a non-separable 3-D discrete complex wavelet representation. Discrete complex wavelet transform (DCWT), that is a non-separable multi-dimensional wavelet transforms, overcome the serious artifacts of the separable M-D wavelet transform. This transform has good directional selectivity and its subband responses are approximately shift-invariant. With the complex wavelet transform, it is more likely that the multi-resolution framework, which has proven very effective for image denoising [2-3] can also, be effectively applied to video denoising. Because this oriented 3-D transform can represent motion information, it provides a tool for video denoising that takes into account the motion of image elements, without explicitly using motion estimation.

The problem of wavelet based video denoising can be expressed as estimation of clean coefficients from noisy data with Bayesian estimation techniques. If the MAP estimator is used for this problem, the solution requires *a priori* knowledge about the distribution of wavelet coefficients. Based on the distribution type, the corresponding estimator (shrinkage function) is obtained. For example, the classical soft threshold shrinkage function can be obtained by a Laplacian pdf [7].

The *primary property* of the wavelet transform states

that the wavelet transforms of real-world signals tend to be sparse [8]. Although long-tailed pdfs are proper for modeling this property of wavelets [9-12] but they are not able to exhibit other statistical properties of wavelets. For example the *secondary property* of wavelets expresses that if a coefficient is large/small, their adjacent in spatial and in scales are large/small. Many researchers have been proposed the bivariate pdfs for modeling the interscale dependency [7], [13-15]. Although usually these pdfs improve the denoising results but they may lead to complicated algorithms. Intrascale dependency states that pdfs using spatial local parameters are able to better capture the statistical properties of wavelets [16-17]. For example Mihcak [17] proposes a Gaussian pdf with local variance for denoising and earns impressive results with his simple algorithm.

In this paper we use a Laplace pdf with local variance to model the heavy-tailed property and interscale and intrascale dependences of wavelet coefficients. This pdf is univariate but since we estimate the local variance of each coefficients using its spatial adjacent and its parent's spatial adjacent, we incorporate both inter- and intrascale dependencies in this estimation. We propose to use this new model in the 3-D discrete complex wavelet transform domain for video noise reduction.

The rest of this paper is organized as follows. After a brief review on the basic idea of Bayesian denoising, we obtain a shrinkage function namely, *SoftL*, using the Laplace pdf with local variance in Section 2. In Section 3 we use our model for wavelet-based denoising of several videos corrupted with additive Gaussian noise in various noise levels. Despite the simplicity of our method, since proposed pdf is able to simultaneously model the heavy-tailed property and interscale and intrascale dependences of wavelet coefficients, our denoising results achieves better performance than several published methods both visually and in terms of root-mean-square-error (RMSE). Finally the concluding remarks are given in Section 4.

2. Proposed approach

In this section, we assume that input video $x=[x(1),\ldots,x(l)]$ (l is the number of pixels in all frames of video) is corrupted by AWGN $\varepsilon(k)$ of zero-mean and standard deviation σ_n. We observe a noisy signal $g(k)=x(k)+\varepsilon(k)$ for $k=1,\ldots,l$ and wish to estimate the noise-free signal $x(k)$ as accurately as possible according to some criteria. Due to linearity of the wavelet transform, the noise remains additive in the transform domain as well. Thus we can write $y(k)=w(k)+n(k)$ for $k=1,\ldots,l$, where $y(k)$ is the noisy wavelet coefficient, $w(k)$ is the noise-free wavelet coefficient and if the wavelet transform is orthogonal, $n(k)$ is independent, identically distributed (i.i.d.) normal random variables $n(k)\sim N(0,\sigma_n)$.

Regardless of the type of the employed DWT, denoising is commonly done by *wavelet shrinkage*. A common shrinkage approach is *thresholding* [2-3], which sets the wavelet coefficients with "small" magnitudes to zero while keeping ("hard-thresholding") or shrinking in magnitude ("soft-thresholding") the remaining ones. Although thresholding with a uniform per subband threshold is attractive due to its simplicity, the performance is limited and the denoising quality is often not satisfactory. Thus wavelet shrinkage methods using separate threshold in each subband have been developed over recent years [7-17].

2.1. Stochastic model for wavelet coefficients

We model image wavelet coefficients as a realization of a doubly stochastic process. Specifically, the wavelet coefficients are assumed to be conditionally independent zero-mean Laplacian random variables, given their variances. These variances are modeled as highly correlated random variables. Indeed, we model the wavelet coefficients as *conditionally independent Laplcian* random variables rather than Laplace variables. To motivate this model, see Fig. 1. This Fig shows the histogram of the wavelet coefficients of a sample video in second scale of HHL subband. The best Laplace pdf with a single variance fitted to this histogram and the best Laplace pdf with local variance fitted to this histogram are shown in this Fig. We can see that our model follows the histogram more closely.

Our proposed denoising algorithm operates in two steps. Initially, we estimate the variance for each coefficient, using the observed noisy data in a local neighborhood. The estimation of $\sigma(k)$ is then substituted for $\sigma(k)$ in the expression for the MAP estimator of $w(k)$. Both steps are summarized in Fig. 2 and are described in more detail below.

2.2. Denoising employing local Laplace pdf

If we use the MAP estimator to estimate $w(k)$ from the noisy observation, $y(k)$, we can write

$$\hat{w}(k) = \arg\max_{w(k)} p_{w(k)|y(k)}(w(k) \mid y(k)) \qquad (1)$$

Using Bayesian rule we get

$$\hat{w}(k) = \arg\max_{w(k)} \frac{p_{y(k)|w(k)}(y(k) \mid w(k)) p_{w(k)}(w(k))}{p_{y(k)}(y(k))}$$

Because the term $p_{y(k)}(y(k)$ does not depend on $w(k)$, the value of $w(k)$ that maximizes right hand side is not influenced by the denominator. Therefore the MAP estimate of $w(k)$ is given by

$$\hat{w}(k) = \arg\max_{w(k)} [p_{y(k)|w(k)}(y(k) \mid w(k)) p_{w(k)}(w(k))]$$

Because $y(k)$ is the sum of $w(k)$ and $n(k)$, a zero-mean

Gaussian pdf, when *w(k)* is a known constant, *y(k)* will be a Gaussian pdf with mean *w(k)*. Therefore, if the pdf of *n(k)* is $p_n(n(k))$, then $p_{y(k)|w(k)}(y(k)|w(k))$ will be $p_n(y(k)-w(k))$.

Thus, (1) can be written as

$$\hat{w}(k) = \arg\max_{w(k)}[p_n(y(k)-w(k))p_{w(k)}(w(k))] \quad (2)$$

We have assumed the noise is zero-mean, Gaussian with standard deviation σ_n,

$$p_n(n(k)) = \frac{1}{\sigma_n\sqrt{2\pi}}\exp(-\frac{n(k)^2}{2\sigma_n^2}) \quad (3)$$

By replacing (3) in (2) yields

$$\hat{w}(k) = \arg\max_{w(k)}[-(y(k)-w(k))^2/(2\sigma_n^2) + f(w(k))] \quad (4)$$

where $f(w(k))=log(p_{w(k)}(w(k))$.

Therefore, we can obtain the MAP estimate of *w(k)* by setting the derivative to zero with respect to $\hat{w}(k)$. That gives the following equation to solve for $\hat{w}(k)$.

$$(y(k)-\hat{w}(k))/\sigma_n^2 + f'(\hat{w}(k)) = 0 \quad (5)$$

We now need a model $p_{w(k)}(w(k))$ for the distribution of wavelet coefficients. Mihcak [17] proposed a Gaussian pdf with local variance to model wavelet coefficients. We use Laplace pdf instead Gaussian pdf as

$$p_{w(k)}(w(k)) = \frac{1}{\sigma(k)\sqrt{2}}\exp(-\frac{\sqrt{2}|w(k)|}{\sigma(k)}) \quad (6)$$

In this case $f(w(k)) = -\log(\sigma(k)\sqrt{2}) - \sqrt{2}|w(k)|/\sigma(k)$ thus $y(k) = \hat{w}(k) + (\sqrt{2}\sigma_n^2/\sigma(k)).sign(\hat{w}(k))$ that is often written in the following way

$$\hat{w}(k) = sign(y(k)).(|y(k)| - \sqrt{2}\sigma_n^2/\sigma(k))_+ \quad (7)$$

where $(a)_+=max(a,0)$. Let's define the *SoftL* operator as

$$SoftL(g(k),\tau(k)) := sign(g(k)).(|g(k)| - \tau(k))_+ \quad (8)$$

The shrinkage function (7) can be written as

$$\hat{w}(k) = SoftL(y(k), \sqrt{2}\sigma_n^2/\sigma(k)) \quad (9)$$

Fig. 3 shows a comparison between *SoftL* and soft thresholding produced from Laplace pdf for sample parameters.

2.3. Parameters estimation

To apply the *SoftL* rule we need to know σ_n and $\sigma(k)$. For each data point *y(k)*, an estimate of $\sigma(k)$ is formed based on a local neighborhood N(k). We can use a square window N(k) centered at y(k). Then we obtain an empirical estimate for $\sigma(k)$ as

$$\hat{\sigma}^2(k) = \max(\sum_{j\in N(k)} y^2(j)/M - \sigma_n^2, eps) \quad (10)$$

where M is the number of coefficients in N(k) and *eps* is a very small positive. This estimation employs the spatial adjacent coefficients in the same scale and realizes the intrascale dependency property of wavelets. But since (10) doesn't use the parent coefficient and its spatial adjacent (in coarser scale), this estimation is not able to model the interscale dependency. For this reason we correct (10) as

$$\hat{\sigma}^2(k) = \max(\sum_{j\in N(k)} \frac{y_1^2(j)+y_p^2(j)}{2M} - \sigma_n^2, eps) \quad (11)$$

where according to Fig. 4, $y_p(j)$ is the expanded parent of $y_1(j)$.

When σ_n is unknown, to estimate the noise variance from the noisy wavelet coefficients, a robust median estimator is used from the finest scale wavelet coefficients [3].

$$\sigma_n^2 = \frac{median(|y_i|)}{0.6745},\ y_i\ \epsilon\ subband\ HH\ in\ finest\ scale \quad (12)$$

3. Experimental results

This section presents video denoising examples in ordinary and complex wavelet domain to show the efficiency of our new model and compare it with other methods in literature.

To evaluate the use of the 3-D DCWT for video denoising, the *SoftL* rule is used with four different transforms. In the first two cases, we applied the 2-D ordinary and the 2-D complex wavelet transforms to each frame of the video. In the second two cases, we applied the 3-D ordinary and the 3-D complex wavelet transform to the 3-D video data. In the example (the test video *160×224×48 Salesman,* $\sigma_n=30$), the RMSE for the 2-D ordinary wavelet transform is 12.95; for the 2-D complex wavelet transform it is 11.65; for the 3-D ordinary wavelet transform it is 11.40; for the 3-D complex wavelet transform it is 10.78. The 3-D complex wavelet transform performed best for this example. A frame from the test video, the noisy video, and each of the processed videos are shown in Fig. 5. The improvement of the 3-D complex wavelet over the 2-D complex wavelet is not clearly visible in the Fig, however, the difference between them is more visible when viewing the video proper.

Fig. 6 shows one frame of the denoised video obtained using Laplace pdf (soft thresholding), circular symmetric pdf (Selesnick's work [7]), Gaussian pdf with local variance (Mihcak's work [17]) and Laplacian pdf with local variance (our method) in 3-D DCWT. The *160×224×48 Salesman* video is used for this purpose and zero-mean, white Gaussian with $\sigma_n=30$ is added to the original image.

We also tested our algorithm using different additive Gaussian noise levels $\sigma_n=10,20,30$ to three, videos, namely, *160×224×48 Salesman*, *80×80×80 Bicyclers*, *128×128×24 MRI* and compared with soft thresholding,

Selesnick's work [7] and Mihcak's work [17]. Performance analysis is done using the PSNR measure. The results can be seen in Table.1. Each PSNR value in the table is averaged over ten runs. In this table, the highest PSNR value is bolded. As seen from the results, our algorithm mostly outperforms the others.

4. Conclusion and future works

In this paper we use *SoftL* function based on Laplace pdf with local variance to model the heavy-tailed property and interscale and intrascale dependences of wavelet coefficients in each subband. This pdf is univariate but since we estimate the local variance of each coefficients using its spatial adjacent and its parent's spatial adjacent, we incorporate both inter- and intrascale dependencies in this estimation. Experiments show our model has better visual results than other methods such as soft thresholding. In order to show effectiveness of new estimator, we compared *SoftL* method with effective techniques in the literature and we see that our denoising algorithm mostly outperforms the others.

Instead of this shrinkage function, other nonlinear shrinkage functions can be used. For example, instead of using Laplace pdf we can use generalized Gaussian distribution, mixture pdfs or bivariate pdfs or instead of using the MAP estimator we can use the *minimum mean squared error* (MMSE) estimator.

Since our proposed estimation for variance is empirical, we can improve the variance estimation using ML and MAP estimator.

5. References

[1] 1. J. Brailean, R. Kleihorst, S. Efstratiadis, A. Katsaggelos, A. Lagendijk, *"Noise reduction filters for dynamic image sequences: A review,"* Proceedings of the IEEE 83(9), pp. 1272–1292, 1995.

[2] D. L. Donoho, I. M. Johnstone, *"Ideal spatial adaptation by wavelet shrinkage,"* Biometrika, vol. 81, no. 3, 1994.

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

[4] P. Roosmalen, S. Westen, R. Lagendijk, J. Biemond, *"Noise reduction for image sequences using an oriented pyramid thresholding technique,"* in Proc. IEEE Conf. on Image Process., pp. 375–378, (Switzerland, Sep. 1996).

[5] V. Zlokolica, A. Pizurica, W. Philips, *"Video denoising using multiple class averaging with multiresolution,"* in International Workshop VLBV03, (Madrid, Spain, Sep. 2003).

[6] I. Selesnick, K. Li, *"Video denoising using 2d and 3d dual-tree complex wavelet transforms,"* in Proc. SPIE Wavelet Applications in Signal and Image Processing, 5207, (San Diego, Aug 4-8, 2003).

[7] L. Sendur, I. Selesnick, *"Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency,"* IEEE. Tran. Signal Processing, vol.50, no. 11, pp. 2744–2756, 2002.

[8] M. S. Crouse, R. D. Nowak, R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Processing,* vol. 46, no. 4, pp. 886-902, Apr. 1998.

[9] M. Hansen, B. Yu,"Wavelet thresholding via MDL for natural images," *IEEE Trans. on Information Theory,* vol. 46, no. 5, 2000.

[10] A. Achim, P. Tsakalides, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Trans. Geosci. Remote Sensing,* vol. 41, no. 8, pp. 1773–1784, Aug. 2003.

[11] M. A. T. Figueiredo and R. D. Nowak, "Wavelet-based image estimation: An empirical Bayes approach using Jeffrey's noninformative prior," *IEEE Trans. on Image Processing,* vol. 10, pp. 1322–1331, Sept. 2001.

[12] H. Gao, "Wavelet shrinkage denoising using the non-negative garrote," *J. Comput. Graph. Stat.,* vol. 7, pp. 469–488, 1998.

[13] V. Strela, J. Portilla, and E. Simoncelli, "Image denoising using a local Gaussian scale mixture model in the wavelet domain," *in Proc. SPIE 45th Annu. Meet.,* 2000.

[14] H. Rabbani, M. Vafadust, I. Selesnick, S.Gazor "Image Denoising Based on a Mixture of Bivariate Laplacian Models in Complex Wavelet Domain," in *Proc. IEEE International Workshop on Multimedia Signal Processing,* Victoria, BC, Canada, October 3-6, 2006.

[15] A. Achim and E. E. Kuruoglu. "Image Denoising Using Bivariate alpha-Stable Distributions in the Complex Wavelet Domain," *IEEE Signal Processing Letters*, 12, pp. 17-20, 2005.

[16] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Processing,* vol. 50, no. 11, pp. 2744–2756, Nov. 2002.

[17] M. K. Mihcak, I. Kozintsev, K. Ramchandran, P. Moulin, *"Low complexity image denoising based on statistical modeling of wavelet coefficients,"* IEEE Signal Processing Letters, vol. 6, pp. 300–303, Dec. 1999.

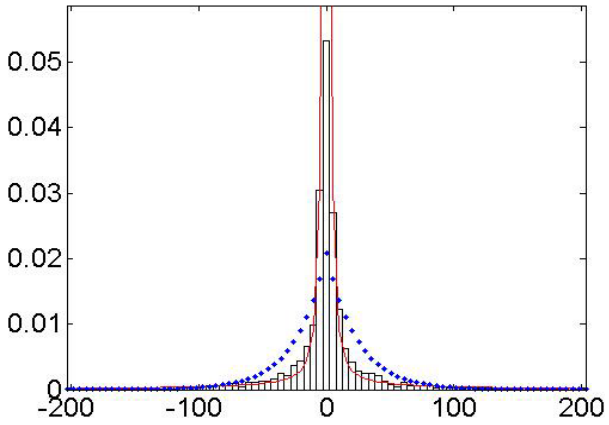

Fig. 1. Histogram of a test video in second scale of HHL subband. Dotted line: the best Laplace pdf with local variance fitted to this histogram. Solid Line: the best Laplace pdf with a single variance fitted to this histogram.

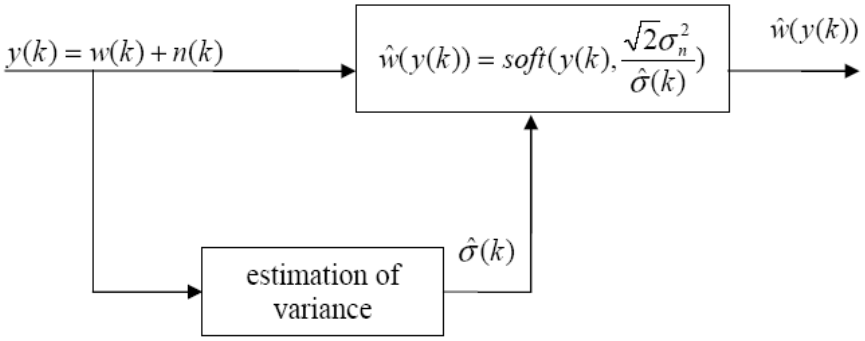

Fig. 2. Block diagram of the denoising algorithm. For each observed noisy co-efficient *y(k),* we find an estimation of the variance of *w(k)* based on a local neighborhood. The estimate of *σ(k)* is then used for MAP estimation of *w(k)*.

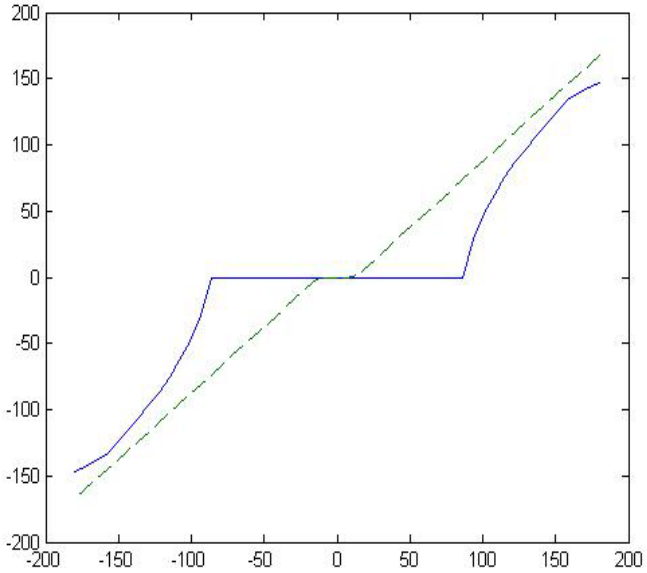

Fig. 3. A comparison between *SoftL* (solid line) and soft threshold function (dashed line).

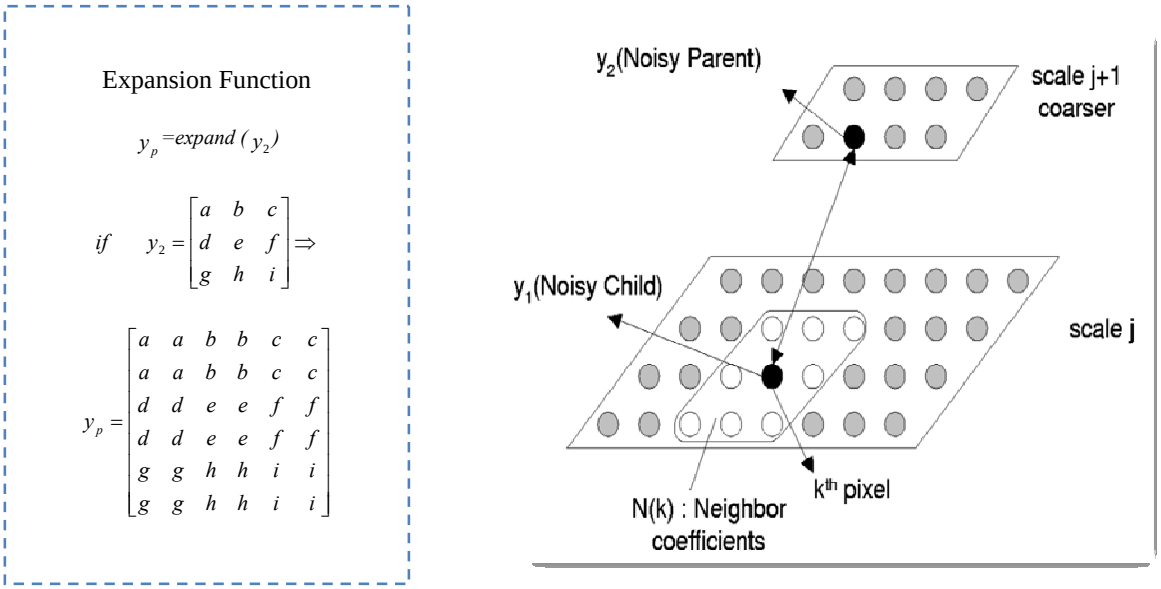

Fig. 4. Left image: the expantion function. Right image: illustration of neighborhood N(k) [16].

Fig. 5. From top left, clockwise: A frame from the test video, the processed video in 2-D DWT, the processed video in 2-D DCWT, the processed video in 3-D DCWT, the processed video in 3-D DWT and the noisy video.

Fig. 6. From top left, clockwise: A frame from the noise-free video, the processed video with hard thresholding, the processed video with Mihcak's work, the processed video with *SoftL* (our method), the processed video with soft thresholding and the noisy video.

Table 1. PSNR Results in Decibel for Several Denoising Methods in 3-D Complex Wavelet Domain.

σ_n	Hard Thresh.	Soft Thresh.	Selesnick's Method	Mihcak's Method	Our Method
Salesman					
10	30.85	31.97	32.02	32.48	**33.25**
20	28.72	28.76	29.35	29.66	**29.90**
30	27.32	26.89	27.40	27.83	**28.14**
Bicyclers					
10	31.04	32.46	32.58	33.36	**33.78**
20	28.26	28.91	29.25	29.86	**30.37**
30	27.23	26.99	27.49	27.97	**28.49**
MRI					
10	32.84	33.79	34.11	**35.19**	35.18
20	30.60	30.34	30.74	31.89	**31.97**
30	29.04	28.93	29.12	29.95	**30.21**