Asian Journal of Information Technology 13 (5): 267-273, 2014
ISSN: 1682-3915


# An Overview of Pixel Domain and Wavelet Domain Filters for Video Denoising

[1]V. Sasikala, [2]N. UmaMaheswari and [3]R. Venkatesh
[1]Department of Computer Science and Engineering,
University College of Engineering, Thirukkuvalai, Nagapattinam, India
[2]Department of Computer Science and Engineering,
[3]Department of Information Technology,
PSNA College of Engineering and Technology, Dindigul, India

**Abstract:** Vision is the most important source of information for humans and video has become a part of the everyday life. Video processing is a particular case of signal processing where the input and output signals are video files or video streams. Video denoising is especially interesting for surveillance systems but als for television and medical video. Many video denoising techniques have been published in the past two decades. They vary in a wide range of complexity, performance and implementation cost. The search for efficient video denoising methods is still a valid challenge at the crossing of functional analysis and statistics. The main focus of this study is to compare various Video Denoising algorithms.

**Key words:** Video denoising, filters, cost, output, signals

## INTRODUCTION

The human vision system is a highly advanced and complex image processing sensor. It automatically tells what people really want and discards the useless details. The video signals are often contaminated by noise during acquisition, storage and transmission. The presence of noise not only results in an unpleasant visual appearance but also impose an adverse effect on the performance of subsequent video processing tasks such as video compression, analysis, object tracking and pattern recognition. Therefore, video denoising is a highly desirable and essential step in video processing systems.

The development of an advanced video denoising scheme is essential. Video denoising techniques can be considered as an extensions of image denoising techniques by providing temporal filtering taking into account the correlation between the neighboring frames. I.e., there exists a high correlation among the neighboring frames of a video, since the motions among such frames are small.

## CLASSIFICATION OF DENOISING ALGORITHMS

Video denoising algorithms can be classified into four categories:

**Spatial domain video denoising:** It is a way of utilizing spatial correlation of video content to suppress noise. It is normally implemented with a weighted local 2D or 3D windows and the weights can be either fixed or adapted based on the image content. However, spatial only denoising is rarely considered in real applications as it often leads to visible artifacts.

**Temporal domain video denoising:** It is an approach of exploiting temporal correlations to reduce noise in a video sequence. A video sequence contains not only spatial correlation but also temporal correlation between consecutive frames. Normally, motion estimation methods which can be based on block matching or optical flow are employed to find the prediction of the reference block. For each reference block, its temporal predictions are combined with the block itself to suppress noise.

**Spatio-temporal video denoising:** This approach exploits both spatial and temporal correlations in video sequence to reduce noise. In many real video applications, spatio-temporal filtering performs better than temporal filtering and the best performance can be achieved by exploiting information from both past and future frames. 3D non-local means and VBM3D are some Spatio Temporal Denoising Methods.

**Transform domain video denoising:** Transform Domain Video Denoising Methods first decorrelate the noisy

**Corresponding Author:** V. Sasikala, Department of Computer Science and Engineering,
University College of Engineering, Thirukkuvalai, Nagapattinam, India

signal using a linear transform (e.g., DCT or Wavelet transform) and then recover the transform coefficients. Then this signal is subjected to inverse transform to get the signal back to spatial domain. Typically, transform domain methods are used together with Temporal or Spatial Domain Denoising Methods.

## OVERVIEW OF SPATIAL DOMAIN ALGORITHMS

**Introduction to spatial-domain denoising:** Denoising Methods where the pixel intensities are used directly in the denoising process are said to be spatial-domain filters. Even within this class of denoising methods, the actual approaches can vary significantly. In general, the most successful approaches can be classified as being either a process where denoising is performed by a weighted averaging of pixel intensities or an explicit model-based approach where parameters of the data model are usually learned from the noisy image itself.

**Weighted Averaging Methods:** The underlying concept behind many spatial-domain denoising filters is to suppress noise through a weighted averaging process. Pixels across edges are averaged in the denoising process, leading to loss of detail and edge sharpness. To restrict such loss of detail in the image, it is important to ensure that the averaging is performed only over photometrically similar pixels. One of the first approaches making use of a data adaptive weight function is attributed to Smith and Brady (1997) (SUSAN) and Tomasi and Manduchi (1998) (bilateral filter). Buades *et al.* (2005a) and Awate and Whitaker (2006) independently proposed a simple modification that lends robustness to this weight function. Instead of comparing intensities of a pair of pixels, intensities of local groups of pixels (patches) are compared using patch based weight function. So, the weight calculation scheme is considerably better at rejecting dissimilar pixels from the averaging process.

**Denoising through data modeling:** In most of these methods the models act as prior information about the clean image and are either learned a priori from noise-free natural images or directly from the noisy image. Denoising is then performed by enforcing these priors on the noisy image. One of the most popular model-based methods is the K-SVD algorithm (Aharon *et al.*, 2006). There the authors propose a patch-based framework where each patch in the image is represented as a linear combination of patches from some over-complete set of bases. Building on the observation that noise-free image patches are sparse-representable (Elad and Aharon, 2006). Researchers enforce a constraint on the number of basis patches (or atoms) that can be used in estimating any given patch.

**Non-Local Means algorithm:** The NLM is the Motion-Estimation-Free Video Denoising algorithm and yet, it is also the simplest. As such it stands as a good candidate for generalization. The NLM is posed originally by Buades *et al.* (2005b) as a single image denoising method, generalizing the well-known bilateral filter (Tomasi and Manduchi, 1998; Elad, 2002). Denoising is obtained by replacing every pixel with a weighted average of its neighborhood. The weights for this computation are evaluated by using block-matching fit between image patches centered around the center pixel to be filtered and the neighbor pixels to be averaged. Recent research has shown how this method can be used for video denoising by extending the very same technique to 3D neighborhoods (Buades *et al.*, 2005b). An improvement of this technique, considering varying size neighborhoods is suggested by Boulanger *et al.* (2007) so as to trade bias versus variance in an attempt to get the best Mean Squared Error (MSE).

**CIFIC video denoising scheme:** Combined interframe intercolor prediction is proposed by Dai *et al.* (2013). The inter frame and inter color correlation is decomposed into three kinds of correlation and a comparison is made between them. First one is pure inter frame correlation which is the correlation between the same color component of the current pixel and its motion compensated pixel in the neighboring reference frame. Second one is pure inter color correlation which is the correlation between any pair of color components at the same spatial position and in the same frame. The last one is the inter frame-inter color correlation which is the correlation between one color component of the current pixel and a different color component of its motion compensated pixel in the reference frame. The correlation coefficients are calculated block by block (Fig. 1).

## OVERVIEW OF TRANSFORM DOMAIN ALGORITHMS

**Introduction Totransform-Domain Methods:** The main motivation of denoising in transform domain is that in the transformed domain it may be possible to separate image and noise components. The basic principle behind most Transform-Domain Denoising Methods is shrinkage-truncation (hard thresholding) or scaling (soft thresholding) of the transform coefficients to suppress the effects of noise. For such thresholding, the challenge

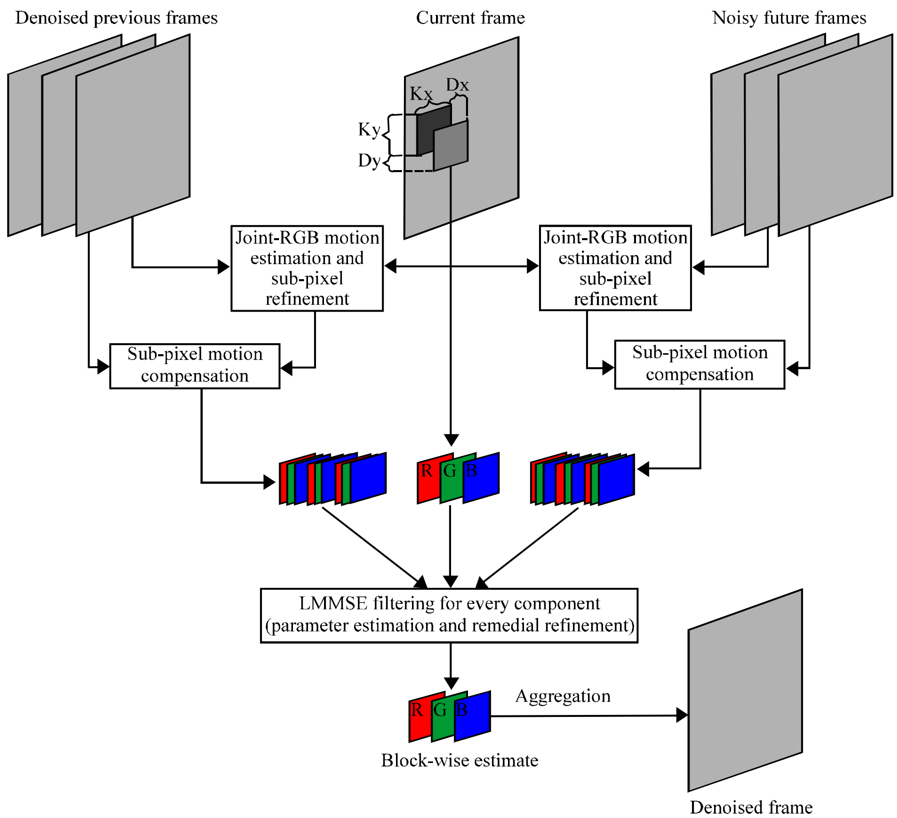


Fig. 1: Framework of CIFIC proposed by Dai *et al*. (2013)

is to develop a suitable coefficient mapping operation that does not sacrifice the details in the image. The final denoised image is obtained by performing an inverse transform on the shrunk coefficients. Apart from the choice of the thresholding operator, the choice of the transform domain is also critical.

**Introduction to wavelet:** Wavelet transform provides an elegant solution for the analysis of non-stationary signals. Wavelets are a useful tool for signal processing applications such as image compression and denoising. There are two broad approaches for the wavelet domain spatio-temporal filtering for video denoising. First approach is the thresholding of the coefficients of the 3D wavelet representation of a noisy video. Video denoising based on the 3D wavelet transform avoids the procedure of motion estimation and detection. However, the main drawbacks of the 3D wavelet transform are its inability to make use of the asymmetry of space and time resolutions that exist in a video and its long time latency in view of the memory requirements being constrained by the length of the filter coefficients.

The second approach is to represent each frame by the 2D wavelet transform and perform spatial filtering using one of the available denoising techniques for still images this is followed by temporal filtering by taking into account the strong temporal correlation that exists in video.

There are many types of 3D and 2D wavelet representations, orthogonal/biorthogonal, real/complex valued and decimated/redundant.

Scalar wavelets are wavelets generated by one scalar function. Multi wavelets are several wavelets with more than one scaling function. Scalar wavelets cannot possess all the properties (short support, orthogonality, symmetry and vanishing moments) at the same time. A multi wavelet system can simultaneously provide perfect reconstruction while preserving length (orthogonality), good performance at the boundaries (via linear phase symmetry) and a high order of approximation (vanishing moments).

The space-frequency localization property of the wavelet domain makes it the most popular choice. There has been several other developments of directional wavelet systems with the same goal, namely Steerable wavelets (Freeman and Adelson, 1991; Simoncelli *et al*., 1992), Gabor wavelets (Lee, 1996), Wedgelets (Donoho, 1999), Beamlets (Donoho and Huo, 2002), Bandlets

(Mallat and Peyre, 2007; Le Pennec and Mallat, 2005), Contourlets (Do and Vetterli, 2005), Shearlets (Labate *et al*., 2005; Guo and Labate, 2007), Wave atoms (Demanet and Ying, 2007), Platelets (Willett and Nowak, 2003) and Surfacelets (Luisier *et al*., 2010). These geometric wavelets or directional wavelets are uniformly called X-lets.

**Fourier Transform (FT):** The FT decomposes a signal in complex exponential functions at different frequencies. The poor time localization is the main disadvantage of the Fourier transform, making it not suitable for all kind of applications.

**Discrete Cosine Transform (DCT):** Discrete Cosine Transform (DCT) is one of the most widely used transformation operation for image and video coding. It is a variation of the Discrete Fourier Transform (DFT). It transforms an N-point time domain signal into N-point frequency domain coefficients. Where the DFT consists of real and imaginary coefficients, the DCT only has real coefficients. The DCT has a very high energy de-correlation ability that is suitable for decomposing highly correlated natural image/video contents.

**Discrete Wavelet Transform (DWT):** Although, the discrete wavelet transform (Smith and Brady, 1997) has established an impressive reputation as a tool for mathematical analysis and signal processing, it has the disadvantage of poor directionality. Wavelets are not very efficient when dealing with multidimensional functions and signals. This limitation is due to their poor directional sensitivity and limited capability in dealing with the anisotropic features which are frequently dominant in multidimensional applications. To overcome this limitation, a variety of methods have been recently introduced to better capture the geometry of multidimensional data, leading to reformulate wavelet theory and applied fourier analysis within the setting of an emerging theory of sparse representations.

**SURE-LET for orthonormal wavelet domain video denoising:** Stein's Unbiased Risk Estimator-Linear Expansion of Thresholds (SURE-LET) introduced by (Luisier *et al*., 2010), the principle is to parameterize the wavelet estimator as a linear expansion of thresholds and minimize an extended version of Stein's unbiased risk estimator to determine the best linear parameter of this expansion. To increase the correlation between adjacent frames, a global motion compensation followed by a selective block matching is first applied, the a multiframe interscale wavelet thresholding is performed to denoise the current central frame.

**The curvelet transform:** The curvelet transform (Ma and Plonka, 2010) is a multi scale directional transform that allows an almost optimal non adaptive sparse representation of objects with edges. The discrete curvelet transform (Starck *et al*., 2002) is very efficient in representing curve-like edges.

**The 3D discrete curvelet transform:** The 3D curvelet functions depend on four indices instead of three, the scale, the position and two angles. The 3D discrete curvelet (Ying *et al*., 2005) aims at frequency partitioning. The curvelet implementation starts from defining a mother curvelet in the fourier domain whose scaled and sheared copies form a partition of a unity. The curvelet coefficients are then obtained by multiplying the fourier samples of the input signal with curvelet window functions at different scales and directions followed by a spatial down sampling (implemented by frequency wrapping).

The curvelet systems have two main drawbacks. They are not optimal for sparse approximation of curve features beyond $C^2$-singularities and the discrete curvelet transform is highly redundant.

**The contourlet transform:** Contourlets as proposed by Do and Vetterli (2005) form a discrete filter bank structure that can deal effectively with piece wise smooth images with smooth contours. The contourlet transform can be seen as a discrete form of particular curvelet transform. The difference between contourlets and curvelets is that the contourlet transform is directly defined on digital-friendly discrete rectangular grids. But, it has less clear directional geometry features than curvelets (Fig. 2).

In an efficient directional multi resolution image representation using contourlet transform starts with a discrete domain construction and then its convergence to an expansion in the continuous domain. This construction results in a flexible multi resolution, local and directional image expansion using contour segments and thus it is named the contourlet transform.

**The surfacelet transform:** Surfacelets are 3D extensions of the 2D contourlets that are obtained by a higher-dimensional directional filter bank and a multi scale pyramid. They can be used to efficiently capture and represent surface like singularities in multi dimensional volumetric data.

Table 1: Video denoising performance using different video sequences

| PSNR (dB) | Mobile | | | Coastguard | | | Tempete | | |
|---|---|---|---|---|---|---|---|---|---|
| Noise (σ) | 30 | 40 | 50 | 30 | 40 | 50 | 30 | 40 | 50 |
| 3DCURV (Ying *et al*., 2005) | 23.54 | 23.19 | 22.86 | 25.05 | 24.64 | 24.29 | - | - | - |
| SURF (Lu and Do, 2007) | 28.39 | 27.18 | 26.27 | 26.82 | 25.87 | 25.15 | 24.20 | 23.26 | 22.61 |
| 3D SHEAR (Pooran and Demetrio, 2012) | 28.68 | 27.15 | 25.97 | 27.36 | 26.10 | 25.12 | 25.24 | 23.97 | 22.81 |
| 2D SHEAR (Easley *et al*., 2008) | 25.97 | 24.40 | 23.20 | 25.20 | 23.82 | 22.74 | 22.89 | 21.63 | 20.75 |
| DWT (Smith and Brady, 1997) | 24.93 | 23.94 | 23.03 | 24.34 | 23.44 | 22.57 | 22.09 | 21.5 | 20.92 |
| VBM3D (Dabov *et al*., 2007) | 25.98 | 24.04 | 21.71 | - | - | - | - | - | - |
| NLM (Buades *et al*., 2005b) | 25.36 | 23.60 | 21.93 | - | - | - | - | - | - |

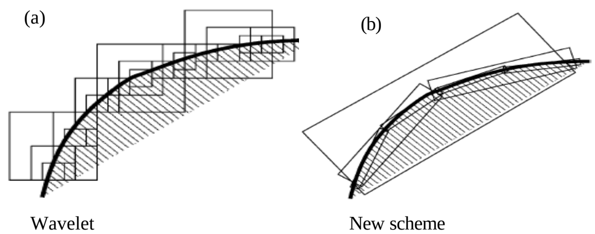


Fig. 2: Wavelet versus new scheme: illustrating the successive refinement by te two systems near a smooth contour which is shown as thick curve separating two smooth regions; a) walvet and b) new scheme

The surfacelets proposed by Lu and Do (2007) includes new set of tools, namely the N-dimensional Directional Filter Banks (NDFB) and surfacelets that can capture and represent signal singularities lying on smooth surfaces. They combined the NDFB with multi scale pyramid and constructed the surfacelet transform. The surfacelet transform is less redundant than 3D curvelet transform and this advantage is payed by a certain loss of directional features.

**The shearlet transform:** The shearlet representation, originally introduced by Guo *et al*. (2006, 2004) has emerged in recent years as one of the most effective frameworks for the analysis and processing of multidimensional data. This representation is part of a new class of multiscale methods introduced during the last 10 years with the goal to overcome the limitations of wavelets and other traditional methods through a framework which combines the standard multiscale decomposition and the ability to efficiently capture anisotropic features. Unlike curvelets, the Shearlets form an affine system with a single generating mother shearlet function parameterized by a scaling, a shear and a translation parameter where the shear parameter captures the direction of singularities.

**3D Discrete Shearlet Transform (3D DShT):** The new algorithm introduced by Negi and Labate (2012) can be described as the cascade of a multiscale decomposition based on a version of the Laplacian pyramid filter followed by a stage of directional filtering. The main novelty of the 3D approach consists in the design of the directional filtering stage which attempts to reproduce the frequency decomposition faithfully provided by the corresponding mathematical transform by using a method based on the Pseudo-Spherical Fourier transform.

**VBM3D algorithm:** The VBM3D, reported by Elad and Aharon (2006) uses a multitude of patches in the three-dimensional neighborhood of each pixel for attenuating the noise. However, the patches are used in a different manner. The most similar patches in the neighborhood are collected and stacked into a 3D array. A 3D wavelet transform is then applied with hard-thresholding used for noise suppression. After the inverse transform is applied, the patches are returned to their original locations and averaged. A second iteration follows with Wiener filtering used to improve denoising results.

## COMPARISON OF DENOISING ALGORITHMS USING DIFFERENT VIDEO SEQUENCES

The denoising algorithms were tested on three video sequences, called mobile, coast guard and tempete for various values of the standard deviation σ of the noise (values σ = 30, 40, 50 were considered). All these video sequences have been resized to 192×192×192. The performance of the denoising algorithms are given in Table 1. The data in Table 1 shows that the 3D discrete shearlet algorithm and surfacelets based denoising algorithms outperforms or is essentially equivalent. But, comparing the running times of these two transforms surfacelets are better.

## CONCLUSION

Together with the overview, this study can be seen as a preparation to the comparative study of filters for video denoising.

## REFERENCES


Aharon, M., M. Elad and A. Bruckstein, 2006. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Process, 54: 4311-4322.

Awate, S.P. and R.T. Whitaker, 2006. Unsupervised, information-theoretic, adaptive image filtering for image restoration. IEEE Trans. Pattern Anal. Mach. Intell., 28: 364-376.

Boulanger, J., C. Kervrann and P. Bouthemy, 2007. Space-time adaptation for patch-based image sequence restoration. IEEE Trans. Pattern Anal. Mach. Intell., 8: 1096- 1102.

Buades, A., B. Coll and J.M. Morel, 2005a. Denoising image sequences does not require motion estimation. Proceedings of the IEEE Conference on Advanced Video and Signal based Surveillance, September 15-16, 2005, Como, Italy, pp: 70-74.

Buades, A., B. Coll and J.M. Morel, 2005b. A review of image denoising algorithms, with a new one. Multiscale Model. Simul., 4: 490-530.

Dabov, K., A. Foi and K. Egiazarian, 2007. Video denoising by sparse 3D transformdomain collaborative filtering. Proceedings of the 15th European Signal Processing Conference, September 3-7, 2007, Poznan, Poland, pp: 146-149.

Dai, J., O.C. Au, C. Pang and F. Zou, 2013. Color video denoising based on combined interframe and inter color prediction. IEEE Trans. Circ. Syst. Video Technol., 23: 128-141.

Demanet, L. and L. Ying, 2007. Waave atoms and sparsity of oscillatory patterns. Appl. Comput. Harmon. Anal., 23: 368-387.

Do, M.N. and M. Vetterli, 2005. The contourlet transform: An efficient directional multiresolution image representation. IEEE Trans. Image Process., 14: 2091-2106.

Donoho, D. and X. Huo, 2002. Beamlets and Multi Scale Image Analysis. In: Multi Scale and Multi Resolution Methods, Barth, T., T. Chan and R. Haimes (Eds.)., LNCE. Vol. 20, Springer, Berlin, pp: 149-196.

Donoho, D., 1999. Wedgelets: Nearly minimax estimation of edges. Ann. Statist., 27: 859-897.

Easley, G., D. Labate and W.Q. Lim, 2008. Sparse directional image representations using the discrete shearlet transform. Applied Comput. Harmonic Anal., 25: 25-46.

Elad, M. and M. Aharon, 2006. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process., 15: 3736-3745.

Elad, M., 2002. On the origin of the bilateral filter and ways to improve it. IEEE Trans. Image Process., 11: 1141-1151.

Freeman, W.T. and E.H. Adelson, 1991. The design and use of steerable filters. IEEE Trans. Pattern Anal. Mach. Intell., 13: 891-906.

Guo, K. and D. Labate, 2007. Optimally sparse multidimensional representation using shearlets. SIAM J. Math. Anal., 39: 298-318.

Guo, K., G. Kutyniok and D. Labate, 2006. Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators. In: Wavelets and Splines, Chen, G. and M. Lai (Eds.). Nashboro Press, USA., pp: 189-201.

Guo, K., W.Q. Lim, D. Labate, G. Weiss and E. Wilson, 2004. Wavelets with composite dilations. Electron. Res. Announc. Am. Math. 10: 78-87.

Labate, D., W.Q. Lim, G. Kutyniok and G. Weiss, 2005. Sparse multi dimensional representation using shearlets. Proc. SPIE., 5914: 254-262.

Le Pennec, E. and S. Mallat, 2005. Sparse geometrical image approximation with bandlets. IEEE Trans. Image Process., 14: 423-438.

Lee, T.S., 1996. Image representation using 2D gabor wavelets. IEEE Trans. Pattern Anal. Mach. Intell., 18: 959-971.

Lu, Y.M. and M.N. Do, 2007. Multidimensional directional filter banks and surfacelets. IEEE Trans. Image Process, 16: 918-931.

Luisier, F., T. Blu and M. Unser, 2010. SURE-LET for orthonormal wavelet-domain video denoising. IEEE Trans. Circuits Syst. Video Technol., 20: 913-919.

Ma, J. and G. Plonka, 2010. The curvelet transform. IEEE Signal Process. Mag., 27: 118-133.

Mallat, S. and G. Peyre, 2007. A review of Bandlet methods for geometrical image representation. Numer. Algorithms, 44: 205-234.

Negi, P.S. and D. Labate, 2012. 3-D discrete shearlet transform and video processing. IEEE Trans. Image Process., 21: 2944-2954.

Simoncelli, E.P., W.T. Freeman, E.H. Adleson and D.J. Heeger, 1992. Shiftable multiscale transforms. IEEE Trans. Inform. Theory, 38: 587-607.

Smith, S.M. and J.M. Brady, 1997. SUSAN-A new approach to low level image processing. Int. J. Comput. Vision, 23: 45-78.

Starck, J.L., E.J. Candes and D.L. Donoho, 2002. The curvelet transform for image denoising. IEEE Trans. Image Process., 11: 670-684.

Tomasi, C. and R. Manduchi, 1998. Bilateral filtering for gray and color images. Proceedings of the 6th International Conference on Computer Vision, January 4-7, 1998, Bombay, India, pp: 839-846.

Willett, R.M. and R.D. Nowak, 2003. Platelets: A multiscale approach for recovering edges and surfaces in photon-limited medical imaging. IEEE Trans. Med. Imag., 22: 332-350.

Ying, L., L. Demanet and E. Candes, 2005. 3D discrete curvelet transform. Proc. SPIE, 5914: 351-361.