Abstract

Images get corrupted with noise during its acquisition, processing, signal conditioning and transmission through noisy channels. The process of removing noise from images is termed as image denoising. Multiresolution methods of denoising were found to be superior compared to many other denoising methods previously existed. When images contain directional features such as contours, curves, edges, lines, etc the Contourlet as well as Curvelet transform methods of denoising could perform better than the earlier denoising methods using discrete wavelet transform (DWT), the pioneer in multiresolution image denoising. Some drawbacks of the Contourlet Transform method was solved by the more recently proposed Non-Sub sampled Contourlet transform (NSCT). The Support Vector Machine (SVM) method of classification of noisy and non-noisy pixels and subsequent thresholding was found to give better performance compared to direct thresholding to remove noise from NSCT decomposed images. Further, use of Orthogonal Matching Pursuit (OMP) after SVM classification can dispense with thresholding and is seen even better than the previous methods of denoising using NSCT using SVM and thresholding.

Keywords: image denoising; multiresolution; NSCT; thresholding; SVM; OMP.
1. INTRODUCTION
Recently, extensive use of images are seen in several technical and non-technical applications. Portrait images, video images, biomedical images, geophysical images, satellite images, micro-array images, forensic images, etc are only some of such images to mention. Due to some pitfalls in the technology of acquiring, processing, signal conditioning, transmitting of images, noise creeps in images. Because of the presence of noise, not only that the images loose visual quality but also that many important features get submerged in noise and any feature extraction from such noisy images becomes non-trivial problems. Hence, image denoising is an essential pre-processing step for any successful higher level image processing such as recognition, identification, authentication, compression, segmentation, medical diagnostic investigation, forensic verification, etc. Image denoising is a well studied problem and plenty of research activities have contributed to it yielding many successful image denoising algorithms. However as the problem is so complex with so many denoising parameters involved in it, more and more sophistications are sought after with a view of improving the denoising qualities. Hence image denoising is a sustainable research activity [1].

Noise in images are mathematically described by some probabilistic functions such as Gaussian, Impulse, Poisson, Rician, Speckle, etc [2]. In general, a combination of these noises can also occur in images. Proper restoration of signal from noisy images require knowledge of the type of noise in images. However denoising algorithm can be said to be successful if only noise is removed without degradation of the important features and singularities such as contours, curves, lines, textures, etc present in the images and therefore image denoising is still a challenging problem.

2. RELATED PAST WORKS IN IMAGE DENOISING
Image denoising can be classified into (1) pixel domain and (2) frequency domain. In pixel domain, each pixel in the noisy image is modified with a view of removing noise from the image. As most of the noises are in the high frequency domain, the denoising algorithm generally implement some kind of low pass filtering. Gaussian filter, averaging filter [2] are some examples of such filters. Draw backs of such filters are that along with denoising, the images get blurred and edges are smeared out resulting in visually un pleasant denoised image with loss of important features. A different type of pixel filter is the order-statistic filters in which the median filter [2] is a prominent one; usually used to remove impulse (salt & pepper) noise. It preserves edges but performance is not good for denoising other types of noises. The bilateral [3] is a nonlinear, edge preserving, noise reducing smoothing filter and it takes in to account of local characteristics based on pixel distance and pixel intensity characteristics for image denoising.

In frequency domain denoising, image is transformed from pixel domain to frequency domain using image transforms, processing is done in frequency domain and then inverse transformed by the respective inverse image transform to get denoised image.
Performance of frequency domain processing is much better than that of pixel domain processing, considering the visual quality as well as noise removal are concerned. The multiresolution capability of some transforms has become attractive in image denoising due to (1) its compatibility with human visual sensation [4] characteristics and (2) its ability to decompose noisy image into many sub bands that differ in frequency domain and thereby separation of signal and noise becomes easier and better. The ability of a transform to deliver sparse decomposed output [5] in which signals of interest have large amplitudes and noise to be removed have small amplitudes is considered as highly desirable characteristic in image denoising, since denoising is more perfect in this case.

Use of Singular Value Decomposition (SVD) [6] and Principal Component Analysis (PCA) [7] are efficient tools for image analysis and denoising in which after decomposition using any of these tools, pixels in the noisy image can be classified as significant ones and non-significant ones and removal of non-significant ones, which are essentially the noisy pixels, can lead to denoising of images. In Higher Order Singular Value decomposition, which is one step ahead of SVD method, similar patches of pixels, considering some similarity in pixel intensity, of noisy image are formed. These patches are grouped to get 3D/4D patterns. SVD applied to these higher order matrix, termed as HOSVD [8], could remove noise in a better way by coefficient truncation of non-significant coefficients resulted from HOSVD analysis of noisy image. In [9], local grouping of similar pixels are done and by applying PCA twice, once by using the input noisy image and then using this processed image as input for the second stage. Further usage of bilateral filter provided good denoising effect. Block matching in three dimension (BM3D) algorithm [10] groups similar pixels and then projects as wavelets and then removing noisy pixels gave good results of denoising. Use of Partial Differential Equation (PDE) is an anisotropic [11] method of denoising to remove noise from noisy images while keeping singularities. The blocking effect of such method could be eliminated using relaxed median filter [12].

A good method of despeckling medical ultrasound images using Contourlet Transform was proposed in [13]. Wavelet methods of denoising medical images are proposed in [14,15]. A new method based on wavelet and cycle spinning with an intention to overcome the drawbacks of discrete wavelet (DWT) was proposed in [16]. The Curvelet Transform and Contourlet Transform are second generation wavelets proposed to overcome some drawbacks of DWT and to achieve denoising many natural images containing curve and edge singularities [17,18].

Natural images invariably contain many singularities like curves, edges, contours, lines, textures, etc., which are information bearing structures and are indispensable to retain the identity of it. A good denoising algorithm should not do any damage to these singularities. The DWT based denoising had been very popular due to its properties such as multiresolution capability, sparsity, availability of efficient and fast implementation. However the DWT performs poorly when images contain singularities. Although the basis elements of DWT can detect point discontinuities, it is rather very difficult to represent curve or contour singularities. The wavelet style
painter is limited to using square shaped brush strokes along the contours which requires many wavelet coefficients from many scales for its representation. To capture contours, it requires many fine dots at finer resolutions. An associated problem with DWT is its limited directional (namely horizontal, vertical and diagonal) decomposition capability. As natural images (like palm print, finger print, seismic, MRI, etc) are composed of many directional information, DWT decomposition do not provide the required features for satisfactory reconstruction. Yet another problem is its limited redundancy due to its critical sampling which leads to artifacts. Use of undecimated DWT \[ \] is one of the possible solutions of these problems.

The Contourlet Transforms proposed by Do and Vitterli [18] provides a satisfactory solution to the above problems. It is a multiscale, multidirectional transform which has distinct capacity of capturing the contours of natural images. Compared to DWT, the Contourlet Transforms require only fewer coefficients to represent discontinuities of images. The basis elements of Contourlet transforms are localized in spatial and frequency domain and are oriented at a variety of directions. They are composed of variety of elongated shapes with different aspect ratios at multiple resolutions and therefore can capture smooth contours and anisotropic features in natural images.

The coefficients of Contourlet transforms of an image are generated [19] by passing the image through Laplacian pyramid (LP) and the output of LP is passed through Directional Filter banks (DFB). NSCT decomposition algorithm is shown in Fig.1

The LP is responsible for the multiscale property of the transform. The LP is a decomposition of original image into a hierarchy of images. Each level corresponds to a different band of image frequencies. The band pass function enhances the image features such as edges that are vital for image interpretation. In LP, the image signal is low pass filtered, down sampled, interpolated and then the difference between input and interpolated output is obtained as \( L_0 \) output of LP. Output of the first down sampler is again filtered, down sampled, interpolated and the difference between the interpolator output and the stage input is obtained as \( L_1 \) output of LP and so on to get \( L_0, L_1, \) etc. outputs of LP. Due to the down samplers, each LP output is a scaled down output of the previous one and thus the multi scale property is obtained. Each output of LP is passed through DFB, providing multi directional output coefficients. The DFB contains fan filter banks and resampling operators. Outputs DFB are frequency of the input and hence multi directional decomposition of Contourlet Transform is realized. In image denoising applications using Contourlet Transforms, random noise is less likely to generate significant coefficients (compared to that using DWT) and hence application of simple thresholding schemes can remove noise from images decomposed using Contourlets

3. NSCT.
One important disadvantage of Contourlet Transform is its translation invariance (coefficients of delayed signal are not time shifted version of that of original signal)
Reason for this time variance is due to down samplers and up samplers in the LP and DFB. Because of the down sampling operation, the contourlet transform is nearly critically sampled, and is nearly irredundant. The redundancy of Lp is 4/3 and the DFB is critically sampled and the overall redundancy of the transform is 4/3. Due to this property of Contourlets, when it is used for denoising employing thresholding operation, pseudo Gibb's phenomenon and artifacts arises in the denoised output. Perfect reconstruction becomes a problem. this problem was solved by the introduction of Non-Sub sampled Contourlet Transform (NSCT) [20].

The NSCT is a fully shift invariant multiscale and multidirectional image transform, having fast implementation. The tight frame approach of filter design (which is considered for DWT) is not considered in the case of filter design in NSCT. Hence the filters in NSCT has better frequency seletivity which leads to better sub band decomposition having good frequency localization, and linear phase filter characteristics.

The shift invariance in NSCT is obtained by using 1) a non sub-sampled pyramidal (NSP) structure and 2) a non sub-sampled directional filter(NSDFB) in place of LP and DFB of Contourlet Transform. The NSP ensures multiscale property of NSCT, where the down sampling operation is not directly done, instead makes use of two channel non sub sampled 2D filter banks. As down sampling is not done, NSCT is a redundant transform and hence it can out perform the non-redundant transforms (eg. DWT, Contourlet, etc) in image denoising. The redundancy achieved by this expansion is \( J+1 \) where \( J \) denotes the number of decomposition levels. Subsequent to the first stage, the filter of the first stage is up sampled. Elimination of the down samplers is also done by a similar process in the NSDFB. At each scale, the NSCT could also attain \( 2^l \) directional decomposition. Due to the shift invariance of NSCT, images denoised using it have high visual quality and is devoid of artifacts which was seen when using Contourlet.
4. IMAGE DENOISING USING NSCT

The image denoising algorithm using NSCT is composed of the following steps:
1) Using NSCT, the noisy image is decomposed into J levels and obtain a low pass band L1 and number of high pass bands \( H_k^s \) for \( k=1,2,...,J \) and \( s = 1,2,...,D \) where \( k \) represents decomposition levels and \( s \) represents the decomposition orientations.
2) The high pass bands contain essential high frequency features such as edges, curves, etc. along with noise, as noise is generally of high frequency. The challenge is to remove noise alone keeping image features not disturbed. Several methods are available to remove noise in this manner. The quality of denoised image depends to a great extent on this second step. Complexity of each method also vary to a large extent. Generally, this step can be called as thresholding coefficients of high frequency bands of NSCT decomposed image in step 1.
3) The low pass band and the processed high pass bands are used to perform inverse NSCT and thereby to get the denoised image.

The step 2 has many choices. The two common and simple methods are
1) Hard thresholding and
2) Soft thresholding. In both cases, a threshold value \( T \) is estimated [21] which is a function of number of pixels in the sub band and its noise variance. Threshold is applied to each pixel in the image. For hard thresholding, if the pixel intensity is less than \( T \), all such pixels are forced to zero; or else it is retained as it is. In soft thresholding, if pixel intensity is less than \( T \), all such pixel values are forced to zero; or else it is shrunk towards zero. The hard thresholding can well retain the local characteristics of the image. However the method gives discontinuities in pixels to the image and therefore some image vision distortion. The soft thresholding method generally yields more visually pleasing image because of its continuity. At the same time, the edges of the soft thresholded image is made too vague by its continuous nature of thresholding. Final choice between the two thresholdings is made after conducting experiments of image denoising using both hard and soft and choosing the better.

Many adaptive thresholding techniques such as Baye's, Bivariate, etc also exist for thresholding noise from high frequency sub bands. In Baye's technique [22], threshold values are calculated from noise variance in the sub image, signal variance for each scale and orientation , exploiting exponential prior of NSCT coefficients at each scale. Soft thresholding is then done using calculated threshold value to remove noise. Generally, the Baye's method yields better denoising and more visually pleasing denoised image. However in high noise activity sub regions, denoising is not satisfactory. In Bivariate shrinkage [23], dependence between the coefficients and its parents in the previous level is captured, in addition to considering dependencies of coefficients in each sub band and Baye's estimation theory is used to get a shrinkage function. this method yields competitive results in addition to providing closed form shrinkage function and easy implementation.
In all the above thresholding methods, it is very important that the threshold calculated is most appropriate. If the threshold value is too small, noise removal is less. On the otherhand, if the threshold value is large, some signal features will be lost during thresholding. Also for more texture region, lower should be the threshold value and for more smooth regions, larger should be the threshold value. For higher noise levels, threshold value should be more for both texture and smooth regions. As most natural images contain both texture and smooth regions, setting an optimum threshold value becomes a non-trivial problem. One solution to this problem is to resort to classification of image pixels as noise related and edge related using Support Vector Machines (SVM) and then thresholding.

5. CLASSIFICATION OF PIXELS IN THE NOISY IMAGE USING SVM

Classification of NSCT coefficients contained in the high frequency sub bands is done using SVM [24] to separate the coefficients into noise related and edge related ones. For this purpose, binary map of NSCT coefficients of each sub image is obtained by determining threshold value T of the considered sub image by Otsu’s method and then converting the coefficients to either ones or zeros depending on the coefficients are larger than T or less than T respectively. From this binary map, spatial regularity in it is determined and is used further examine the role of valid NSCT coefficients whether it is isolated noise or part of spatial feature. The number of supporting binary values around a particular non-zero value of pixel of binary sub image under consideration is used to make the judgement. The support value is the total number of all valid NSCT coefficients which are spatially connected to the current pixel under consideration. The feature vectors $F_{k}^{s1}$ is obtained as the NSCT coefficients with maximum support value and $F_{k}^{s2}$ obtained as the NSCT coefficients with support value zero, which are randomly selected. These feature vectors $F_{k}^{s1}$ and $F_{k}^{s2}$ are used for training the objective $O_{k}^{s1}$ and $O_{k}^{s2}$ of all selected coefficients of sub image presently under consideration. SVM model can be obtained by training. By using the well trained SVM model, all high frequency NSCT coefficients of subimage are classified into noise related coefficients and edge related coefficients; edge related when actual output is one and noise related when output is zero. The classified noisy pixels are replaced by their median value while the non-noisy pixels are kept without any change.

The noisy pixels thus classified by SVM are thresholded to remove noise from the sub image. Any of the thresholding methods such as hard or soft or adaptive thresholding discussed previously can be adopted for this purpose. However they would give some what different final results of denoising, depending on the noise density present in the noisy image. The reconstruction of the denoised image is obtained by using inverse NSCT with the denoised NSCT high frequency sub bands and low frequency sub band as input for inverse transformation.
6. SIGNAL RECOVERY USING ORTHOGONAL MATCHING PURSUIT (OMP)

It is seen that with a small number of non-zero coefficients in a sparse or compressible [25] code, true signal can be accurately recovered by a few linear measurements which stisfy incoherence property. Let \( A \in \mathbb{R}^{n \times k} \) be an overcomplete dictionary of \( k \) bases and a signal \( b \in \mathbb{R}^{n} \) be represented as a sparse linear combination with respect to \( A \). For sparse signal \( x \in \mathbb{R}^{k} \) where \( b \) can be written as \( b = Ax \) and can be solved to extract true signal by using method such as OMP or basis pursuit or iterative hard thresholding.

The OMP algorithm [26] can be used after SVM classification for denoising and thus avoiding thresholding altogether and thereby eliminating the drawbacks of thresholding in image denoising. The initial signal estimation is chosen as random measurement matrix, usually the Gaussian random matrix. The OMP algorithm, at each step selects the column from the sparse data which is most strongly correlated with current residuals. Solution of a least square problem is done before each new selection. The selected column at each iteration is added to the set of columns already selected from the previous iterations and therefore the selected columns grow at each step. Before a built in stopping criterion is satisfied, the OMP algorithm will select all significant components. Also, the residuals after each step in the algorithm are orthogonal to all selected columns and so no column is selected more than once.

7. SIMULATION RESULTS

Random noise \( \sigma = 30 \) was added to gray (Lena) image. After NSCT decomposition of 3 levels and SVM classification, OMP method of denoising was applied to the decomposed high frequency sub bands. Thereafter, performing the inverse NSCT on the NSCT decomposed low frequency sub band and OMP denoised high frequency sub bands, final denoised image is obtained. The original, noisy and denoised images are shown in Fig.1. The PSNR, SSIM and output noise level (NL) of the denoised image was also estimated and is tabulated in Table 1. Experiment was repeated for \( \sigma = 20 \) and \( \sigma = 10 \) and results for \( \sigma = 10 \) are also tabulated in the Table 1. In addition, results of denoising experiments conducted on some other images (House, Pepper, Barbara, Cameraman) can also be seen in the Table 1. Further, the table also contains the results of experiments of NSCT with thresholding (Baye's) and hence the results of the two methods can be compared. It can be seen that the OMP method after SVM classification outperforms the threshold method after SVM classification in all respects.

For colour images, conversion from RGB colour space to YUV colour space was done first (as processing done directly in RGB domain resulted in colour artifacts due to high correlation among R, G and B components) and thereafter denoising each component was done using NSCT decomposition, SVM classification and OMP denoising. For comparison, NSCT decomposition, SVM classification and thresholding was also done. Representative results are shown in Fig.2 and Fig.3 and
Table 2. From the results it can be seen that the NSCT-SVM-OMP method gives much better denoising in terms of visual quality as well as PSNR, SSIm and NL estimation. All simulations were done using Matlab and relevant tool boxes.

8. CONCLUSION
Image denoising using NSCT decomposition, SVM classification and OMP denoising is found to give much better results, compared to NSCT, SVM, Thresholding method, considering image visual quality and denoising metrics such as PSNR, SSIM and Output Noise Level, both in the case of gray and colour images. On an average, there is an improvement of around 5 dB for the NSCT-SVM-OMP method compared to NSCT-SVM-Thresholding. As thresholding is totally discarded, the artifacts usually seen in denoising using thresholding is not met in the OMP method. Both these methods outperform compared to NSCT decomposition and thresholding. The PSNR of NSCT-SVM-OMP method and that of NSCT-SVM-Threshold method exceeds by about 7 dB and 2dB respectively, compared to NSCT-Thresholding method of denoising. However SVM takes around 45 sec and OMP takes around 25 sec of excess time for its execution.

Fig.2. Denoising of grey images

Fig.3. Denoising of colour image Barbara image
**Fig. 4.** Denoising results on Parrot image with Gaussian noise of $\sigma = 20$

**Table 1.** PSNR, SSIM, NL for gray images

<table>
<thead>
<tr>
<th>Image</th>
<th>Input noise level $\sigma = 10$</th>
<th>Input noise level $\sigma = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSCT, SVM, Threshold</td>
<td>NSCT, SVM, OMP</td>
</tr>
<tr>
<td></td>
<td>PSNR (dB)</td>
<td>SSIM</td>
</tr>
<tr>
<td>Lena</td>
<td>28.6</td>
<td>0.94</td>
</tr>
<tr>
<td>House</td>
<td>28.5</td>
<td>0.93</td>
</tr>
<tr>
<td>Pepper</td>
<td>27.6</td>
<td>0.91</td>
</tr>
<tr>
<td>Barbara</td>
<td>28.2</td>
<td>0.92</td>
</tr>
<tr>
<td>Cameraman</td>
<td>27.5</td>
<td>0.90</td>
</tr>
</tbody>
</table>

**Table 2.** PSNR, SSIM, NL for colour images

<table>
<thead>
<tr>
<th>Image</th>
<th>Input noise level $\sigma = 10$</th>
<th>Input noise level $\sigma = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSCT, SVM, Threshold</td>
<td>NSCT, SVM, OMP</td>
</tr>
<tr>
<td></td>
<td>PSNR (dB)</td>
<td>SSIM</td>
</tr>
<tr>
<td>Parrot</td>
<td>34.56</td>
<td>0.92</td>
</tr>
<tr>
<td>pepper</td>
<td>33.82</td>
<td>0.92</td>
</tr>
<tr>
<td>Mandrill</td>
<td>32.51</td>
<td>0.89</td>
</tr>
<tr>
<td>Barbara</td>
<td>31.23</td>
<td>0.82</td>
</tr>
</tbody>
</table>
REFERENCES


