Advances in Computational Research, ISSN: 0975–3273, Volume 2, Issue 1, 2010, pp-06-14

A review on natural image denoising using independent component analysis (ica) technique

Potnis Anjali*, Somkuwar Ajay and Sapre S.D.

Department Of Electronics and Communication Engineering, University Institute of Technology, Barkatullah University Bhopa, M.P., 462026, India

Department Of Electronics and Communication Engineering, Maulana Azad National Institute of Technology, Bhopal, MP, India- 462007, anju_pot@yahoo.com, asomkuwar@rediffmail.com

Abstract - Denoising of natural images is the fundamental and challenging research problem of Image processing. This problem appears to be very simple however that is not so when considered under practical situations, where the type of noise, amount of noise and the type of images all are variable parameters, and the single algorithm or approach can never be sufficient to achieve satisfactory results. Fourier transform method is localized in frequency domain where the Wavelet transform method is localized in both frequency and spatial domain but both the above methods are not data adaptive .Independent Component Analysis (ICA) is a higher order statistical tool for the analysis of multidimensional data with inherent data adaptiveness property. The noise is considered as Gaussian random variable and the image data is considered as non-Gaussian random variable. Specifically the Natural images are considered for research as they provide the basic knowledge for understanding and modeling of human vision system and development of computer vision systems. This paper reviews significant existing denoising methods based on Independent Component Analysis and concludes with the tabular Summary of denoising methods and their salient features / applications.

Key Words- ICA, ISA, MICA, MFT-ICA

Introduction

Denoising of images is a challenging and extremely relevant research problem as the type of images, type of noise and amount of noise all are variable in the practical situations. All denoising methods are ultimately a type of low pass filters. Filtering of the images can be done in the spatial domain or in the transform domain. Filtering in the transform domain is more efficient and introduces fewer artifacts. The focus of recent research has been on the higher order statistical methods and the nonlinear transform domain filtering .There are numerous methods and approaches for denoising of images, however only the prominent and relevant methods based on ICA technique are discussed here. First time the ICA technique was presented by [1].The easy and popular low pass filter method smoothens out the image so that the edges are not prominently visible .Denoising techniques based on Fourier transform method is localized in frequency domain and the wavelet transform method is localized in both frequency and spatial domain but both the methods are not data adaptive, however if the filtering approach is data adaptive it comes out with promising results, and that is the inherent property of ICA techniques. Data adaptiveness plays an important role in image denoising process because denoising of images is also dependent on the image (type) which is to be denoised. For all the methods discussed here, images are corrupted by additive Gaussian White noise, which is an appropriate representation of noisy images acquired by various image capturing and scanning methods. Gaussian white noise also includes all the frequency components in the visible light range .The multiplicative noise reduction method is also discussed to provide idea as to how the ICA can be applied to solve such complex noise removal problem. The amount of noise added in the image is quantitatively measured by the variance of the observed data .Variance of the noisy image is decided by assumption in a particular range or is calculated by different variance calculation methods and the optimization methods, the method given by [2, 3] is very popular, which uses the wavelet HH band coefficient for the noise variance calculation. Performance of denoising algorithms strongly depends on the amount of noise present and the sample size (image patch /image window) considered for processing in single step. If the image patch considered is big, the computational complexity will be increased and it is also difficult to implement the algorithm, however if the image patch considered is small, computational complexity will be less and the algorithm is easier to implement even though the overall computational burden increases. The patch size window size also affects the amount of artifacts introduced in the processed image. The visual quality of the processed image is more important although most of the algorithms also compare mean square error (MSE) and peak signal to noise ratio (PSNR) in dB of the denoised and noisy images. The following methods are discussed to understand the approaches to denoised the Natural images: - 1.Principal Component Analysis (PCA) and Adaptive PCA 2. Sparse code shrinkage and Improved Sparse code shrinkage methods 3.

ICA and orthogonal ICA mixture model 4.Subspace ICA 5. Topographic ICA 6. Wavelet ICA 7. Multi resolution Fourier transform (MFT) MFT-ICA 8. ICA for multiplicative noise. In addition to above orientation –adaptive Gaussian scale mixture Model for image denoising in the ICA domain , in wavelet domain , Sparse code shrinkage based on the normal inverse Gaussian Density Model, Gradient Adaptive methods, Denoising source separation and Hebbian learning and Bubblean coding etc. are also used for denoising however due to the constraints of space only important methods and their salient points have been discussed. Efforts have been made to avoid the repetition of basic theories, definitions, standard notations and other common details while discussing the various methods.

1.PCA and Adaptive PCA :-Principal component analysis(PCA) is a linear data adaptive type of transform technique which is also known as Hotelling transform .In this transform linear subspace fit for the given data is to be calculated i.e. optimizing for the minimum mean square distance between the data points and their projection on the subspace . If it is assumed that the dimension of the data is *m* and the dimension of the transform subspace is *n*, then it is required to find the orthogonal vector w_i ,where (i=1,2,3…. n) .The principal components are found by optimizing the direction of maximum data variance under the constraint of the orthogonality to previously found direction , mathematically it is given as

$$w_i = \arg\max_{w_i} E\{(w_i^T x)^2\} \qquad (1.1)$$

The maximization is performed under the constraints

$$w_i^T w_j = 0 \text{ , } j < i \qquad (1.2)$$

$$||w_i|| = 1 \qquad (1.3)$$

The principal component vectors are found by calculating the Eigen vector of the correlation matrix $C_x = E\{x\, x^t\}$ arranged in the decreasing order of the corresponding Eigen values ,where the Eigen values are the variance in the direction of Eigen vector . Eigen vectors λ_i are represented as

$$\lambda_i = E\{(w_i^T x)^2\} \qquad (1.4)$$

If the calculated principal component vectors w_i are arranged in to the matrix W such that w_i^T is the i^{th} row of W then we have the PCA transformed matrix W. This transform have many useful applications like calculating the uncorrelated vectors, principal component, the dimensionality reduction, preprocessing for the ICA method etc. The adaptive PCA technique [3] is used for generating 2-D basis sets, which have vectors lined up along the edges and not across them .The selection of 2-D locally adaptive basis sets is the main contribution of the paper [3]. 2-D basis functions are generated by the principal component method applied on image patches .The largest eigenvector is in the direction of the local image edge. The decomposition of the image is by locally adaptive principal component analysis technique, thresholding the coefficient and then reconstructing the denoised image by the reverse transform. The local basis Principal components are recalculated at different locations. The process is started with the estimation of noise variance present in the image, using the most popular expression given by [3] is used as a starting variance .The noisy image is represented as $Y = X + n$ (1.5)

Where Y is the noisy image observations, X is the denoised image or the source image and n is the noise present in the image. Noise considered is the additive white Gaussian noise of zero mean and variance σ. The image is partitioned into overlapping patches, each patch contains trained region, denoised region and overlap region as shown in figure.1 The principal components of Y are locally adaptive basis function , Y is the observed image patch matrix consist of locally collected noisy image patch, the size of the image patch is M pixels i.e. i= 1,2,……M. The number of column in Y matrix is MXM. Total dimension of Y matrix is $M^2 X N$. Where N is the number of patches considered .Next the variance of the Principal component coefficient is calculated using maximum likelihood estimator as given by [3]. This basis decomposition is more efficient than the wavelet decomposition [3]. This is more specific for the high structured regions or for the images like natural images. The algorithm is inherently shift invariant in nature [3].If the denoising region is too large, blocking artifacts in the denoised image is quite evident even the PSNR is improved. To average out the blocking artifacts between different denoised regions, add an overlap region and average it. The paper [3] compares the proposed method with three different methods i.e. wavelet Hidden Marko

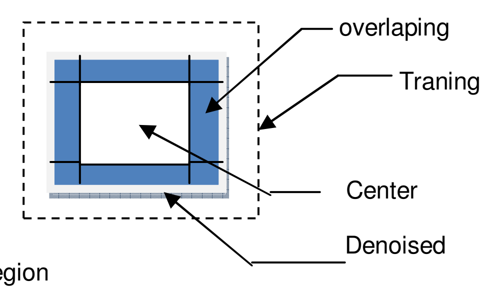

Fig:-1 Denoising is done patch wise using the PC base algorithm. The overlapping area is labeled as shown to reduce the artifacts

Tree model method of [3], spatially adaptive image denoising method and SI–Adaptive shrinkage method and shows the improvement in the PSNR. Table 1 shows PSNR achieved by this method for

Gaussian noise variance σ 50 , 25and 15 as given in [3] for the four test images .

Table1- Comparison of PSNR (dB) for various noise variance

Images	Adaptive Principal Component method (PSNR in dB) for noise variance σ 50	25	15
Lena	28.00	31.26	33.60
Boat	26.36	29.62	32.25
Barbra	26.21	29.91	32.63
Ring	27.18	32.82	36.32

2. Sparse code shrinkage and Improved Sparse code shrinkage method: - Sparse code shrinkage is a type of redundancy reduction and is equivalent to Independent component analysis method for image data [4].It is a data adaptive noise removing technique where shrinkage parameter is calculated from the data itself. Sparse coding is a linear transform in which each of the components of the transform is rarely significantly active. The representation of the data is on the basis of estimation of maximum likelihood of the observed data[4]. In most of the cases it is assumed that number of transform component is equal to number of observed nongaussian variables for the ease of calculation .Maximum likelihood estimation is carried out by considering the different parameters , and the shrinkage strategy is key to denoise the images . Sparse distribution has heavy tail and peak at zero which is also called the super gaussianity. Shrinkage is nothing but thresholding method of independent components for denoising of images. In general two type of component thresholding is possible Soft thresholding and hard thresholding. As only few of the components are significantly active at a time and the components with the small absolute values are purely noisy so by soft thresholding the unwanted components are made zero or replaced by the approximately calculated threshold value from the parameters considered for the modeling of the image data.

Let the observed random variable are gives as $x = (x_1,x_2,x_3,\ldots\ldots.x_n)^T$ that is input data or noisy image window and $s= (s_1,s_2,s_3,\ldots..s_n)^T$ the vector of the linearly transformed component variables .If W is a n X n transformation matrix then the linear representation is given as $s=Wx$,where it is assumed that the number of transformed components are equal to number of observed components .

It is assumed that the distributions of sparse components are nongaussian and the noise is having the Gaussian distribution. Let the nongaussian random variable is denoted by s and the Gaussian noise is represented by f with zero mean and some variance value ,so the noisy image data is represented by

$$Y= s + f \qquad (2.1)$$

In the above equation s is the denoised image which is the required solution of this problem .If the source s is represented by probability density function p ,which gives the negative likelihood of it as $f = -log(p)$ for the given data density as given in [4,5],then the maximum likelihood estimation method gives the estimation for s' as

$$s'=arg\ min\ _u 0.05\sigma^{-2}(y-u)2+f(u) \qquad (2.2)$$

Here f is a convex function and completely differentiable gives the solution as $s' = g(y)$, where the function g is given as[6]

$$g^{-1}(u)=u+\sigma^2 f'(u) \qquad (2.3)$$

This nonlinear estimator is the key to this method. The probability density of the s is required to be modeled to use the above nonlinear estimator .Two important parameterization methods given in [6] are considered for most of the densities encountered in the natural image. The density family suitable for the suppergaussian density is given as

$$p(s) = C\ exp\ (\ -as^2/2 - b|\ s|\ ) \qquad (2.4)$$

Where a, b are greater than zero and to be estimated as given in [6], C is the proportionality constant or the scaling constant .For the above density the nonlinearity is given as

$$g(u) = 1\ /\ (1+\ \sigma^2 a\ \ (sign(u)max(\ 0,\ \ |u|\ -\ b\ \sigma^2\ \ ))) \qquad (2.5)$$

Another density is more sparse then Laplace density and is more complex then the above given in the [6].

The method can be summarized as - first start with noise free data as training set of x , using some sparse coding method calculate the orthogonalize matrix W so that the components s_i in $s=Wx$ have as sparse distribution as possible . Estimate a density model $p_i\ (s_i)$ for each sparse component using equation 2.4 ,now for each noisy observation x' of x the sparse component is represented as $y=Wx'$, apply the shrinkage nonlinearity $g_i(.)$ as given by equation 2.5 denote the obtained components by $s'= g_i(y_i)$.Invert the relation $s =Wx$ which will give the denoised image . Here the row vector of W is called the ICA filter.

The improved spares code shrinkage method for natural image denoising [7] suggest the steady –

transition threshold and proposes a compensation operation to avoid the losing details in the filtering process .The new sliding and nonsquare window method greatly reduces the computational complexity . Fixed – point Fast ICA algorithm proposed by Hyvarinen is used as for the ICA transform .The components are multiplied by a compensation factor ß(1.05 to 1.20) as given in [7] before inverting the transformation .The method can be summarized as –take the noise free image set of the same class of images which are to be denoised and estimate the sparse coding transform W by estimating the ICA transform matrix and orthogonalized it, than observer the noisy image , transform it by sparse coding transform W .Apply the estimated nonlinearity g to each component ,multiply each component by the compensation factor ß ,perform the inverse transformation to get the denoised image .

Performance of the improved sparse code shrinkage method for the random images as given in [7] , shows average value of improvement in PSNR by this method for different window size is given in table-2

Table 2- Average PSNR for different window size and noise level

Noise Level	Noisy Image	8X8	4X8	4X4	6X6
0.1	32.27	37.10	37.17	37.00	37.03
0.2	26.25	33.16	33.22	33.06	33.17
0.3	22.73	31.13	31.13	30.84	31.08
0.4	20.21	29.63	29.28	29.28	29.63
0.5	18.26	28.62	28.57	28.10	28.51

3. ICA and orthogonal ICA mixture model:-PCA is purely second order representation and processing of the data but it is observed that more information in the natural images is in the higher order of the data .Therefore the ICA method gives better representation of the natural images than PCA. In ICA transform data is represented as the statistically independent random variables .The image data is considered as the nongaussian random variable and the noise is additive white Gaussian noise. This assumption makes this problem suitable for solving efficiently by linear ICA method[8,9] .The observed random variable vector(noisy image) is represented by X and modeled as

$$X= As+ \eta \qquad (3.1)$$

s_i is statistically independent random vector , A is mixing matrix and η is representing the noise which is independent of s_i .The representation of the data without noise is nothing but the reconstructed denoised image $X' = As'$.

The dimension of observed random vector X and the required vector s gives three different conditions of ICA (1) over complete case when dimension of observed random vector X is greater than the required source vectors s_i (2) Complete case when both are same (3) undetermined case when dimension of observed random vector X is less than the required sources vectors.

The goal of this method is to find the optimized mixing matrix A and the independent sources s from the observed random vector X. For the case when the dimension of the observed vector and the source vector are same then the ICA model is considered as the complete noiseless model and in this case the solution is as given in [8] for the uncorrelated one dimension projections which are maximally non- Gaussian. The processing starts with whitening of the data and finding the orthogonal non–normal projection .Denoising is a two step process (1) noisy image patches are observed, which are randomly selected and stored in an observed matrix X row wise, it is then linearly transformed to get $g=VX$,such that the components of g are having unit variance and are uncorrelated i.e. $I=E\{ g .g^{-1}\}$.This can be achieved by setting $V =C_x^{-1/2}$ where $C_x=E\{X.X^{-1}\}$is the correlation matrix of the data . The second step(2) is to find the orthogonal matrix B, which gives the maximization of the cost function and so X is calculated as $X= Bg$.

In ICA mixture model based image denoising, the data density in each class is given by parametric nonlinear function that fits to the non-Gaussian structure of the data [8,9]. The modeling of image is based on different data density function .Data in each class is generated by multivariate Gaussian sources known as Gaussian mixture model [10] which uses Gaussian kernel The fast ICA with the hyperbolic tangent nonlinearity is used for the ICA estimation . The ICA transform can be orthogonalzed to achieve the super gaussianity better then PCA .

4. Independent Subspace Analysis (ISA):- It is the nonlinear version of ICA, in this analysis the components are grouped and the independence is found in such a way that the groups are independent from each other but the components inside the group are dependent[11] .It is observed that with in the group a component follows spherical symmetrical sparse density given as :

$$P(s) = \prod_k \exp\left(-f\left(\textstyle\sum_{i\in G} s_i^2\right)\right) \qquad (4.1)$$

Where $k=[\ 1,2,\ldots\ldots..G]$,is number of subspace assumed for the analysis .So it is clear that the subspaces are independent of each other ,and its individual probability density product can be expressed as above equation (4.1) of they are independent .The algorithm used for the learning the image basis is maximum likely hood estimation by Gradient assent method. The Figure no.2 shows that in a group the features have same orientation ,spatial frequency ,and location .This model is used for the denoising of images which have localized type of noise [11].

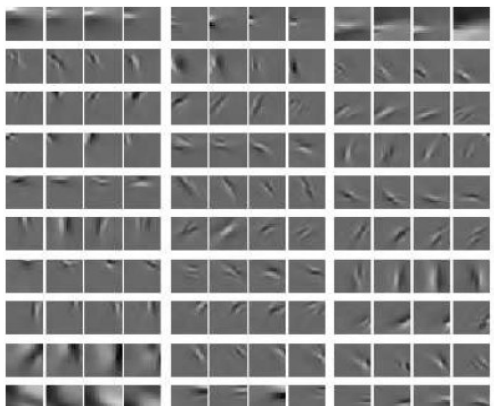

Fig. 2- image bases of independent subspace analysis

5. Topographic Independent Analysis (TICA):- In real data samples it is observed that the groups of components are not as independent as required for the success of ICA .It is observed that pair of components are independent and dependence exist between related components .This is modeled by arranging the components on a grid i.e. 1D or 2D and assumed that the near by components on the grid have energy correlation [12] ,but distant components are independent. The probability density function is given as

$$P(s) = \prod_i \exp\left(-f\left(\sum_{j \in N_i} s_j^2\right)\right) \quad (5.1)$$

Where i and j both have values from 1 to n .

This model is more or less similar to ISA with only difference that this has overlapping group of components .In this process the near by features have

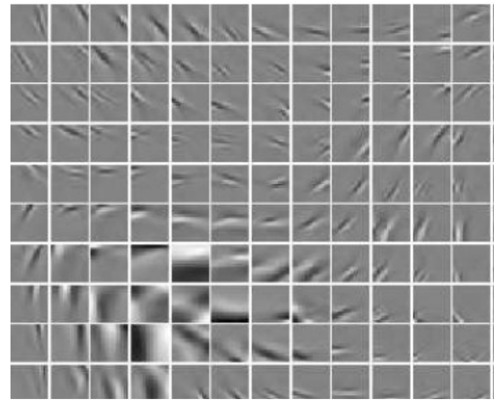

Fig. 3- image bases of topographic ICA

same orientation, spatial frequency ,and location. This extension is very efficient in case of natural image processing . The bases of topographic ICA are given in figure .3 which shows near by features have same spatial frequency and location.

6. Wavelet-ICA: - Here in this paper an adaptive technique is proposed for the denoising of natural images using soft thresholding method in the wavelet domain [13]. The optimum value of threshold is determined by ICA technique. The negentropy of the processed image which is the measure of nongaussianity is considered for the optimum selection of threshold value. Wavelet transform converts the image in the wavelet domain where noise is more efficiently separable. The denoising is done by the thresholding method, the value of thresholding parameter is very critical to decide. In the proposed method the value of threshold is selected by the measure of non gaussianity of the processed image. The measure of nongaussianity of the image is found out by the ICA technique. As it is assumed that the image has nongaussian distribution and the high value of nongaussianity can be considered as image with less noise added, this makes the proposed method more efficient and gives better results compared to other thresholding techniques.

Thresholding of wavelet coefficient is done to denoise the given image in efficient way [13] than the conventional filtering method. Hard thresholding is simple to implement but not efficient and results in more artifacts compared to soft thresholding. For hard thresholding Wavelet coefficients are to be threshold by the criteria that if the value of wavelet coefficient is more than the assumed threshold value then there will be change in that coefficient and the coefficient value is replaced by the threshold value and if the value of the coefficient is below the threshold value then that coefficient will be replaced by zero as in [14].Soft thresholding method performs better as thresholding is not fixed type and results in less artifacts. Here value of wavelet coefficient is compared with the thresholding value and if it is less than the thresholding value then the value of that coefficient is replaced by zero if it is greater than the threshold value, coefficient is replaced by subtraction of the thresholding value with the value of wavelet coefficient [14]. The value of starting threshold value is critical to decide and the over all performance of the method depends on it. This initial threshold value η is given in terms of noise variance and the size of image to be considered, to calculate the wavelet transform.

$$\eta = \sigma \sqrt{(2 \log M)} \qquad (6.1)$$

Where σ^2 is the noise variance and M is the block size to calculate the wavelet transform. The value of noise variance for the given noisy image is calculated by median estimator [13,14] using the following equation

$$\sigma^2 = \text{Median}\ (\, y_{ij}\,)/\ 0.6745 \qquad (6.2)$$

Where $y_{ij} \in$ subband HH_1 of the test image.

In the proposed method ICA is used to select the optimum value of threshold .It is assumed that the image is degraded by the Gaussian noise 'n' and the observed image e is a mixture of noise and the image i represented as

$$e = i + n \qquad (6.3)$$

The noise is assumed as Gaussian noise having zero mean and variance σ^2 which is the basic requirement of ICA technique . For the nongaussianity measure negentropy is considered here and is represented as

$$j(y) = 1/12(\, E(y^3)^2\ ) + 1/48\ (E\{y4\} - 3(E\{y^2\})^2)^2 \qquad (6.4)$$

as given in [13] the high value of negentropy gives the image with minimum Gaussian noise present in it.

The proposed method does not require the clean image .Noise variance is calculated form the noisy image which can be considered as foremost advantage of the proposed method. The value of noise variance is adapted to image under consideration. The graph of Figure no. 4 as given in [13] shows variation in the normalize negentropy with respect to the wavelet thresholding values.

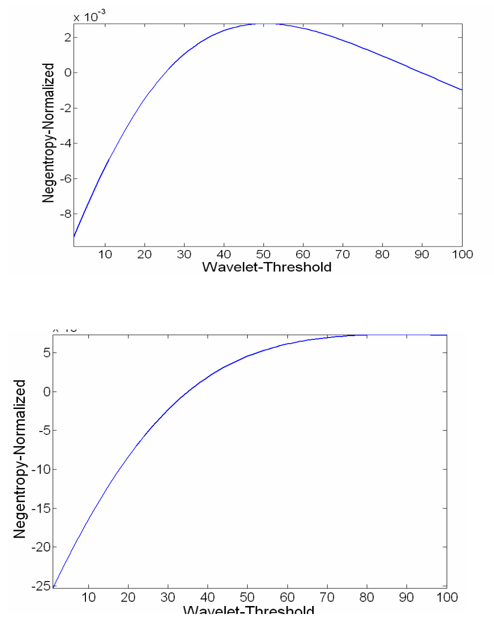

Fig. 4- wavelet thresholding Vs the negentropy and normalized negentropy

7. MFT-ICA: - Multi-resolution Fourier transform (MFT) ICA is the combination of two methods to exploit the benefits of both the approaches i.e. computational efficiency of the MFT and data adaptiveness of ICA .Directional information is an important component of natural images. Directional basis analysis is used for observing the statistical structure of image.MFT is based on multidirectional selective filter .This analysis is coupled with the ICA which adaptively decomposes an image into a set of directional bases [15] . The noisy image *v'*, noise free image v and the noise η are related as

$$v' = v + \eta \qquad (7.1)$$

The ICA decomposed signal has following relation as given in [15]

$$P(A \cap B)=P(A).P(B) \qquad (7.2)$$

Where *A* and *B* are the independent components of the signal as given in [15].The estimation of the ICA component is for the search of uncorrelated directions in which the components have a sparse distribution .In the first step perform the ICA on each subband of the MFT with the limited number of components specified .The rejected components are nothing but the noisy components .The second step is to apply a 2D Gaussian filter obtained from the Fourier spectrum of the bases found in the above step .As most of the components are localized in orientation as well as frequency the resultant inverse reproduction is nothing but the noise free image .The effectiveness of the algorithm is compared with the recent wavelet transform and is represented for various noise levels as shown in figure .5 as given in[15] . The test image Lena and Jaguar are considered in the paper [15]and the results are compared for PSNR value of denoised image patch for various noise level.

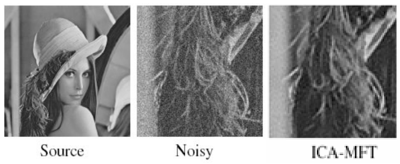

Fig. 5- (a) results of ICAMFT for Lena

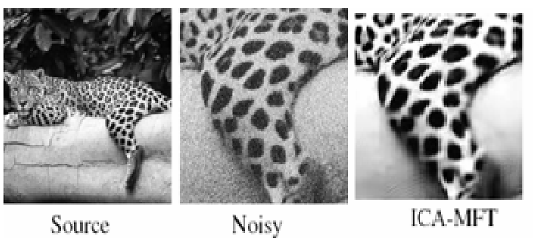

Fig. 5- (b) results of ICAMFT for Jaguar images

8. Multiplicative Noise Reduction: - In this paper the modified ICA is used for multiplicative noise removal [16]. Images contaminated by the multiplicative noise are very complex to denoised as the multiplicative noise is not as independent as the normal white Gaussian noise and changes the statistic of the image data under consideration in very different way. This requires modification in the standard ICA model which is based on simple linear and nonlinear mixture of independent sources . The Multiplicative ICA (MICA) model is discussed in [16] to denoise the natural images. The second and higher order statistics are both used together to solve this complex problem. The linear mixture of independent sources contaminated by multiplicative noise is represented by

$$Y_i=v_i x_i \quad \text{for } i= 1,2,\ldots\ldots\ldots n \qquad (8.1)$$

and

$$x=As \qquad (8.2)$$

Where $Y=[Y_1,Y_2,\ldots\ldots\ldots Y_n]^T$ is noisy image patches represented in the vectors and $s=[s_1,s_2,\ldots\ldots s_n]^T$ is the vector of n independent sources, the component of multiplicative noise $v=[v_1,v_2,\ldots\ldots\ldots v_n]^T$, the noise free data is given by vector x and A is the mixing matrix . The inverse of the mixing matrix is called the unmixing matrix and that is the solution of this type of problem .This solution is achieved in two steps first is whitening which is very essential step for the dimensionality reduction and at the same time it improves the convergence rate of this problem. The goal of ICA is to find a linear transformation matrix B of the data such that the output is as independent as possible. The matrix B is represented as

$$B=UW \qquad (8.3)$$

Where W is the whitened matrix and $\boldsymbol{U}$ is the unmixing matrix (unitary matrix). The problem is solved by the second, third and fourth order statistical structure of the image. Covariance, third and fourth order cumulants are considered as contrast function, the ICA problem is solved by the linear and nonlinear minimization with the quasi – Newton method [16]. Here the method uses the third order moment and is called Third order MICA(TMICA)and the method based on fourth order statistic is known as FMICA . FMICA is more complex then TMICA. It is observed in this paper that TMICA is giving more accurate result than FMICA because the calculation of estimation function is more accurate in case of TMICA . Results of TMICA and FMICA are compared with the standard Fast ICA method as given in [16]. The Figure no:6 shows clear improvement in the result compared to FAST-ICA for the wide range of multiplicative noise as given in [16].

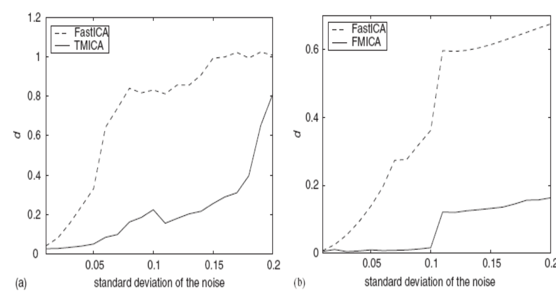

Fig. 6- variation in the parameter d with respect to standard deviation of the noise.

The parameter used for the comparison is d and the standard deviation of the noise present in the data .The parameter d is defined as the minimum distance from the global transformation to the identity or any permuted sign switched version of the identity as given in[16] .

9. Conclusion

ICA is a higher order statistical method suitable for multidimensional and multivariate data analysis and processing. Denoising of natural images by ICA methods are strongly data adaptive as denoising processes do not require the noise free image in general, which is the uniqueness of this method and makes it more efficient than wavelet and other transform domain filtering methods.PCA is a second order blind source separation method based on the covariance of the data and mainly used for the dimensional reduction and whitening of the data for the further processing. PCA method discussed here for denoising also suggests the improvement in the simple PCA method for the efficient natural image denoising. Higher order statistical methods for the signal processing are more efficient and accurate but computational complexity and time consuming . Sparse code shrinkage(SCS) method is a popular method for image denoising compared to all other methods discussed here .Efficiency of SCS method depends on the accuracy of sparse coding and the shrinkage parameters assumed for the dimension reduction. .SCS method based on the maximum likelihood estimation is most popular and efficiently implemented [5,6,7]. ICA methods discussed here have considered the Gaussian mixture model which can be linear or non linear ICA depending on the initial assumption applying for denoising. It is found that Fast ICA algorithm given by [17] is used by the most of the methods for the independent component analysis. It is also observed one of the following :- that mean square error or PSNR or the measurers of nongaussianity like negentropy or normalized kurtosis values of the denoised image are considered for comparison of the performance of the methods, at the same time it is realized that the visual inspection also plays important role . Type of data set, normalization of local variance, window size of image, the amount of noise present and type of noise in the image affects the performance of the algorithm and are the major constraints of the process. As most of the methods consider the white Gaussian noise introduced in the image for the experiment purpose but the non Gaussian noise can also be eliminated by making some modification in the basic processes. In the limiting case the non Gaussian variable can be considered as the Gaussian variable by central limit theorem. A table is prepared to conclude the study and gives the salient feature of each method and the applicability of the method. It is clear from the table that each method have different application and constraint in the practical situation .To over come the constraints and to develop a unified and global method for image denoising is still the real challenge in front of signal processing and image processing community.

References

[1] Comon P. (1994) *Signal Processing* 36:287-314.

[2] Donoho D. and Johnstone I. (1994) *Biometrika,* 81, 425– 455.

[3] Darian Muresan D. and Thomas W. Parks *"Adaptive Principal Components and Image denoising",ICIP*-2003.

[4] Hyvaerinen A. (1999) Neurocomputing.

[5] Hyvaerinen A. (1999 *Neural Computation*, 11, 1739-1768.

[6] Hyvaerinen A., Hoyer P. and Oja E. (1999) *Sparse code shrinkage Denoising by nonlinear maximum likelihood estimation, in Proc. Advances in Neural Information Processing Systems*, 1999.

[7] Ying yu ,Jain Yang ,DanXu

[8] Lee T.-W. and Lewicki M. S. (2000) *Proc. Int. Workshop ICA,* 239–244.

[9] Jose A. Guerrero-Col´on, Eero P. Simoncelli, Javier Portille (2008) *Proceeding International conference on image processing ,San Diego, CA.*

[10] Lee T.-W., Lewicki M. S. and Sejnowski T. J. (1999) *Advances in Neural Information Processing Systems*. Cambridge, MA, MIT Press, 508–514 .

[11] Hyvaerinen A., Jarmo Hurri ,Patrik O.Hoyer (2009) "*Natural Image statistics – A probabilistic approach to early computational vision", Springer.*

[12] Aapo Hyvaerinen, Patrik O. Hoyer, and Mika Inki. (2001) *Neural Computation* 13(7), 1527-1558.

[13] Isar Hussian and Hujun. (2008) ISCCSP, Malta.

[14] Donoho D.L. (1995) *IEEE transactions on Information theory,* 41 (3), 613- 627.

[15] Heechan Park, Graham R. Martin, Zhen Yao (2007) *Image denoising with Directional Bases in ICIP*-2007 , I-301 to I-304.

[16] Blancoa D., Mulgrewb B., McLaughlinb S., Ruiza D.P., Carriona M.C. (2006) *Neurocomputing* 69, 1435–1441.

[17] Aapo Hyvaerinen and Oja E. (1997) *Neural Computing.*, 9,1483–1492.

SN	Methods	Silent features /Applications
1	PCA and Adaptive PCA	• Selects 2D locally adaptive basis sets, thereby reducing high frequency components and improving denoising algorithm. • Performs best with high frequency content or textures.
2	Sparse Code Shrinkage and Improved sparse code shrinkage	• Reduces computational complexity by using sliding window and nonsquare window technique. • The method is most suitable for Sharp object boundaries and small texture areas.
3	ICA and orthogonal ICA mixture model	• Orthogonal ICA gives better results than simple ICA. • The method is most suitable for Sharp object boundaries and small texture areas.
4	Subspace ICA	• It is a non linear projective subspace analysis method. • Basis vectors associated with the single subspace have almost same frequency and orientation.
5	Topographic ICA	• It is the Generalization of the subspace ICA model. • More reliable compared to simple ICA. • Performs well for the natural images.
6	Wavelet ICA	• Combines Wavelet method with ICA • Nongaussianity is used as measure for the adaptive selection of the wavelet threshold.

		• The method is very much suitable for the higher noise level.
7	MFT-ICA	• Computationally efficient then ICA. • Combines the Advantages of MFT and ICA methods • Performs better for texture.
8	ICA for multiplicative Noise	• Second and higher order statistics are combined for processing the image. • The method uses the covariance, third and fourth order cumulants as the cost function for the ICA .

Mrs Anjali Potnis- Mrs. Anjali Potnis received her Bachelors degree in Electronics and Communication Engineering in the year 1994 and Masters degree in Digital Communication in the year 2002 from MANIT ,Bhopal(M.P)India. Currently she is pursing her PhD from the same Institute. Her research interests are Signal processing, Image processing and Digital Communication. Presently she is working as senior faculty in the Department of Electronics and Communication Engineering , Barkatullah University Institute of Technology ,Barkatullah University , Bhopal (M.P.)India.

Dr Ajay Somkuwar- Dr. Ajay Somkuwar has done PhD from IIT ,Delhi . Presently he is working as Assistant Professor in the Department of Electronics and Communication Engineering ,MANIT, Bhopal (M.P.) He have more then 15 years of teaching experience .His areas of interest are Signal processing , Image processing and communication .

Dr S.D. Sapre- He is working as Professor in the Department of Electronics and Communication Engineering ,MANIT, Bhopal (M.P.) He has more then 22 years of teaching experience .His areas of interest are Signal processing, Image processing, Information Theory and Coding . He is the Author of Book Communication systems published by Tata McGraw Hill's India.