International Journal of Engineering Researches and Management Studies

ADAPTIVE VIDEO DENOISING USING DISCRETE WAVELET TRANSFORM

Shylaja.S.L*[1], Vinayadatta V Kohir[2]
*[1]Research Scholar Visvesvaraya Technological University,India
[2]Poojya Doddappa Appa College of Engineering, Gulbarga-585102India

Keywords: Video Denoising, Motion Detection, DWT, Adaptive Spatial Filter, Recursive Temporal Filter.

ABSTRACT
This paper proposes a spatial video de-noising method based on adaptive filtering and temporal filtering as it requires less computation time and suitable for real-time applications. video de-noising rely on spatial adaptive temporal filtering is offered. Initially in the adaptive spatial filtering, DWT (discrete wavelet transfer) is applied on all frame. These co-efficient are used to alter trained co-efficients stored in knowledge base by means of adaptive spatial filter. On the further hand, the motion estimation is performed on every frame and the estimated results are used to guide the temporal filtering on noisy frame. Later on processing, two de-noised frames are obtained, one obtained from recursive temporal filtering and another obtained from adaptive spatial temporal filtering. Finally, by weighting between two de-noised frames, a satisfactory result can be obtained. Later all the de-noised frames are combined to get the de-noised video. In this work, we introduce a trade-off between the quality of denoised video and time required for denoising. The results show that proposed method is suitable for real time applications while giving high quality videos at relatively low noise level.

INTRODUCTION
Over the past few decades, advance of signal processing and networking technologies, applications that utilize digital video ,,have been increased. Example includes video conferencing, digital TV broadcasting and multimedia services. Several video processing technologies have been developed to tackle dissimilar problems. one of the method is video de-noising, from where the noise is removed from digital video.

There are many ways noise could get into a digital video, typical ways are through acquisition system and transmission over networks. Noise in a digital video is undesirable not only because it scarifies the perceptual quality but decreases the compression efficiency of a predictive video. Thus, video denoising is important as it could increase the perceptual quality.

Video signals are considered as a sequence of two dimensional images, which are projected from a dynamic three dimensional(3d) scene onto the image plane of a video camera. Luminance and chrominance are two attributes that describe the color sensation in a video sequence of a human being. Luminance refers to the perceived brightness of the light, while chrominance corresponds to the color tone of the light. Numerous still images and video denoising algorithms have been developed to enhance the quality of the signals over the few years[1]. Many of the algorithms are always based on probability theory, partial differential equations, statistics , linear, non -linear filtering, spectral and analysis. However, image de-noising can be extended to a video by applying it to each video frame independently to other. Depending on many signal-processing problems many algorithms have been proposed mainly for image de-noising [2]. A human observer cannot resolve fine details within any image due to the presence of speckling.

Video denoising is normally done with some linear or non-linear operation on a set of neighboring pixels and the correlation between those pixels available in spatio-temporal sense. The good video de-noising can be easily achieved by exploiting information from both past and future frames. But this leads to an additional delay of at least one frame which is undesirable in some real time applications. Only for this reason, many of the algorithms exploit information from usually the current present frame and one or two previous frames. Basically In image frame de-noising algorithms are focused to find the best compromise between noise removal and preservation of important denoised image frames. Here each frame is independently processed. That is why for optimal filter performance the spatio-temporal properties of the processed noisy image frames are taken into consideration.

An accurate modeling of noise is necessary in order to estimate noise-free spatio-temporal sequence structures. To differentiate in between the noise and the noise-free spatio-temporal correlations in the image frames the information concerning the noise and the noisy input frames are summoned together. In this way, the spatio-temporal structures can be estimated

Video de-noising algorithms may be roughly classified based on two different criteria: whether they are implemented in the spatial domain or temporal domain and whether motion information is directly implemented [09]. The Motion compensation may be used to avoid ghosting artefacts when blending together pixels from several frames.

In this paper, spatial filtering of individual frames is done and is combined with temporal filtering. The denoising artifacts and residual noise differ from frame to frame which degrades the visual quality. Hence temporal filtering is combined with adaptive spatial domain denoising. The rest of this paper follows explanation of all this approach in detail

EXISTING SYSTEM

Image and video denoising has been studied for decades. As it is beyond the scope of this paper to provide a thorough review, we shall focus on reviewing the present work closest to ours. Image sparsity can be manifest itself in the different forms. When images are decomposed into sub-bands,the sparsity leads to image coring algorithms on the wavelets co-efficients [4, 5]: large-magnitude coefficients that more likely correspond to true image signal should be retained, whereas small-magnitude coefficients that more likely correspond to noise should be shrunk. When the prior of natural images is incorporated in denoising [12], image sparsity is reflected by the heavy-tailed robust potential functions associated with band-pass filters: pixels in a neighborhood are encouraged to be more similar, but sometimes occasional dissimilarity is allowed. Other denoising techniques such as PDE’s [11] and region-based denoising [14] also implicitly formulate sparsity in their representation

Roth and Black [3] used a model, called the Fields of Experts (FoE) model, which is a generic Markov random field model of image priors over extended neighborhoods. The goal of the Field of Experts model is to develop a framework for learning rich, generic prior models of natural images (or any class of images). In contrast to example-based approaches, this model develops a parametric representation that uses examples for training, but does not rely on examples as part of the representation. Such a parametric model has advantages over example-based methods in that it generalizes better beyond the training data and allows for more elegant computational techniques. The key idea is to extend Markov random fields beyond FRAME by modeling the local field potentials with learned filters. To do so, the authors (Roth and Black) exploited ideas from the Products-of Experts (PoE) [3] framework.

Wilfred L. Rosenbaum, M. Stella Atkinsa, Gordon E. Sarty [2] applied waveletshrinkage denoising algorithms and Nowak's algorithm for de-noising the magnitude images. The wavelet shrinkage to denoising , methods were performed using the both hard and soft thresholding. It was said that changes in mean relative SNR are basically , statistically associated with the type of threshold and type of wavelet. Nowak's data-adaptive wavelet filtering was found to provide the overall good performance as compared to direct wavelet shrinkage.

Jiecheng Xie [4] mentioned the denoising method based on a doubly stochastic process model of wavelet coefficients that gave a new spatially differing threshold using the MDL principle. This method outperformed the traditional thresholding method in both MSE error and compression gain.

V. K. Nath and D. Hazarika have proposed video de-noising method which combines Wiener-bilateral filtering and selective recursive temporal filtering. The best combination of wavelet based local Wiener filtering and spatial domain bilateral filtering is referred to as Wiener bilateral filtering. Wiener-bilateral filter is used for spatial filtering where the bilateral filter parameters are selected,which are based on empirical study. The motion detection and temporal filtering is performed over the spatially filtered frames [03].

Denoising is still one of the most fundamental, widely studied, and largely unsolved problems in video processing. The purpose of denoising (or restoration) is to estimate the original video (or a “better” representative of it) from noisy data. Many methods for video denoising have been suggested, but the wavelet transform has been viewed by many as the preferred as noise removal technique [2]. Rather than a complete transformation into the frequency domain, as in DCT or FFT, the wavelet transform produces coefficient values

International Journal of Engineering Researches and Management Studies

which represent both time and frequency information. The hybrid spatial frequency presentation of the wavelet coefficients allows for analysis based on both spatial position and spatial frequency content.

Devi et. al. have proposed a video denoising algorithm which uses an efficient wavelet based spatio-temporal filter. The filter first applies 2D two dimensional discrete wavelet transform (DWT) and then applies 1-D discrete cosine transform (DCT) in the temporal direction in order to reduce the redundancies which exist among the wavelet coefficients in the temporal direction. The sub-band coefficients with huge magnitudes occur in clusters in different locations corresponding to the edge locations even after applying the above spatiotemporal filter. They have presented two different low complexity wavelet shrinkage based methods to denoise the noisy wavelet coefficients in different sub-band

DISCRETE WAVELET TRANSFORM

Wavelets are the signals which are basically local in time and scale and generally have an irregular shape. A wavelet is a waveform of which effectively limited duration that have an average value of zero(0). The word 'wavelet' derived from the information that they are integrate to zero; they wave up and down across the axis. Many wavelets are also demonstrate in a property ideal for compact signal illustration: orthogonality. This property ensures that data is not over represented. A signal can be disintergated into many shifted and scaled demonstration of the original wavelet. A wavelet transform can be used to decompose a signal into component wavelets. Once this is finished the coefficients of the wavelets can be decimated to remove some of the details. Wavelets have the great lead of being able to separate the good details in a signal. Lesser wavelets can be used to segregate a very fine details in a signal, while huge wavelets can identify the coarse details. In addition, there are many more different wavelets to select from. Many types of wavelets are: Morlet, Daubechies, etc. One particular wavelet may produce a more sparse representation of a signal than another, so different kinds of wavelets must be examined to see which is most well matched to image compression.

In this section we briefly review the wavelet decomposition and its use in noise filtering. A brief review of wavelets can be seen in a [9]- 11]. A. Discrete Wavelet Transform The Discrete wavelet transform (DWT)which can be seen as a filter bank algorithm that are iterated on the low pass output [10]. In the discrete wavelet transform, an image signal is passed through an analysis filter bank followed by decimation operation. The analysis filter bank consists of low-pass and high-pass filter at each level of decomposition. When the signal passes through these filters, it splits into bands. The low pass filtering produces an approximation of the signal, while the high pass filtering reveals the details that are expressed by wavelet coefficients. For reconstruction, the approximation and detail coefficient are upsampled and then filtered with a low pass and high pass filter followed by summation of the output. The above mentioned Discrete wavelet transform is crucially sampled (non-Discrete) and it is well known that the noise suppression improves when it is implemented in Discrete representation.

GAUSSIAN NOISE

Gaussian noise is statistical noise having a probability density function (PDF)that is equal to that of the normal distribution, which is can be also known as the Gaussian distribution In other words, the values which the noise can take on are Gaussian-distributed.

In communications, indiscriminate interference generated by the movement of electricity in the line. It is related to white noise, but restricted to a narrower variety of frequencies. You can actually see and perceive sound Gaussian noise when you tune your TV to a channel that is not operating. Contrast with white noise and pink noise.

A random distribution of artifacts in analog video images so as to makes everything look soft and slightly blurry. On the close up inspection, one can see tiny specks in unsystematic patterns. originate on films shot with older cameras as well as films and videotapes that have been archived for a long time, dynamic noise reduction (DNR) circuits can eliminate much of the Gaussian noise when the analog material is changed to digital

PROPOSED SYSTEM

The main idea about the FOE model is that with small patch size its easy to learn the number of parameters, added this model works for any size of images and this model is translational invariant which is one of the

International Journal of Engineering Researches and Management Studies

generic image priors. This we have achieved by combining the sparse coding along with Markov Random Field Models.

In our proposed methodology,an efficient method of denoising a noisy video is performed in two phases.**1]Training 2]Testing.**

A**] Training:** In the training stage, frames are originated from an input video. Discrete wavelet transform (DWT) is applied on only I frames. The main video frames of size mxn are contaminated by additive white Gaussian noise with zero mean and standard σ. The consequential of co-efficients are trained and are stored in the knowledge base.

we use full product of t-distribution [POT] model which is proposed by welling et.al to compute field of expert at pt X

$$P(x) = \frac{1}{Z(\Theta)} \prod_{i=1}^{N} \varphi_i [w_i^T x; a_{[i]}]$$

$where\ \square = \square_1, \square_2, \square_3 \dots \dots \square_N\ and\ \square_i = [a_i, w_i]\ ,\ and\ \varphi_i are\ the\ experts$

$$\varphi_i[w_i^T x, a_i] == [1 + \frac{1}{2}[w_i^T x]^2]^{-a_i}$$

is the partition [normalizing] function. Here we have assumed that are to be positive which are needed for proper distribution of experts but its very much clear that the experts by themselves are not normalized.
Here, the probability density form (Gibb's) can be written as

$$p(x) = \frac{1}{z(\Theta)} \exp[E_{POE}(x, \Theta)]$$

Now the sum of Energy value is calculated using the equation

$$E_{POE}(x, \Theta) = -\sum_{i=1}^{N} \log \varphi_i(w_i^T x, a_i)$$

Which is also equal to

$$P(x) = \frac{1}{Z(\Theta)} \prod_{p} \prod_{i=1}^{N} \varphi_i [w_i^T x; a_{[i]}]$$

Here we consider the product over all neighborhoods 'p' and the overlapping patches are highly correlated and the learned filter and the parameters must account for correlation. This POE model which is translational invariant is referred as Field of Experts [FOE] Model.

The energy gradient of FOE by parameters W at point X and update the values of W.

$$de(i) = \log\big(1 + (0.5 * wx(i)^2)\big)$$

The generated coefficients are stored in the Knowledge base.

Here we are maximizing the likelihood for FOE model by using contrastive divergence to achieve efficient sampling strategies to run for smaller fixed number of steps by initializing the sampler at the data points. By doing so for a very less iterations the sampler can draw samples that are very closer to the target distribution starting from the data distribution.

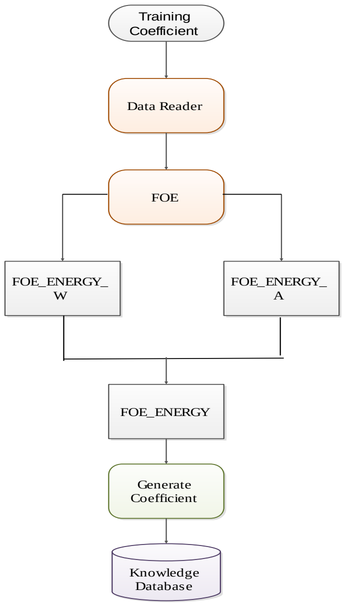

Fig.2 Training Coefficient

B.TESTING STAGE

Next in the testing stage, a noisy video is taken as input and the frames are extracted from the respective video. Figure 3 displays

the original input frame. Figure 4 depicts the frame polluted with white Gaussian noise. DWT based denoising is done in two stages: adaptive spatial filter and temporal filter. Next, inverse DWT is applied on the outcomes of these filters and results are sent to weighting block, where the resultant frame from both adaptive spatial filter and temporal filter are combined. This process is applied to every frame which is extracted from input video

Figure 3: Original Frame

Figure 4: Original Frame affected by Noise

the original input frame. Figure 4 depicts the frame polluted with white Gaussian noise. DWT based denoising is done in two stages: adaptive spatial filter and temporal filter. Next, inverse DWT is applied on the outcomes of these filters and results are sent to weighting block, where the resultant frame from both adaptive spatial filter and temporal filter are combined. This process is applied to every frame which is extracted from input video.

Initially, on each noisy frame DWT is applied which results in a set of co-efficients. These co-efficients are passed to adaptive spatial filtering.

IJERMS
International Journal of Engineering Researches and Management Studies

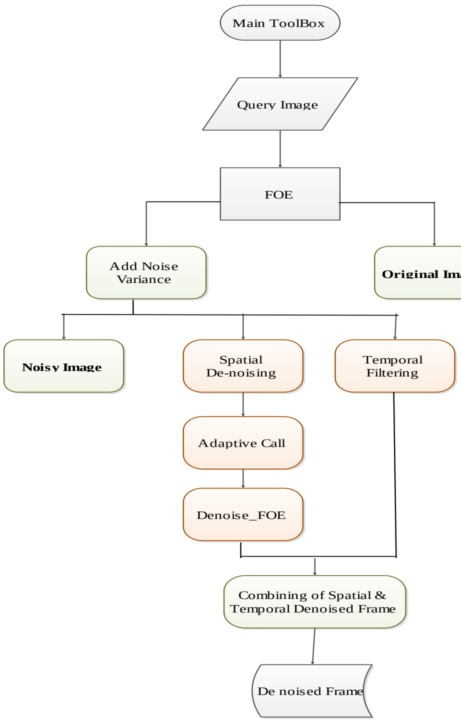

Fig 6:Testing Phase

Initially select a video and Generate i^{th} Frame by using the converter. Compute YCBCR for ith frame.Select FOE(Field of energy) patch size (5X5 or 3X3) & Noise variance. Apply spatial, Temporal filtering to Nosiy Image. Wieghing the average of the denoised image (both temporal and spatial filter) provides denoised output.

I) Adaptive Spatial Filter

Spatial filtering is especially useful at higher noise levels but even for smaller noise levels it can significantly improve the video quality. In order not to reduce the resolution of the input image sequence one has to adapt filtering to the presence of spatial details, i.e. take into account the (local and/or global statistical distribution of the image.

This filter modifies the co-efficients that are obtained during the training phase according to the DWT co-efficients. Later these modified co-efficients are passed to denoising module as decribed in (12)

Initialization

$$Noise = \left(\left(\frac{1}{sig^2}\right) * \frac{I-N}{255}\right)^2$$

$$de(i+1) = de(i) - conv(w,(wout/(1+(0.5*wout)^2$$

Gradient iterations is given by

$$d(i) = d(i) + \left(\frac{1}{sigma^2}\right) * \left(im - denoised_{image}\right)$$

$$Denoised_{im} = denoised_{imeta} * DI$$

II) Temporal Filter

Temporal filtering is an approach of exploiting temporal correlation to reduce noise in a video sequence. Temporal filtering is based on a simple block based motion detection and selective recursive time averaging of spatially filtered frames.

Blocks are generated from the frame the wavelet coefficients are obtained by applying DWT. The motion detector evaluates the absolute difference between the pixel value from the current frame and previous frame, where the considered pixels have the same spatial positions. If the absolute difference is above a predefined threshold then the temporal filtering is switched off otherwise recursive time averaging of frames is done as shown in fig 7.

Firstly, we calculate the Mean Absolute Difference (MAD) between the pixels in the corresponding blocks in the current and the previous frames.

$$D_{i,j}^{k} = \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}\left|g_{m,n}^{k,i,j} - g_{m,n}^{k-1,i,j}\right| \qquad (18)$$

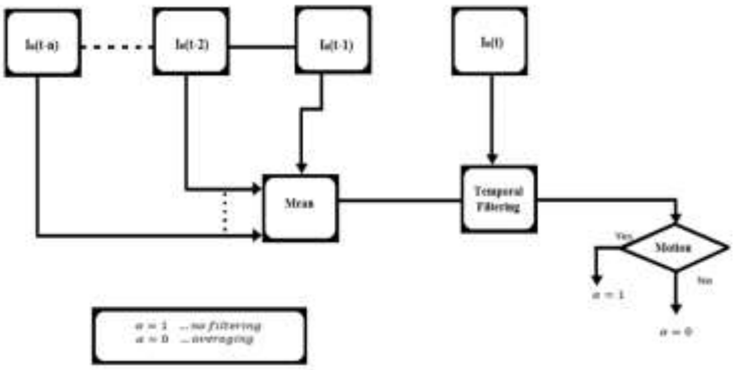

Figure 7 Block diagram of Temporal Filter

where k is the frame number, i, j , are the spatial coordinates of a block, m, n are the coordinates of a pixel inside the block, and N is the block size. In the filtering step, we determine whether motion exists in each block by comparing the absolute block difference with a threshold T.

Figure 8 depicts the frame which is temporally denoised. The outcome of this filter is passed to IDWT in order to obtain denoised frame.

Figure 8: Temporally Denoised Frame

II) Inverse Discrete Wavelet Transform:

In order to obtain denoised frame, co-efficients from both spatial filter and temporal filter have to undergo Inverse Discrete Wavelet Transform, which reconstructs the data. To preserve the features from both the images, coefficients from the approximation band are averaged,

International Journal of Engineering Researches and Management Studies

$$I_F^a = mean(I_P^a, I_R^a) \qquad (19)$$

where I_F^a is the approximation band of the fused image.

Finally, IDWT is applied on the four fused sub bands to generate the final enhanced image I_F.

$$I_F = IRDWT(I_F^a, I_F^v, I_F^d, I_F^h) \qquad (20)$$

Weighted Average

After spatial filtering and temporal filtering, we have two denoised frames. One is obtained from spatial filtering, in which the still regions are well denoised. Another is from temporal filtering, in which the motion regions are denoised to some extent. So, we integrate the two denoised frames by weighting them based on motion estimation results. Then, the weighted denoised frame can be calculated as follows

$$I_{Denoise} = w \times I_{FT} + (1 - w) \times I_{FS} \qquad (21)$$

Figure 9: Weighted Frame

Here, *w* represents weight matrix, represents denoised frame obtained from temporal filtering, represents denoised frame obtained from spatial filtering and represents the weighted frame which is final denoised frame. The output of this weighting process is as shown in Figure 9.

In the proposed sequential spatio-temporal filtering, motion detection and selective recursive temporal filtering is performed over spatially denoised frames. The motion adaptive temporal filtering is benefited from the use of high-quality spatial denoising [7]. The inter-frame differences due to remaining noise and artifacts after spatial wavelet denoising are relatively small compared to the actual inter-frame differences produced by motion. Therefore, simple block based motion detection technique is considered. The recursive motion adaptive filtering is given in Fig. 12

RESULTS AND COMPARISON

The performance of the proposed denoising algorithm is tested on four different gray-scale videos: "Foreman", "Mother daughter", "Akiyo". These test videos have been corrupted with Gaussian noise of the following standard deviation values: σ=30. The spatial filtering part is implemented with a non-decimated wavelet transform with 4 decomposition levels.

In the testing phase, denoising process follows two paths-temporal and spatial. For spatial denoising adaptive spatial filter is employed and in order to denoise in temporal domain recursive temporal filter is being used. The measurement is all made in terms of PSNR as defined in Equation (22).

$$PSNR = 10 \log_{10}\left(\frac{255^2}{MSE}\right) \qquad (22)$$

IJERMS
International Journal of Engineering Researches and Management Studies

Table 1: Result of Our Proposed System on Akiyo ,Foreman& mother-daughter for various frames

No of Frames/Video Sequence	Foreman σ=30	Mother-Daughter σ=30	Akiyo σ=30
2	35.9015	39.3625	38.5223
4	35.7129	38.5826	40.3623
5	35.8298	40.2323	41.2352
10	35.9015	39.5698	39.2352
20	35.8016	38.5323	40.6974
30	36.8025	41.3632	39.6637
50	36.5523	42.6213	41.6432

Table 2: Result of Our Proposed System on Akiyo,foreman & mother-daughter for different video formats

Video Formats/Video Sequence	Foreman σ=30	Mother-Daughter σ=30	Akiyo σ=30
SQCIF	35.3233	38.6323	39.2335
QCIF	36.3256	38.6634	40.3652
SCIF	34.2356	39.5623	39.3236
SIF	35.3623	38.6523	40.3656
CIF	38.3223	41.3223	41.3636
DCIF	38.5233	41.6355	42.3663

In order to improve the mean PSNR gain for different test sequences, the parameters of the temporal filter were optimised.

CONCLUSION

In this paper a novel spatio-temporal technique for video denoising, is presented, based on Discrete wavelet transform. Discrete wavelet transform ensures high quality video denoising. The video undergoes denoising in both spatial and temporal domain by employing adaptive spatial filter and recursive temporal filter respectively. These system comprises of two stages-training and testing. In training part, Discrete wavelet transform is applied on each frame and the resulting co-efficients are trained. The trained co-efficients are stored in the knowledge base. Next, in testing part, there are two steps. Firstly, spatial filtering is done by taking Discrete wavelet transform of individual frames. The trained coefficients are modified based on the Discrete wavelet coefficients. Further, temporal filter is employed in order to remove noise due to motion. This includes block based motion detection and recursive time averaging of spatially filtered frames. Finally, by weighting the

IJERMS
International Journal of Engineering Researches and Management Studies

denoised frames from both spatial filtering and temporal filtering, a satisfactory result is obtained.we have worked out for different video formats namely Forman,mother and daughter and akiyo.

REFERENCES

1. G. Varghese and Z. Wang, “Video Denoising Based on a Spatiotemporal Gaussian Scale Mixture Model”, IEEE Transactions on Circuits and Systems for Video Technology, Volume 20, Issue 7, pp. 1032-1040, 2010.
2. D. Van De Ville and M. Kocher, “Non-Local Means with Dimensionality Reduction and SURE-Based Parameter Selection”, IEEE Transactions on Image Processing, Volume 20, Issue 9, pp. 2683-2690, 2011.
3. V. K. Nath and D. Hazarika, “Noise Reduction in Video Using Wiener-bilateral filter”, Journal of Computer Engineering, Volume 3, Issue 1, pp. 1-5, 2012.
4. K. Devi, D. Hazarika and V. K. Nath, “Video Denoising Based on Exploiting Intrascale and Interscale Dependency in Wavelet Domain”, International Journal of Computer & Communication Technology, Volume 4, Issue 3, pp. 11-15, 2013.
5. Jeena Joy, Salice Peter and Neetha John, “Denoising using Soft thresholding”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering , Volume 2, Issue 3, pp. 1027-1032, 2013.
6. S. Mishra and P. D. Swami, “Spatio-Temporal Video De-noising by Block-Based Motion detection”, International Journal of Engineering Trends and Technology, Volume 4, Issue-8, pp. 3371-3382, 2013.
7. M. Vatsa, R. Singh and A. Noore, “Denoising and Segmentation of 3D Brain Images”, In the Proceedings of IPCV, 2008.
8. Y. Jin, E. Angelini and A. Laine, “Wavelets in Medical Image Processing: Denoising, Segmentation, and Registration,” Handbook of Biomedical Image Analysis, pp. 305–358, 2005.
9. E. Angelini, A. Laine, S. Takuma, J. Holmes and S. Homma, “LV Volume Quantification via Spatio-Temporal Analysis of Real-Time 3D Echocardiography” ,IEEE Transactions on Medical Imaging, Volume 20, Issue 6, pp. 457–469, 2001.
10. T. M. Moldovan, S Roth and M. J. Black, “Denoising Archival Films using a Learned Bayesian Model”, IEEE International Conference on Image Processing, pp. 2641–2644, 2007.
11. D. Donoho, “De-noising by Soft-thresholding, “IEEE Transaction on Information Theory”, Volume 41, Issue 3, pp. 613–627, 1995.
12. S. Roth and M. J. Black, “Fields of Experts: A Framework for Learning Image Priors”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 2, pp. 860–867, 2005.
13. [Hwea Yee Lee, Wai Lam Hoo, and Chee Seng Chan, “Color video denoising using epitome and sparse coding”, Elsevier, pp.751-759,2015