An Artificial Neural Network Based Lossless Video Compression using Multi-Level Snapshots and Wavelet Transform using Intensity measures

S.Ponlatha[#1] , Dr. R.S. Sabeenian*[2]

[1]Associate Professor/ ECE,AVS Engineering College, Ammapet, Salem-636003.
ponlathasenthil@gmail.com

[2]Professor/ECE & Centre- Head,SIPRO, Sona College of Technology,Salem
sabeenian@gmail.com

Abstract:-Transmission of uncompressed video segments requires more bandwidth and need more storage for video data. Earlier approaches produces compressed videos , where there exists data loss and they produce less effective results at decompressing video segment. We propose a new lossless approach to perform compression of video segment, where snapshot level video compression is done by splitting the video into tiny snapshots. The snapshots are trained with the neural network and the trained values are used to perform video compression. The input video is de-framed to generate sequence of scenes and compressed with the help of wavelet transform. The efficiency of wavelet helps to reduce the signals into small set of signals. For each frame generated from the video, the features of the scene are extracted and used to identify the frames of scenes. A feature variance matrix is generated by identifying the variance of features between subsequent frames of any scene identified. Based on computed variance matrix, the unaffected pixels are neutralized and the pixels with variance are kept original. Then the frames are transformed with wavelet, and the transformed signals are applied with neural network to come up with output signals. The proposed method has produced higher rate of compression and produces efficient results in decompression. Also the proposed method overcomes the problem of loss introduced in video compression by other methods.

Keywords: Lossless compression, Video Compression, Wavelet Transform, Sub sequence, Neural Networks.

I.INTRODUCTION

A. Lossless and Lossy compression

There are two basic categories of compression; lossless and lossy. Lossless compression is a class of algorithms that will allow for the exact original data to be reconstructed from the compressed data. It means that a limited amount of techniques are made available for the data reduction, and the result is limited reduction of data. GIF is an example of lossless images compression, but is because of its limited abilities not relevant in video surveillance. Lossy compression on the contrary means that through the compression data is reduced to an extent where the original information cannot be obtained when the video is decompressed. The difference is called the artifacts.

B. Wavelet Transform

Wavelets have generated an interest in researchers in both theoretical and applied areas, especially over the past few years. The wavelet transform can be used to reduce the image size without losing much of the resolutions computed and values less than a pre-specified threshold are discarded. Thus it reduces the amount of memory required for an image.

C. Neural Networks

A neural network can be defined as a “massively parallel distributed architecture” for storing experimental knowledge and making it available for use. It refers to a computational paradigm in which a large number of simple computational units are interconnected to form a network, performing complex computational tasks [1].

A neural network is a system of interconnecting neurons in a network working together to produce an output function. Neural networks can accept a vast array of input at once and process it quickly, so they are useful in image compression [2]

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the biological nervous systems, such as the brain. The key element of this paradigm is the structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons.

D. Video Segmentation

The video segments are incorporated in many applications with the support of modern communication technology [3]. The communication standards works on the top of resources like bandwidth, which is a dominant resource and has to be used efficiently. For a video communication in mobile applications, the whole video has to be transferred to other end, and the video compression has various standards which are proposed earlier. A video is a collection of frames or snapshots and has to be compressed to low bit rate without damage the content and meaning of the frame. For example for a single snapshot it needs 1024 bytes to store and transfer which has to be minimized.

The frames of a video segment are more related and have little amount of changes in the pixels of the frames. For a single video segment[4], there will be number of frames according to type of video. There are high redundancies in the same frame, the spatial redundancies, and between successive frames, the temporal redundancies. Most of video compression techniques aim to exploit these correlations and high redundancies in order to achieve better compression.

The video compression methods can be classified as Object based, Wave form based, model based, fractal based. The Object based compression is performed by extracting the objects from video frame and their shape, texture are extracted which will be coded differently[5]. The compression ratio of each object will be different and yield good compression ratio. In wave form method, the video frame is transformed to small size of frame without missing any feature, this also achieves higher rate of compression. The model based compression performs 3D structural analysis to compress the frames and the fractal based methods uses image coding methods to compress the video frames.

E. Video Quality Measure

In order to evaluate the performance of video compression coding, it is necessary to define a measure to compare the original video and the video after compressed. Most video compression systems are designed to minimize the *mean square error (MSE)* between two video sequences Ψ1 and Ψ2, which is defined as

$$\mathrm{MSE}=\sigma^2_e=\frac{1}{N}\sum_t \sum_{x,y}[\Psi 1(x,y,t)-\Psi 2(x,y,t)]^2 \qquad (1)$$

where N is the total number of frames in either video sequences.

Instead of the MSE, the *peak-signal-to-noise ratio (PSNR)* in decibel (dB) is more often used as a quality measure in video coding, which is defined as

$$\mathrm{PSNR} = 10 \log_{10} C\ /\mathrm{MSE} \qquad (2)$$

where $(2^n-1)^2$ is the square of the highest-possible signal value in the image, and *n* is the number of bits per image sample. PSNR can be calculated easily and quickly and is therefore a very popular quality measure, widely used to compare the ‘quality’ of compressed and decompressed video images. For a given image or image sequence, high PSNR usually indicates high quality and low PSNR usually indicates low quality. However a particular value of PSNR does not necessarily equate to an ‘absolute’ subjective quality.

We work on the wavelet transform, which performs transformation of signals into small region by performing scaling and transforming signals. The wavelet transform method produces high quality results in signal processing where the size of signal is compressed and represented with small values. What happens when compression video frames is, some of the values of few pixels gets damaged which is considered as loss. The video compression method has to take care of this loss and has to produce lossless compression.

II.BACKGROUND

There are various video compression method discussed in literature and we discuss few of them here according to our problem statement.

Video Compression Using Neural Network [6], implements different algorithms like gradient descent back propagation, gradient descent with momentum back propagation, gradient descent with adaptive learning back propagation, gradient descent with momentum and adaptive learning back propagation and Levenberg-Marquardt algorithm.

Object-based Video compression using neural networks [7], It consists one predicting video objects motions throughout the scene. Neural networks are used to carry out the prediction step. A multi-step- ahead prediction is performed to predict the video objects trajectories over the sequence. In order to reduce video data, only the background of the video sequence is transmitted with the different detected video objects as well as their initial properties such as placement and dimensions.

Adaptive Filtering and Video Compression Using Neural Networks [8], present an Adaptive Filtering technique for the removal of useless noise from video. After cancelling the unwanted noise, we get a filtered video. After that we take the filtered video as input to the neural network. Finally, the back propagation algorithm is used to compress the video.

High efficiency video coding: the next frontier in video compression [9], was launched to achieve major savings-e.g., reduction by about half for 1,280 × 720 high-definition (HD) and higher-resolution progressives can video-for equivalent visual quality relative to the bit rate needed by the widely used H.264/ MPEG-4 Advanced Video Coding (AVC) standard. For high resolution video where such additional compression is most urgently required, implementations of the current draft standard are already meeting or exceeding the targeted goal. We review the architecture and building blocks of HEVC[10], which were carefully selected with regard to compression capability versus complexity and to enable parallelism for the signal processing operations. Given the benefits that HEVC provides, it is likely to become the new primary reference for video compression.

Royalty free video coding standards in MPEG [11], report on the recent developments in royalty-free codec standardization in MPEG, particularly Internet video coding (IVC), Web video coding (WVC), and video coding for browser, by reviewing the history of royalty-free standards in MPEG and the relationship between standards and patents. A Video Saliency Detection Model in Compressed Domain [12], propose a novel video saliency detection model based on feature contrast in compressed domain. Four types of features including luminance, color, texture, and motion are extracted from the discrete cosine transform coefficients and motion vectors in video bitstream. The static saliency map of unpredicted frames (I frames) is calculated on the basis of luminance, color, and texture features, while the motion saliency map of predicted frames (P and B frames) is computed by motion feature[13]. A new fusion method is designed to combine the static saliency and motion saliency maps to get the final saliency map for each video frame. Due to the directly derived features in compressed domain, the proposed model can predict the salient regions efficiently for video frames.

In [14], the author proposes two schemes that balance energy consumption among nodes in a cluster on a wireless video sensor network. In these schemes, we divide the compression process into several small processing components, which are then distributed to multiple nodes along a path from a source node to a cluster head in a cluster. We[15] conduct extensive computational simulations to examine the truth of our method and find that the proposed schemes not only balance energy consumption of sensor nodes by sharing of the processing tasks but also improve the quality of decoding video by using edges of objects in the frames.

III.PROPOSED APPROACH

The proposed lossless compression method has four stages: First the video is segmented into frames, second: scenery identification is performed, Third: variance matrix is computed between successive frames, Fourth: Video compression is performed using wavelet transform.

A. Video Segmentation

Video segmentation is performed at the first stage of video compression. At this stage, the video is split into frames according to the value of M. M- specifies the number of frames to be generated at each second or the snapshots to be produced. The segmented video will be used for further processing. For a given video V, we generate frame set FS.

1) Algorithm for Video Segmentation:

step1: start

step2: initialize scenery set VS, Frame set Fs

step3: read input video V.

step4: Play video V.

for each second of V.

generate snapshot Ss. //Here we generate snapshots at particular second

FS = FS+ Σ($\int_1^N (V \times M)$). // Fs is sum of all snapshot of video at all seconds

M- number of frames per second.

end.

step5: stop.

The above segmentation algorithm reads the input video and de-frames the video into N number of snapshots according to the video quality. The number of frames to be generated for single second of video is set manually, which is decided according to the time complexity and accuracy of compression needed. Also the problem of restoring original video also must be considered which deciding the value for N.

2) Scene Identification:

The video may composed of any number of scene included and the feature variance or the common features will be there in the frames of any scene. The scene identification performs the identification process of scenes using the frames available. The method computes the intensity pattern available in each frame and computes the intensity mean of each frame. Once the intensity mean of all frames are computed, then each frames intensity mean will be compared with previous frames to compute the distance between the two frames. The method maintains a intensity threshold and if the distance is within the threshold then the two frames are considered as in the same scene. Otherwise the scene will be completed and the new frame is assigned as a starting frame of another scene. This will be performed for each frame generated in the previous stage.

step1: start

step2: initialize Intensity pattern Ip, Scene set SS.

step3: read Frame set Fs.

step4: for each frame F_i from FS

compute intensity pattern Ip. //we generate intensity pattern here

Ip (i)= $\sum_{1}^{x} Mean(R, G, B)$ //computing mean value of red, blue ,green

end

step5:for each frame F_i from Fs

compare intensity pattern with neighbor frames. //comparing the pixels with neighbor frames

if Distance(Ip_i, Ip_j)<Intensity threshold then //

the frames are in same scene.

SS(i) = SS(i)+ F_i. // here we identify the frames of same scene

else

create new Scene S.

S(i) = S(i)+F_i.

end

end.

step6: stop.

B.Variance Matrix Computation

At this stage, we compute the variance between two frames of a video sequence. The variance value represent the change of pattern occurred at one frame from its previous frame. The pixel values which are differ from previous frame only will be retained and others will be removed. This helps to represent the frame with small bits and easier for compression also this itself produces higher compression ratio.

1) Algorithm for Variance Matrix Computation:

step1: start

step2: initialize variance matrix Vm, frame matrix fm, intensity pattern Ip.

step3: read scene frames set SS.

step4: for each frame S_i from SS

compute variance matrix Vm_i. for each pixel Pi from S_i

compute variance v $= \int_{1}^{X} Ip(m) - Ip(n)$ // variance between frames using intensity pattern

if v>vth then //vth-variance threshold'

keep the pixel with the frame F_i.

else

neutralize the pixel from the frame F_i.

$S_{i(Pi)} = 0$.

end.

end.

step5: stop.

C. Video Compression

Once variance matrix is computed then the frames of the sequence will be applied with wavelet transform. The wavelet transform performs scaling and transformation of signals and produces low size signal which represent the originals signals. For each video clip, we apply wavelet transform on the frames of the clips. The transformed signals are applied with artificial neural networks and the function designed with multilayer perception combines the transformed signal and encodes the level of transformation applied to produce the output signal. The artificial neural network encodes the level of transformation applied at wavelet method to produce the output signal. The compressed images are combined to reproduce the video which can be decompressed later.

1) Algorithm for Video Compression:

step1: start

step2: initialize compressed video CV.

step3: read frame sequences Fs.

step4: for each frame Fi from Fs

if(i=1)

$F_i = DWT(F_i)$

Apply neural network with F_i.

$F_i = ANN(F_i)$.

else

$F_i = DWT(Vm(F_i))$.

Apply neural network with F_i.

$F_i = ANN(F_i)$.

end.

end.

step5: generate video clip from frame set Fs.

step6: stop.

IV. RESULTS AND DISCUSSION

The proposed approach has been implemented and evaluated with other methods in Matlab and we have used different video data set to compare the results produced by the proposed method and others.

Fig 1: Snapshot generated from the input video clip.

The figure 1, shows the single frame or snapshot generated by the proposed method.

Fig 2: Result of Compressed image with wavelet transform.

The figure 2 shows the result generated by wavelet transform and the resultant image has retained the quality as well.

Image/Video compression results are presented for various test scenarios with the help of a motion picture obtained from geographic sites which contain a set of 98 frames. The set of frames are tested for, the retraining frames method and the self-adaptive method. The results obtained from the above tests were useful in drawing some conclusions regarding the aforementioned techniques. The Compression Ratio and Peak-Signal-to-Noise Ratios (PSNR) are calculated based on the following formulas for all the test scenarios.

$$Compression\ Ratio = \frac{(K\times L\times M\times N)}{\mathrm{T}} \qquad (3)$$

K- Number of variant pixels

L- Level of wavelet transfer

M- Number of pixels

T-Total number of pixels

PSNR=10× log (1/ N).

Initially, the "Leopard" image is trained for different epochs (5000,10000,20000) and 4 hidden nodes in the network. The network is also tested for a still image with the above parameters. We can see that as the training was increased to 20000 epochs, the weights seem to be better adapted to the particular image, and thus the quality of the reconstructed image is higher compared to the one trained for 5000 or 10000 epochs. Nevertheless, training for 20000 epochs has higher processing requirements compared to the other two cases. Moreover, Graph1 illustrates that as the compression ratio increases; the difference in terms of PSNR between the three different cases becomes negligible.

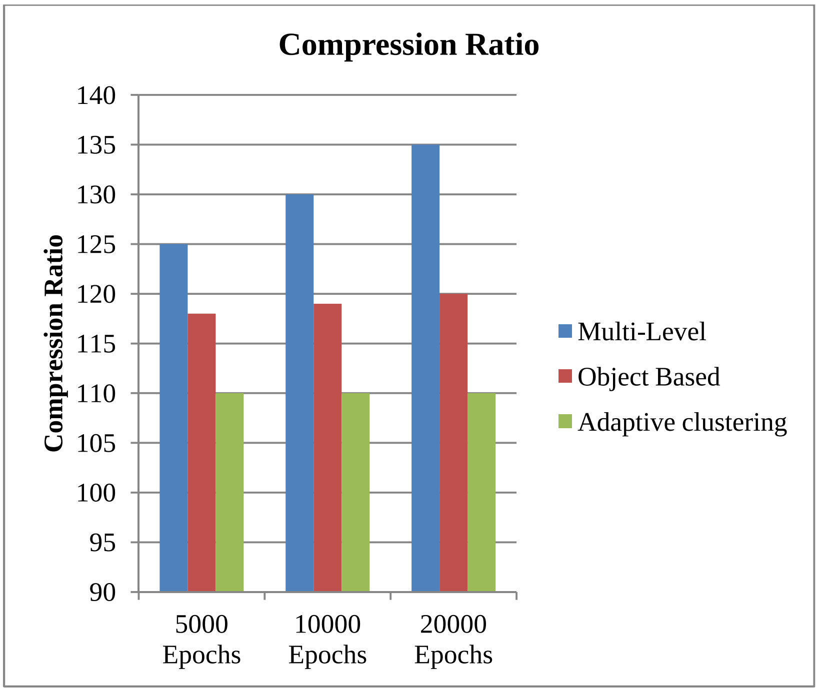

Fig 3: Compression ratio of different methods

The figure 3, shows the compression ratio generated by different methods at different epochs, for each of the epochs, the frames are generated and compressed. The proposed method has produced efficient compression ratio compared to other methods. The method has been evaluated with more epochs from 1 epoch. In all the cases, the proposed method has produced efficient compression than others.

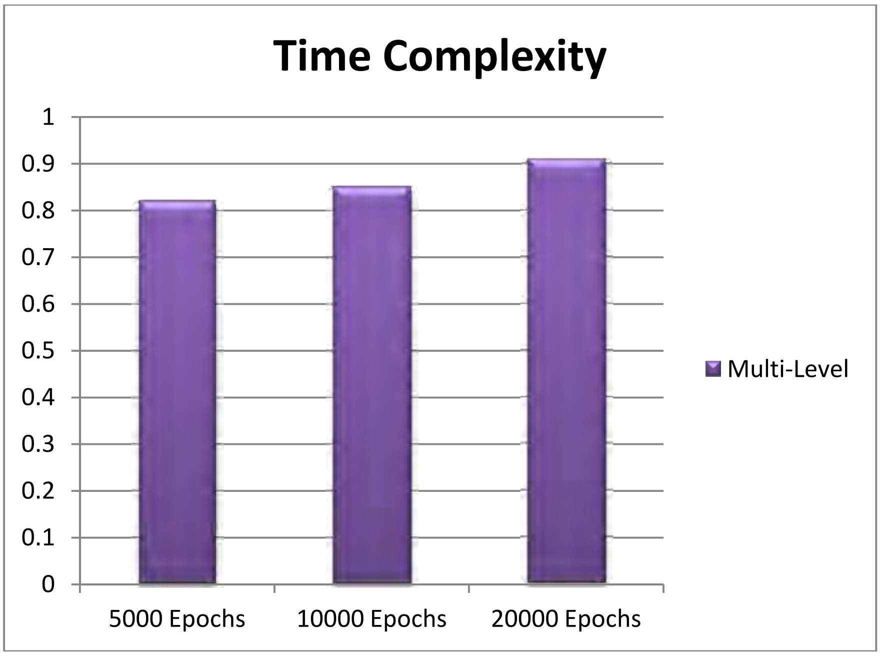

Fig 4: Time complexity graph

The figure 4 shows the time taken value of the proposed method for different epochs of jpeg images. It shows that the proposed method has taken less time to compress the image. The proposed multi-level snapshot algorithm takes less time to compress the video than earlier approaches. To compress a video with 100,300,500 second or 5000,10000 & 20000 epoch , the proposed method takes less time only.

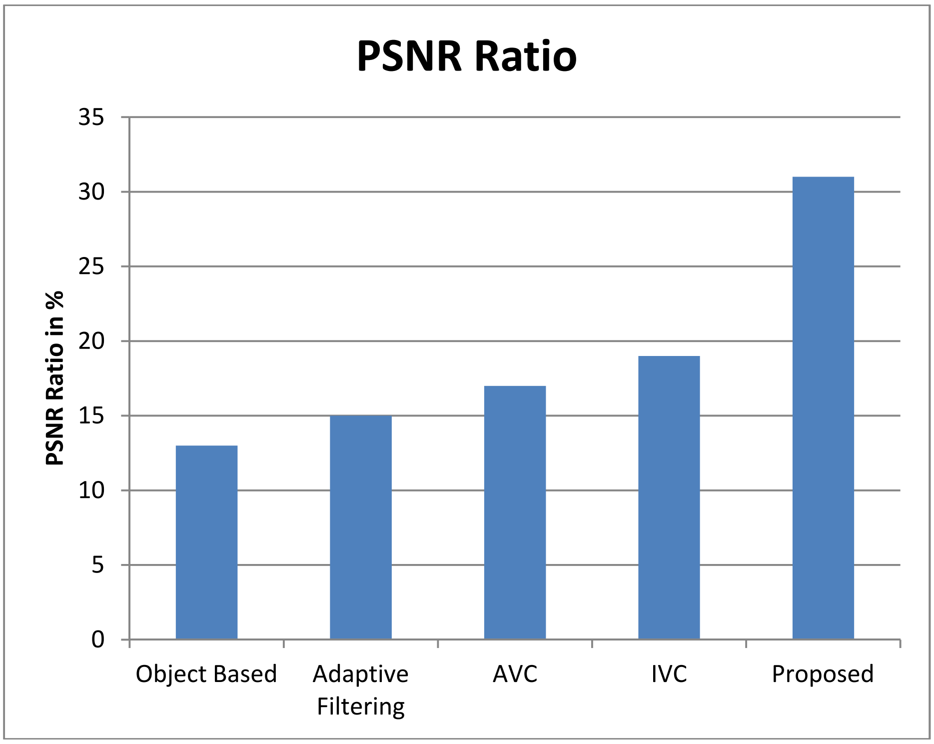

Fig 5: Comparison of PSNR values.

The figure 5 displays the comparison of PSNR value produced by different methods. It shows clearly that the proposed approach has produced more peak signal to noise ratio than others.

V. CONCLUSION AND FUTURE WORK

A lossless approach for video compression with wavelet transform and the support of intensity measures are proposed. The video is split into frames and then scene identification is performed using intensity pattern change. Based on identified scenes, the variance of subsequent image or frames within the scene is computed. The pixel which varies with intensity is identified and kept original and others are neutralized. This makes the size of frame lesser, because only the variant pixels are used and this reduces the space also. After this, the frames are compressed using wavelet transform which increases the compression ratio. The proposed approach has produced efficient compression ratio and produced high quality video. The proposed method has overcomes the problem of energy loss occurred in previous methods and the following are planned for future study: In order to improve the compression efficiency and accuracy in high definition (HD), we would suggest to continue to develop the proposed programming system by implementing in JAVA platform can reduce the time complexity, we can also use fuzzy logic and other algorithms of neural network for obtaining MSE and PSNR along with better compression ratio.

REFERENCES

[1] J. Humphreys and A. Hunter, "Multiple object tracking using a neural cost function", Image and Vision Computing 27, 2009, pp. 417–424..

[2] Souwmaya Ghorbel, "Object-based Video compression using Neural networks", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 1, July 2011

[3] H.B. Kashani, S. A. Seyedin, and H. S. Yazdi, "A Novel Approach in Video Scene Background Estimation", International Journal of Computer Theory and Engineering, Vol. 2, No.2, April, 2010, pp. 274- 282.

[4] L. Zhicheng, Q. Shiyin, and I. Laurent, "Visual attention guided bit allocation in video compression", Image and Vision Computing 29 , 2011, pp. 1–14.

[5] S.Ponlatha, R.S.Sabeenian "Comparison of Video Compression Standards" International Journal of Computer and Electrical Engineering , Vol. 5, No. 6, December 2013, pp-549-554

[6] Sangeetha Mishra," Video Compression Using Neural Network", International Journal of Computer Applications & Information Technology Vol. I, Issue II, September 2012.

[7] N. Ling, "Expectations and challenges for next generation video compression," in Proceedings of the 5th IEEE Conference on Industrial Electronics and Applications (ICIEA '10), pp. 2339– 2344, June 2010.

[8] Kalyan Chatterjee , "Adaptive Filtering and Video Compression Using Neural Networks", International Journal of Inventive Engineering and Sciences (IJIES), Volume-1, Issue-4, March 2013.

[9] Ohm J, "High efficiency video coding: the next frontier in video Compression" , IEEE Transactions on Signal processing, vol:30, issue:1, PP:152-158, 2013.

[10] Paurazad,HEVC: "The New Gold Standard for Video Compression: How Does HEVC Compare with H.264/AVC?", IEEE Transaction on consumer electronics magazine, vol:1, issue.3, 2012.

[11] Choi K, "Royalty-Free Video Coding Standards in MPEG", IEEE Transaction on Signal Processing, vol:31, issue:1, PP:145-155, 2014.

[12] Yuming Fang,," A Video Saliency Detection Model in Compressed Domain", IEEE Transaction on circuits and systems on video technology, vol.24, issue1, 2014.

[13] L. Chang, C. Y. Pong, and Q. Guoping, "Object motion detection using information theoretic spatiotemporal saliency", Pattern Recognition 42, 2009, pp. 2897-2906.

[14] C. Yeo and K. Ramchandran, "Robust distributed multiview video compression for wireless camera networks," IEEE Transactions on Image Processing, vol. 19, no. 4, pp. 995–1008, 2010.

[15] Phat Nguyen Huu,Vinh Tran-Quang, and Takumi Miyoshi "Video Compression Schemes Using Edge Feature on Wireless Video Sensor Networks", Journal of Electrical and Computer Engineering Volume 2012, 2012.

BIBLIOGRAPHY

S.Ponlatha is currently working as a Associate Professor in ECE Department in AVS Engineering College ,Salem. She received her Bachelors in Engineering from Madras University and her Masters in Engineering in Communication Systems from Anna University and she is Pursuing PhD Degree under Anna University, Chennai in the area of Digital Image Processing. She published papers in International Journals & Conferences. She is a Member of ISTE, IEEE, IETE.

Dr.R.S.Sabeenian is currently working as a Professor in ECE Department in Sona College of Technology,Salem, Tamil Nadu, and India. He received his Bachelors in Engineering from Madras University and his Masters in Engineering in Communication Systems from Madurai Kamaraj University. He received his Ph.D. Degree from Anna University, Chennai in the year 2009 in the area of Digital Image processing. He is currently heading the research group named Sona SIPRO (SONA Signal and Image PROcessing Research Centre) centre located at the Advanced Research Centre in Sona College of Technology, Salem. He has published more than 65 research papers in various International, National Journals and Conferences. He has also published around seven books. He is a reviewer for the journals of IET, UK and ACTA Press Singapore. He received the “Best Faculty Award” among Tamil Nadu, Karnataka and Kerala states for the year 2009 given by the Nehru Group of Institutions, Coimbatore and the “Best Innovative Project Award” from the Indian National Academy of Engineering, New Delhi for the year 2009 and “ISTE Rajarambapu Patil National Award” for Promising Engineering Teacher for Creative Work done in Technical Education for the year 2010 from ISTE. He has also received a Project Grant from the All India Council for Technical Education and Tamil Nadu State Council for Science and Technology, for carrying out research. He received two “Best Research Paper Awards” from Springer International Conference and IEEE International Conference in the year 2010.He was also awarded the IETE Biman Behari Sen Memorial National Award for outstanding contributions in the emerging areas of Electronics and Telecommunication with emphasis on R&D for the year 2011.The Award was given by Institution of Electronics and Telecommunication Engineers (IETE), New Delhi. He is the Editor of 6 International Research Journals Research,Journal of Information Technology, Asian Journal of Scientific Research, Journal of Artificial Intelligence, Singapore Journal of Scientific Research, International Journal of Manufacturing Systems and ICTACT Journal of Image Processing. He is also associated with the Image Processing Payload of the PESIT.