Video Watermarking Scheme using Motion Vectors With RBF Neural Network

Anurag Mishra
Department of Electronics
Deendayal Upadhyay College
University of Delhi

Latesh Sharma
M.Tech. (Elect. & Comm.)
GGSIPU
New Delhi

Abstract— **Multimedia processing with real time constraints is one of the challenging tasks of the present day research. Uncompressed and compressed videos are important component of this domain. In this paper, we propose a novel uncompressed digital video watermarking scheme based on 3-level – DWT and Radial basis function (RBF) neural network by computing its motion vectors in uncompressed domain. To achieve this, first, the video is divided into frames and then the frames of maximum motion are determined. Second, every selected video frame is transformed using 3-level DWT. The watermark is embedded in the LL3 sub-band wavelet coefficients. An Radial Basis Function neural network is used to memorize the relationship between the coefficients so that the watermark is embedded and extracted correctly. Three video processing attacks are executed over signed video frames. Experimental results calculate the average peak signal to noise ratio (PSNR) for visual frame quality, normalized correlation (NC(W,W')) and Bit Error Rate (BER) for assessment of embedding algorithm. The average PSNR is high which shows that visual quality of frames post embedding is good. It is concluded that the proposed motion vector and RBF neural network based video watermarking scheme is robust against the selected attacks. Time complexity calculations indicate that this scheme is suitable for real time watermarking application.**

Keywords— *Motion vectors, 3-Level DWT, RBF neural network, Uncompressed AVI Video.*

I. INTRODUCTION

With the evolution of multimedia applications, the threats of copyright violation and destruction of digital content such as digital text, image, video and audio have become the order of the day. Therefore, it has become necessary to legally protect this content by applying robust copyright protection and authentication schemes. Digital watermarking is one such solution. Numerous watermarking algorithms have been developed for videos, some of these may be in compressed domain and some may operate in uncompressed domain. The watermarking techniques must fulfill the twin criteria – visual quality of watermarked image/frames and the robustness of the embedding algorithm. It has been often found that these two criteria are mutually exclusive. If we intend to increase robustness, it can be done at the cost of decreasing visual quality of the signed content. Therefore, the problem of watermarking, in general, has become an optimization

problem over a long period of time. Moreover, real time constraints require processing a given video at run time. In view of this, fast executing watermark embedding and extraction schemes are much in demand.

The transform domain based watermarking schemes usually result in robust watermark embedding as the low frequency or mid frequency band coefficients are modulated in this case. Among these kind of algorithms, DWT based schemes are found to produce better results in comparison to DCT based ones, both in terms of visual quality and robustness [1].

As mentioned above, the problem of digital watermarking is presently being tackled by using optimization cum soft computing techniques; Several researchers worldwide have extensively relied upon the use of artificial neural network (ANNs) to implement image and video watermarking [2-5]. Piao et. al.[2] have presented blind image watermarking algorithm based on human visual system (HVS) in DWT domain using Radial Basis Function Neural Network (RBFNN). Similarly, Xuefang Li. et. al.[3] have developed a video watermarking scheme based on 3D-DWT and Neural Network in which the watermark is embedded in the LL band and the network is trained using artificial neural network. This scheme is proved to be invisible and robust against various video processing attacks. Mahar EL' ARBI et. al.[4] have developed a novel video watermarking algorithm based on multi-resolution motion estimation and Artificial Neural Network. In this case, a multi-resolution motion estimation algorithm is adopted to preferentially allocate the watermark to the coefficients containing the motion. A Back propagation Neural Network (BPNN) is used to memorize the relationship between coefficients in 3*3 sized blocks of the image. The authors claim that their proposed scheme is robust against common video processing attacks [5]. Recently, a novel image watermarking algorithm is developed for gray scale images using a newly developed Single Layer Feed-forward Network (SLFN) commonly known as Extreme Learning Machine (ELM) [6]. This machine is very fast and completes its training in few milliseconds. The authors have shown that this machine can be successfully utilized for developing real time watermarking applications. In addition to this, Chetan et. al. [7] have developed DWT based blind digital video watermarking scheme for video authentication based on scene change and various attacks are implemented to check the

robustness of the proposed scheme. On the similar lines, Aroh Barjatya [8] and Zhang et. al. [9] have respectively developed algorithm for motion vector estimation and used it for video watermarking.

In this paper, we propose a Radial basis function neural network (RBFNN) based approach to implement watermarking of frames of three given standard uncompressed video using their motion vector computation. The three selected videos comprise of 100 frames each. A frame selection criteria based on the motion vector computation is first applied. A data is developed by using the LL3 sub-band coefficients obtained after applying 3-level DWT on the blue channel of the selected video frames. RBFNN is trained using this data which produces a row vector of normalized coefficients. A preprocessed binary watermark is embedded within the coefficients of row vector so as to obtain signed video frames. The frames are reorganized within the video so that the resultant video becomes watermarked video. Extraction of watermark is also carried out by using reverse process to that of embedding. The watermarking scheme is explained in detail in the section 3 of this paper. It is concluded that the proposed watermarking scheme is suitable for developing real time watermarking applications.

The issue of robustness of the proposed watermarking scheme is examined by executing three different video processing attacks. These are - scaling attack, Gaussian noise attack and JPEG compression attack. Plots of PSNR, NC and BER are plotted with respect to noise variance, scaling (%) and JPEG compression quality. The obtained results indicate that the proposed watermarking scheme is robust against the selected attacks.

The remaining paper is organized as follows: Section II explains the basics of the radial basis function neural network (RBFNN) and its training. Section III gives experimental details. The motion vector based frame selection criterion is an integral part of this section and is described in section A. Section B gives watermark embedding algorithm and section C gives watermark extraction algorithm along with their evaluation parameters. Results and discussion are presented in section IV. We concludes the complete experimental work and its findings and analysis with conclusion and at the end we had listed the references.

II. RADIAL BASIS FUNCTION NEURAL NETWORK AND ITS TRAINING

The RBFNN is a three layer neural network as shown in Fig. 1as referred in [10]. The idea of RBFNN is derived from the theory of function approximation [9]. Their main features are as follows:

- It is a two-layer feed-forward network
- The hidden nodes implement a set of radial basis functions (e.g. Gaussian functions)
- The output nodes implement linear summation function as in a multilayer perceptron
- The network training is divided into two stages:- first the weights from the input to hidden layer are determined, and then the weights from the hidden to output layer are determined
- The training / learning is fast
- The network is good at interpolation

The RBFNN is called so because its activation function depends only on the distance from the center vector and is radially symmetric about center vector. Inputs Weights RBF Output Output Function Weights
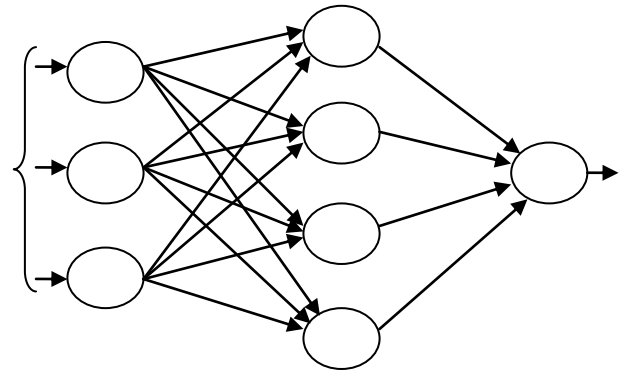

Fig. 1: A typical RBFNN Architecture

Eqn. 1 specifies the weights related to RBF[9][12]: It is easy to determine equations for the weights by resolving Eqn. 1.

$$f(x^p) = \sum_{q=1}^{N} w_q \varphi(||x^p - x^q||) = t^p \qquad (1)$$

In Eqn. 2, the distances $\varphi(||x^p - x^q||)$ between data points p and q are fixed by the training data,

$$\varphi_{pq} = \varphi(||x^p - x^q||) \qquad (2)$$

so Eqn. 3 is simply an array or matrix of training data dependent constant numbers and the weights w_q are the solutions of the linear equations.

$$\sum_{q=1}^{N} \varphi_{pq} \ w_q = t^p \qquad (3)$$

This can be written in the form of a matrix by defining the vectors $t = \{t^p\}$ and $\mathbf{w} = \{w_q\}$, and the matrix$\varphi = \{\varphi_{pq}\}$, so the equation for $\mathbf{w}$ simplifies to$\varphi_w = t$.

Determination of Weights: It can be said that any standard matrix inversion technique can be used to give the required weights provided the inverse of the matrix W exists. Thus,

$$W = \varphi^{-1} t \qquad (4)$$

Where the inverse matrix φ^{-1} is defined by$\varphi^{-1}\varphi = I$. These inverse matrices can very well be computed by efficient computational algorithms and code meant for the same purpose. It can be shown that for a large class of basis functions $\varphi(.)$ the matrix φ is indeed non- singular (and hence invertible) providing the data points are distinct. Once the weights are determined, we have a function $f(\mathbf{x})$ that represents a continuous differentiable surface that passes exactly through each data point.

Commonly Used Radial Basis Functions: A range of theoretical and empirical studies have indicated that many properties of the interpolating function are relatively insensitive to the precise form of the basis functions. Some of the most commonly used basis functions are:

1. GAUSSIAN FUNCTION:

$$N(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \qquad \text{Width parameter } \sigma > 0$$

(5)

2. MULTI-QUADRIC FUNCTION:

$$N(r) = (r^2 + \sigma^2)^{\frac{1}{2}} \qquad \text{Parameter } \sigma > 0$$

(6)

3. GENERALIZED MULTI-QUADRIC FUNCTION:

$$N(r) = (r^2 + \sigma^2)^{\beta} \qquad \text{Parameters } \sigma > 0, 1 > \beta > 0$$

(7)

4. INVERSE MULTI-QUADRIC FUNCTION:

$$N(r) = (r^2 + \sigma^2)^{-\frac{1}{2}} \qquad \text{Parameters } \sigma > 0$$

(8)

Note that in the present work, Gaussian Function is used as the Radial basis Function to train the neural network.

Properties of the Radial Basis Functions: The Gaussian and Inverse Multi-Quadric Functions are localized in the sense that $N(r) \to 0$ as $|r| \to \infty$ 1 (9)

but this is not strictly necessary.

$$N(r) \to \infty \qquad \text{as} \qquad |r| \to \infty \qquad (10)$$

Note that even the Linear Function $N(r) = r = |x - x^p|$ is still non-linear in the components of **x**. In one dimension, this leads to a piecewise-linear interpolating function which performs the simplest form of exact interpolation. It is a well known fact that for neural network mappings, there are good reasons for preferring localized basis functions.

In order to train the network, we establish the relationship between the coefficients of the 3_level DWT as the input and the quantized value of LL3 band coefficients as the target for RBF neural network model.

III. EXPERIMENTAL DETAILS

The entire experimental work is classified into three sub sections. In the first sub section, the motion vector estimation based frame selection criterion is presented in the form of an algorithm. In the second sub section, watermark embedding in the selected blue channels of the video frames is described in detail along with its block diagram. Watermark extraction is given in the third subsection. After having recovered the watermarks from the signed frames, these are subject to comparison with the original ones. For this purpose, two different metrics – Normalized Correlation (NC(W,W')) and Bit Error Rate (BER) are applied. The visual quality of the signed frames is evaluated by PSNR.

A. Selection of video frames using motion vector estimation
While embedding the watermark into the uncompressed video, it is observed that to embed the watermark in the entire video is a bit costly in time. Therefore, a frame selection criterion is applied on the basis of motion vector estimation. Thus, we decide to apply a selection criterion on frames so that not more 10% of the total frames of the given video are actually watermarked. The selection criteria are based on motion vector estimation. Listing 1 stipulates the steps used for this selection criterion.

Listing 1: Algorithm for Frame Selection

1) The motion vectors between two frames are computed by block matching approach in that full search method is utilized
2) As a result of step 1, each frame is divided into blocks of 8*8 due to which the range of motion lies between 0-98. Therefore, to determine the number of locations of maximum motion, set the threshold(th1) equal to 75
3) After executing step 2, the number of locations having motion greater than 75 are determined. After this we set another threshold (th2) equal to 50 which will provide the combinations of frames fulfilling the set threshold (th2) criteria.
4) Since we have to select only 10% or less number of frames for watermark processing we set the criteria on the frames selected in step 3 such that we will select greatest frame out of all combinations and if number of frames are still greater, then every 2nd frame, 3rd frame or so on…is selected so as to get 10% of the total number of frames.
5) At the end, a total of 10% or less number of frames are finally selected for watermark embedding after computing step 4. The selection of threshold may depend upon the nature and type of video in question

B. Watermark Embedding
The watermark is a 32*32 binary image. Listing 2 gives algorithm for watermark embedding.

Listing 2: algorithm for watermarking embedding
1) Divide the video into frames for watermark embedding according to frame selection criteria given in section 3.1
2) Decompose all frames into its R G and B channels
3) Consider only blue channels and apply 3-level DWT on selected frames and obtain there LL3 sub-band coefficients
4) Use the LL3 sub-band coefficients as input and the quantized value of LL3 sub-band as the target to RBF neural network with Q=16. Hence, train the RBF neural network (RBFNN). The RBF neural network (RBFNN) produces a row vector of size 1*1024 as its output
5) Perform watermark embedding using formula given in Eqn. 11

$$I'(1,i) = RBF\left(\text{Round}\frac{C(1,i)}{Q}\right) + W(1,i)$$

(11)

where W= Watermark
C = DWT coefficient of LL3 Band
I' = watermarked blue channel coefficients

6) Compute Inverse Discrete Wavelet Transform (IDWT) of watermark blue channel (I') to transform the watermarked frame into spatial domain
7) Concatenate the three channels – Red, Green and modified Blue obtained in Listing 2 to obtain the watermarked frame

The visual quality of signed frames of the video is quantified by computing average PSNR of the given video. The average PSNR is calculated using formulation given in Eqn. 12.

$$AVG_{PSNR} = \frac{\sum_{i=1}^{T} PSNR}{T}$$

(12)

Where   T = no. of frames for which average PSNR has to

be calculated

This equation uses the quantity PSNR which is primarily a full reference metric to access the visual quality of an individual frame / image. The PSNR is given by Eqn. 13.

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

(13)

Which makes use of another quantity – Mean square error (MSE) of the frame in question. The MSE is given by Eqn. 14.

$$MSE = \frac{1}{m*n} \sum_{i=1}^{m} \sum_{j=1}^{n}(I - I')^2$$

(14)

Where   I = original image
I' = Watermarked image
m * n = image size
MSE = Mean square error

For better visual quality the Avg. PSNR should be as large as possible.

C. Watermark Extraction

It is highly unlikely that the watermark which is embedded into the video is same as the recovered one. An assessment of the similarity between the embedded and extracted watermark is done by computing two different coefficients- Normalized correlation or NC (W, W') and Bit Error Rate or BER(W, W'). Listing 3 gives watermark extraction algorithm.

Listing 3: Algorithm for watermarking extraction
1) Divide the watermarked frames obtained in section 3.2 into R, G and B channels
2) Transform the blue channel of watermarked frame obtained in step 1 using 3-level DWT and obtain its LL3 sub-band coefficients
3) Train the network with the LL3 sub-band coefficients (C') obtained in step 2, as the input to the RBF neural network and Quantized value of LL3 sub-band coefficient as target to the RBF neural network. Use Q = 16 in the present case. Extract the watermark according to the formula as given in Eqn. 15

$$W' = RBF\left(round \frac{C'(1, i)}{Q}\right) - T(1, i)$$

(15)

Where   W' = Extracted watermark
C' = LL3 sub-band coefficient
T= Target

For better extraction, the Normalized Correlation parameter should be as large as possible. On the contrary, a large value of BER indicates poor watermark recovery. The mathematical formulation of NC and BER are given in Eqns. 16 and 17 respectively [6].

$$NC(W, W') = \frac{\sum_{i=1}^{x} \sum_{j=1}^{y}[W(i, j) * W'(i, j)]}{\sum_{i=1}^{x} \sum_{j=1}^{y}[W(i, j)]^2}$$

(16)

$$BER(W, W') = \frac{1}{xy} \sum_{j=1}^{y} |W'(j) - w(j)|$$

(17)

Where

W = Original Watermark
W' = Extracted Watermark
x * y = Size of the watermark

IV.   RESULTS AND DISCUSSIONS

In this paper, we consider three uncompressed colored AVI video to examine the embedding, extraction and robustness modules of the proposed watermarking scheme. The selected video are: (1) Car of frame size 480*360, frame rate of 29 fps, (2) Foreman and (3) News both of frame size 352*288 with the frame rate of 30 fps. Each video contains 100 frames. The watermark is a normalized image of size 32*32. These items are shown in Fig. 1 (a-d) respectively.
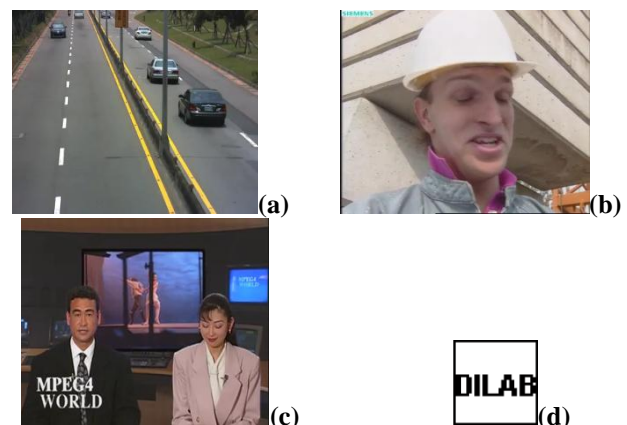
Fig. 1(a-d): Central frame for (a) Car, (b) Foreman and (c) News respectively and (**d**) the original watermark

PSNR$_{car}$ = 38.9693 db db   **PSNR$_{foreman}$ = 43.9095 db**

 (a)   **(b)**
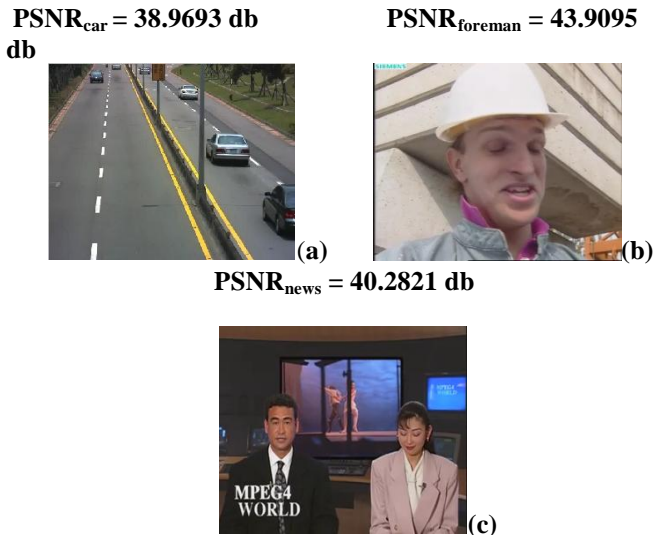
PSNR$_{news}$ = 40.2821 db

 (c)

Fig. 2(a-c): Central watermarked video frames of (a) Car, (b) Foreman and (c) News

The average PSNR for the selected frames of CAR is **38.9693** db, FOREMAN is **43.9095** db and NEWS is **40.2821 db**. These values are also mentioned on top of their respective frames in Fig. 2. Fig. 3 (a-c) depicts extracted watermarks out of the signed video sequences. Their computed NC and BER values are also mentioned besides them.

 (a)   **(b)**   **(c)**

NC = 0.9964   NC = 1   NC = 0.9743

BER = 0.0059   BER = 0   BER = 0.0211
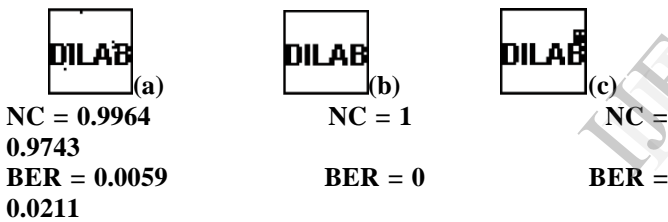
Fig. 3(a-c): Extracted watermarks obtained from (a) Car, (b) Foreman and (c) News

Table 1 compiles the selected video frames after implementing the selection criteria in section 3.1.

**Table 1:** Selected video frames of Car, Foreman and News

| Video	Selected frames
Car	4 , 12 , 19 , 27 , 34 , 42 , 49 , 59
Foreman	3 , 5 , 15 , 72 , 78 , 80 , 91 , 93 , 95
News	2 , 44 , 58 , 70 , 75 , 82 , 88 , 92 , 96 , 100

A. Robustness studies

To examine the issue of robustness of the proposed watermarking scheme, three different video processing attacks are carried out over the signed video frames. These attacks are **scaling** (Scaling % = 20, 40, 60, 80 and 100), **Gaussian noise** (noise variance = 0.001, 0.002, 0.004 and

0.008) and **JPEG compression** (QF = 95, 90, 85, 80 and 75).

Fig. 4 (a) depicts the plot between **scaling (%)** and the computed PSNR values. It is clear that with increase in the scaling % the PSNR increases for all three video sequences. Fig. 4 (b) depicts plot of normalized correlation with scaling %. This quantity also increases with the increase in the scaling % for all three video sequences. As Bit Error Rate or BER is usually found to behave inverse of NC(W,W'), Fig. 4(c) depicts the expected behavior and is decreasing with the increase in scaling % for all three video sequences. Agarwal et. al.[6] have also reported a similar behavior in case of scaling attack..
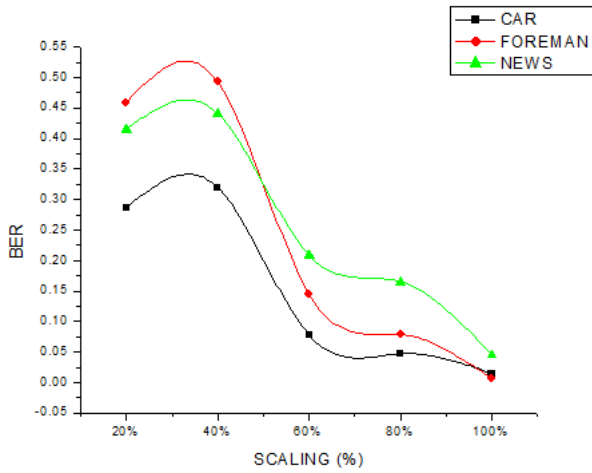

Plot (a)

Plot (b)

Plot (c)
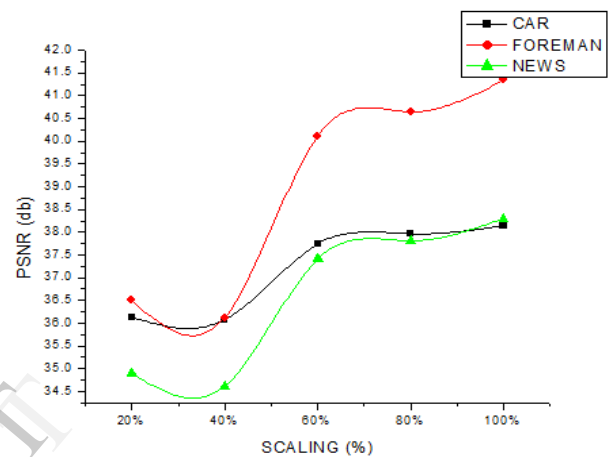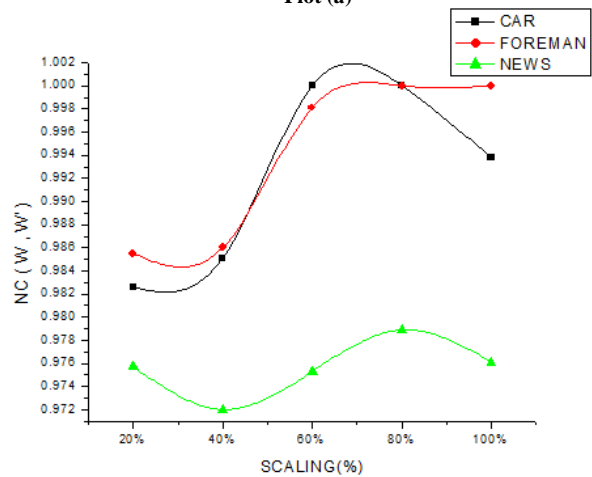
Fig. 4: Plot of (a) PSNR, (b) NC and (c) BER w.r.t Scaling (%)

Plot (c)

Fig. 5: Plots of (a) PSNR, (b) NC and (c) BER w. r. t Gaussian noise variance

Large values of PSNR and NC(W,W') clearly indicate that the proposed watermarking scheme resists the scaling attack carried out over the signed frames. It is found to exhibit good visual quality as well as successful watermark recovery even after implementing large scaling ratio.
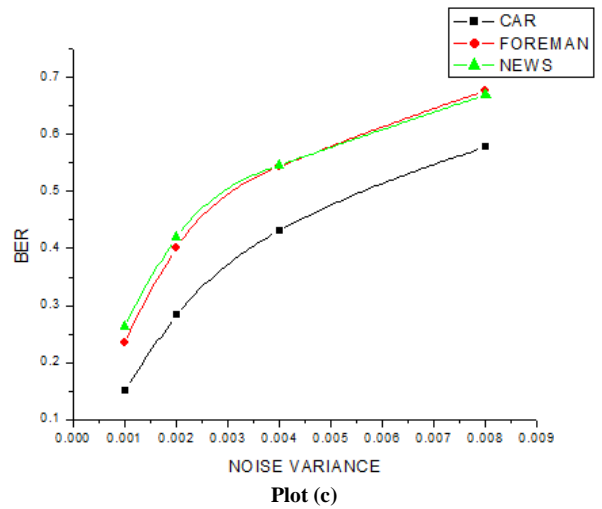
Plots shown in Fig. 5 clearly indicate that both PSNR and NC(W,W') substantially decrease as a result of increase in the **Gaussian noise variance**. As a result, the BER increases with increase in the variance which indicates successful watermark recovery after carrying out Gaussian noise attack. Note that this behavior is observed for all three videos sequences listed in the present work.
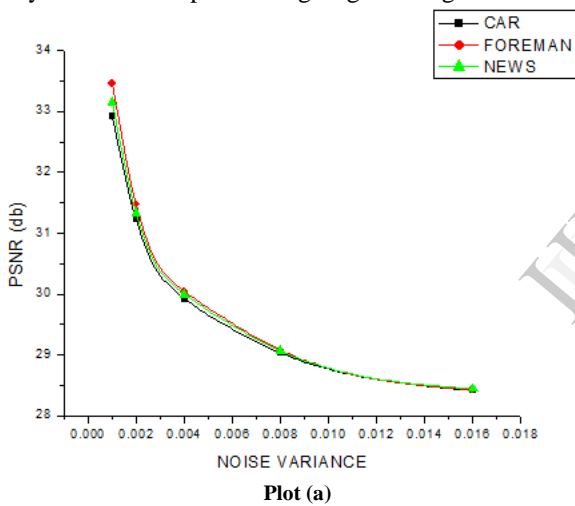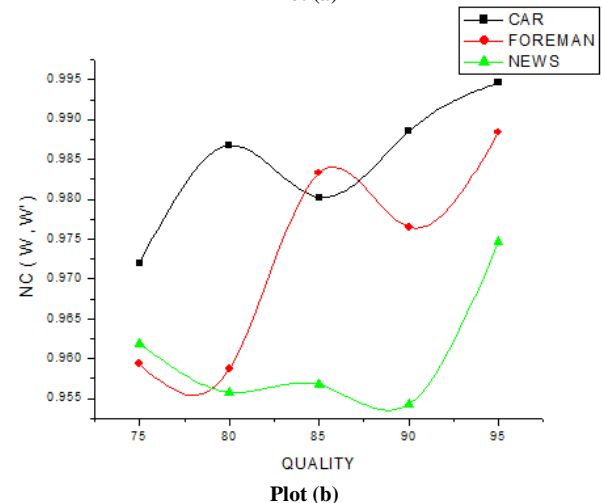

Plot (a)

Plot (a)
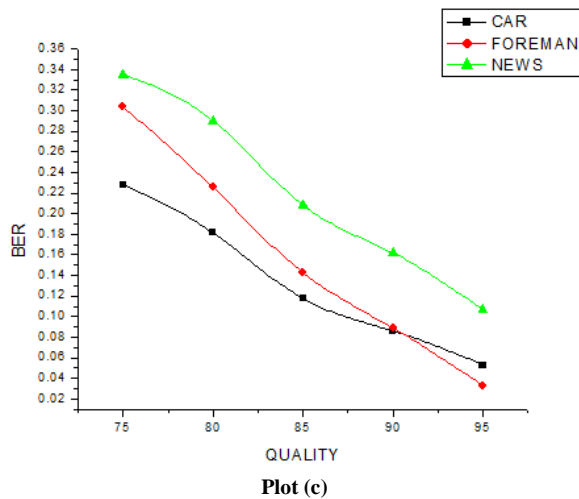

Plot (b)

Plot (b)

Plot (c)
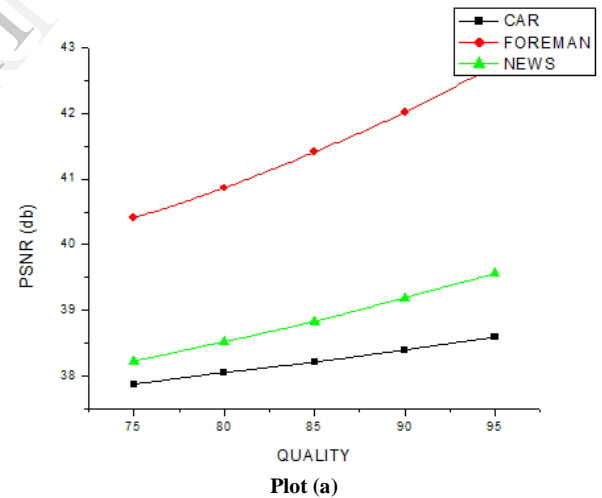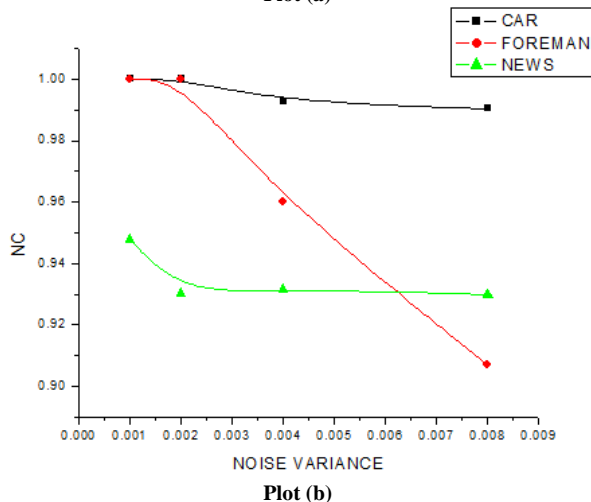
Fig. 6: Plots of (a) PSNR, (b) NC and (c) BER w. r. t Quality Factor (QF) of JPEG Compression

Fig. 6 shows the plots of PSNR, NC and BER with respect to **JPEG** Quality factor. The PSNR and NC(W,W') increases with the increase in the quality factor. On the other hand, the BER shows a decrease with the increase in quality factor.

Table 2 depicts the average embedding and extraction time spans for selected frames in the three video sequences. Note that the embedding time (sec's) includes the training of the RBF neural network. The average time per frame for embedding and extraction comes out to be 24.96 sec, 21.62 sec for Car, 23.71 sec, 22.43 sec for Foreman and 23.33 sec, 22.37 sec for News.

Table 2: Avg. Embedding and Extraction time per frame for selected video frames

video / Time(sec)	CAR	FOREMAN	NEWS
Embedding Time	24.96	23.71	23.33
Extraction Time	21.62	22.43	22.37
Total Time	46.58	46.14	45.70

This indicates that the selected frames of the given video are successfully watermarked within a time span ranging between seconds to minutes for the entire video. This makes the proposed motion vector and RBF neural network based video watermarking scheme a suitable candidate for developing real time video watermarking applications...

CONCLUSION

In this paper, we successfully demonstrate a novel uncompressed RGB, AVI format digital video watermarking scheme based on 3-level DWT and Radial basis Function (RBF) neural network by computing its motion vectors. All the operation is carried on the blue channel in order to retain the color information. The results obtained after embedding gives higher PSNR values for three video sequences from which good visual quality of watermarked frame can be depicted. The extracted watermark results in higher NC(W,W') and lower BER which indicate the good quality of extracted watermark. Robustness studies show that the scheme is highly robust against the considered video processing attacks. From the experimental results we conclude that the proposed watermarking scheme is suitable for developing real time video watermarking applications.

REFERENCES

[1] Munesh Chandra, shikha Pandey, "A DWT Domain Visible Watermarking Techniques for Digital Images", 2010 International Conference on Electronics and Information Engineering.

[2] Cheng-Ri Piao, Seunghwa Beack, Dong-Min Woo, and Seung-Soo Han, "A Blind Watermarking Algorithm Based on HVS and RBF Neural Network for Digital Image," l.Jiao et al.(Eds.):LNCS 4221 , pp.493-496,2006

[3] Xuefang Li,Rangding Wang, "AVideo Watermarking Scheme based on 3-DWT and Neural Network" ,2007 IEEE.

[4] Mahar EL'ARBI,Chokri BEN AMAR,Henri NICOLAS, "Video Watermarking Based on neural network",IEEE-2006.

[5] Nallagarla Ramamurthy and S. Varadarajan, "The Robust Digital Image watermarking scheme with back Propogation Neural Network In DWT Domain", IJCSNS International Journal of computer Science and Network Security,Vol.13 No.1,January 2013

[6] Charu Agarwal ,Anurag Mishra, , arpita Sharma, Chetty Girija ,"A novel Scene Based Robust Video Watermarking Scheme in DWT Domain Using Extreme Learning Machine" Springer International publishing Switzerland, Volume 16, 2014, pp 209-225.

[7] Chetan K, R,Raghanedra K, "DWT Based Blind Digital Video Watermarking scheme for Video Authentication" International Journal of Computer Applications(0975-8887)Volume 4-No.10,Auguest 2010.

[8] Aroh Barjatya,,student Member,IEEE "Block Matching Algorithms for Motion estimation" DIP 6620 Spring 2004 Final Year Project

[9] Jun Zhang ,JeGu Li and Ling Zhang, " Video watermarking technique in motion vector", 2001 IEEE

[10] S.chen , C.F.N.Cowan , And P.M.Grant "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks" , IEEE Transection on Neural Network,Vol.2,no.2, march 1991

[11] T.Jayamalar,Dr.V.Radha, "Survey on Digital Video Watermarking Technique and Attacks on Watermarking," International Journal of Engineering Science and Technology,Vol.2(12),2010,6963-6967.

[12] John A. Bullinaria, 2012, "Radial Basis Function Networks: Introduction", neural computation: Lecture 13 http://www.cs.bham.ac.uk/~jxb/INC/l1.pdf