

# Digital Video Watermarking Based on Quantization Index Modulation

Shilpa P. Metkar, Bhakti D. Kadam

Department of Electronics and Telecommunication, College of Engineering, Pune  
metkars.extc@coep.ac.in, bhaktikadam30@gmail.com

**Abstract-** Nowadays, due to advent in emerging technologies and networking services, it has become extremely important to protect and authenticate the digital data (images, audio and video) which is used for processing and distribution. A watermarking technique based on 3-D Discrete Cosine Transform (DCT) and Quantization Index Modulation (QIM) for video security, authentication and copyright protection is discussed in this paper. Video watermarking embeds a permanent message signal in a video sequence in order to protect the video from illegal copying. The presented watermarking algorithm works in frequency domain utilizing the concept of QIM. The binary watermark is transformed into its bit planes. The first four Most Significant Bit planes are extracted and embedded into video frames. Blind watermark extraction is adopted using Inverse QIM and Minimum Distance Decoder. The proposed algorithm is evaluated for robustness and various attacks like JPEG compression, noising, luminance modification, filtering and rotation.

**Keywords-** Blind Watermark Extraction, Digital Watermarking, Minimum Distance Decoder, Discrete Cosine Transform (DCT), Quantization Index Modulation (QIM)

## I. INTRODUCTION

Digital video is one of the popular multimedia data exchanged on the internet nowadays. Due to the rapid advent of networking services and internet, digital videos have become easily accessible and replicable. New emerging technologies are capable of producing illegal copies of video and making them available to people. Hence, it has become very necessary for creators and owners of the video to protect their copyright. Also, digital videos are very susceptible to attacks like frame dropping, frame averaging, filtering, noising, etc. In this context, digital watermarking provides most efficient solution to protect and authenticate any form of digital data and to prevent its unauthorized access. Watermarking protects digital data by embedding a permanent message signal i.e. watermarks into the host signal.

Watermarks can be classified as visible and invisible. A visible watermark is a visible translucent image laid on the primary image. It holds right on the primary image allowing primary image to be viewed. Invisible watermark is overlaid image which cannot be seen but which can be algorithmically detected. Watermarking systems are of two types- blind and non-blind. Blind watermarking systems do not require original data at the detector side while non-blind systems require original data at the receiver. Here, we have proposed robust and blind video watermarking system.

This paper is organized as follows. Section II briefs about literature survey for watermarking techniques. Section III explains watermark embedding and extraction process. Experimental results are discussed in Section IV. Section V concludes the paper.

## II. LITERATURE SURVEY

A number of watermarking techniques are proposed in literature, which exploit different ways of embedding watermarks. These techniques can be classified as spatial domain and frequency domain [1]. The spatial domain watermarking techniques embed the watermark by modifying the pixel values of the host video frames directly. In transform domain technique, the watermark is inserted by altering the transformed domain coefficients. The inverse transform is finally applied to get watermarked domain. Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Discrete Frequency Transform (DFT) are the three methods of transform [2].

I. J. Cox *et al* presented a secure algorithm in which a watermark is constructed as an independent and identically distributed Gaussian random vector that is imperceptibly inserted in a spread-spectrum-like fashion into the perceptually most significant spectral components of the data [3]. C. T. Hsu and J. L. Wu proposed an image authentication technique by embedding digital "watermarks" into images. In this approach, watermark is embedded with visually recognizable patterns into the images by selectively modifying the middle-frequency parts of the image [4]. M. A. Suhail and M. S. Obaidat proposed a watermarking algorithm based on discrete cosine transform (DCT) and image segmentation. The image is first segmented in different portions based on the Voronoi diagram and feature extraction points. Then, a pseudorandom sequence of real numbers is embedded in the DCT domain of each image segment [5]. B. Chen and G. W. Warnell has considered the problem of embedding one signal (e.g. a digital watermark), within another "host" signal to form a third, "composite" signal. The embedding is designed to achieve efficient tradeoffs among the three conflicting goals of maximizing information-embedding rate, minimizing distortion between the host signal and composite signal, and maximizing the robustness of the embedding. They introduced new classes of embedding methods, termed quantization index modulation (QIM) and distortion-compensated QIM (DC-QIM) and convenient realizations in the form of dither modulation. QIM is "provably good" against arbitrary bounded and fully informed attacks, and achieves provably better rate distortion-robustness tradeoffs than spread spectrum and low-bit(s) modulation methods [6].

## III. PROPOSED METHOD

The proposed method works in uncompressed as well as compressed domain. The system block diagram is as shown in Fig. 1. The algorithm can be divided into three steps- Watermark Bit Stream Formation, Watermark Embedding and Watermark Extraction.

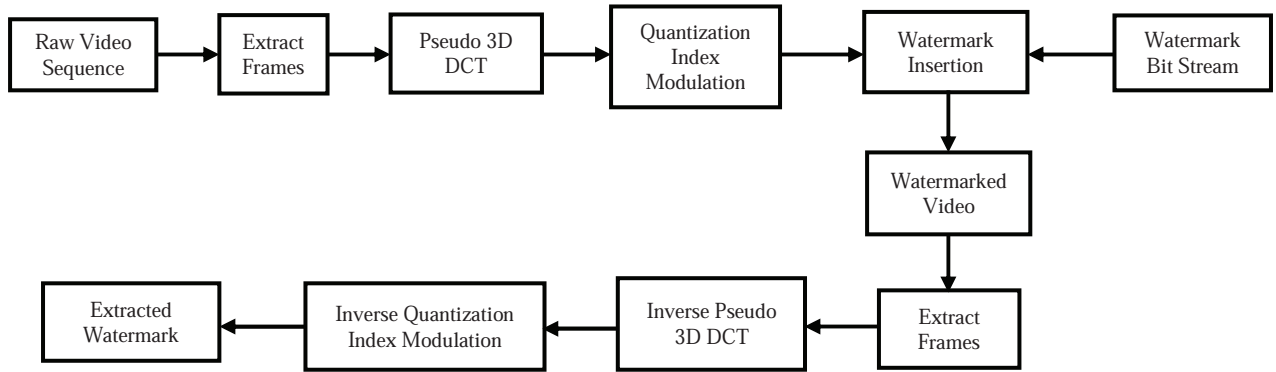


Fig. 1. System Block Diagram

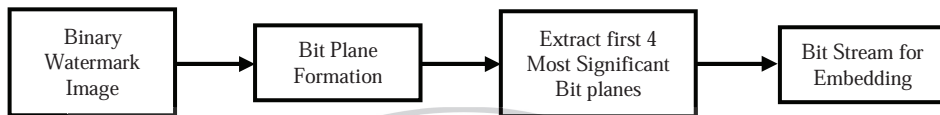


Fig. 2. Watermark Bit Stream Formation

**A. Watermark Bit Stream Formation:**

The original binary watermark image is transformed into its eight bit planes using bit plane slicing technique as each image pixel can be represented using 8 bits. The higher-order bits (especially the top four) contain the majority of the visually significant data. The other bit planes contribute to more subtle details in the image. Then, the first four most significant bit planes are separated and bit stream of '1's and '0's is formed in order to embed the watermark into host video frames. The process of watermark bit stream formation is as shown in Fig. 2.

**B. Watermark Embedding:**

**I. Pre-processing (Pseudo 3-D DCT)**

In the embedding process, first we take several successive raw frames as a group. In our approach, we take four frames as a group (GOP). Each frame in the group is divided into 8x8 sized blocks. These blocks are transformed into DCT domain using 2-D DCT. Next, the DC value of each block located in the same position of successive frames for a group is transformed into DCT domain using 1-D DCT. This process is

known as pseudo 3-D DCT which reduces computation complexity [7]. After the transforming process, we obtain a new DC value and three AC values. The sum of all absolute AC values with weight is obtained using (1),

$$Sum(i, k) = \sum_{l=1}^3 Ws(i, k, l) |AC(i, k, l)| \quad (1)$$

where,  $Sum(i, k)$ ,  $Ws(i, k, l)$  and  $AC(i, k, l)$  denote the sum of all AC values, the corresponding weight value, and the  $l$ th AC value corresponding to the  $k$ th block of successive frames within the  $i$ th group, respectively. By repeating the process for all the blocks, we get a sequence of sums for every block.

**II. Threshold Estimation:**

After computing  $Sum(i, k)$ , we calculated the threshold  $T(i)$  which is dependent on characteristics of video and number of bits to be embedded. The formula used for calculating the threshold is derived from the concept of probability distribution. Steps for the calculation of threshold are:

1. Calculate the Mean and Standard Deviation of  $Sum(i, k)$ .
2. Apply the formula of Gaussian kernel to  $Sum(i, k)$  as in (2),

$$Sumf(i, k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(sum(i,k)-\mu)^2}{2\sigma^2}} \quad (2)$$

where,  $\mu$ =Mean  
 $\sigma$ =Standard Deviation

3. Calculate the probability of every  $Sumf(i, k)$  value.
4. Calculate the threshold probability.

The threshold probability is calculated for each group considering the number of bits to be embedded in the group. The  $Probth(i)$  is calculated as per (3),

$$Probth(i) = 1 - \frac{\text{Number of bits to be embedded in a GOP}}{\text{Number of blocks in a frame}} \quad (3)$$

5. Calculate value of R using (4),

$$Probth(i) \geq \sum_{t=0}^R Prob(t, i) \quad (4)$$

where,  $Prob(t, i)$  denotes the probability value of  $Sumf(i, k)$ .

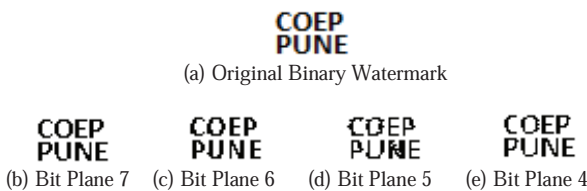


Fig. 3. Watermark 1 and Extracted Bit Planes

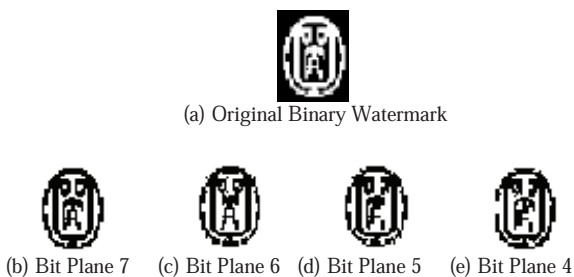


Fig. 4. Watermark 2 and Extracted Bit Planes

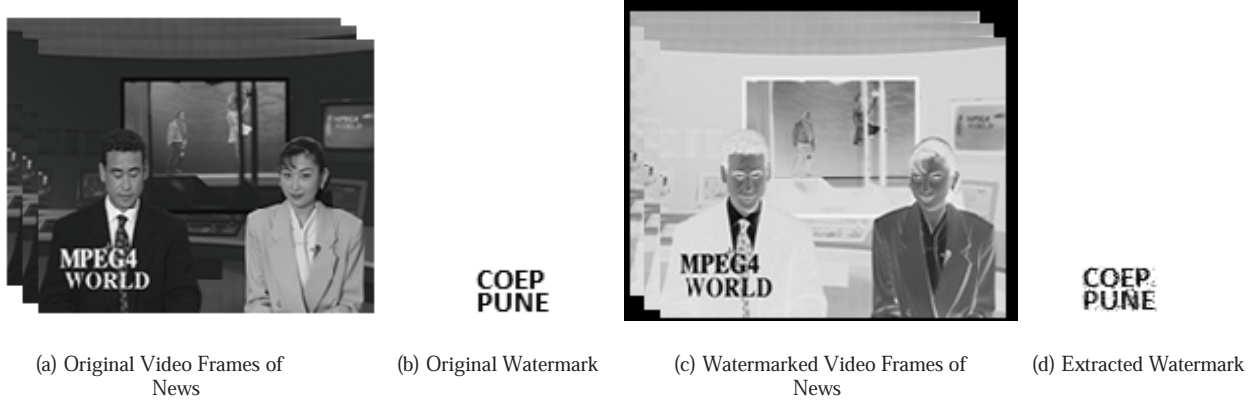


Fig. 5. Result of Proposed Algorithm in Uncompressed Domain



Fig. 6. Result of Proposed Algorithm in Compressed Domain

6. Calculate the threshold  $T(i)$  using (5),

$$T(i) = R + \frac{512}{R \times \log_{10}\left(\frac{512}{R}\right) \times \log_{10}(\text{Maxsum}(i))} \quad (5)$$

where,  $\text{Maxsum}(i)$  is the maximum sum value in the  $\text{Sum}(i, k)$ .

### III. Quantization Index Modulation:

Once the threshold is calculated, we calculated the quotient  $Q(i, k)$  with the help of (6) and (7),

$$t = \frac{\text{Number of Blocks in a frame}}{\text{Number of bits to be embedded in a GOP}} \quad (6)$$

$$Q(i, k) = \text{round}\left(\frac{\text{Sum}(i, k)}{t}\right) \quad (7)$$

Next, utilize quantization index modulation to embed watermark bits [6], [8]. Here, we are using a simple case of QIM for embedding only one bit information using two quantizers  $Q_0$  and  $Q_1$ . The embedded value determines the selection of quantizers with step size  $\Delta$ . We have used the calculated threshold as  $\Delta$ . The host signal is quantized using (8), (9) and (10).

$$Q_i(s) = Q(s - d_i) + d_i, \quad i = 0, 1 \quad (8)$$

where,  $Q(s) = \Delta \left\lfloor \frac{s}{\Delta} \right\rfloor$ ,  $d_0 = -(\Delta/4)$ ,  $d_1 = (\Delta/4)$

If the watermark bit is 0,

$$Q_0(s) = Q(s + \Delta/4) - \Delta/4 = \Delta \left\lfloor \frac{s + \Delta/4}{\Delta} \right\rfloor - \Delta/4 \quad (9)$$

If the watermark bit is 1,

$$Q_1(s) = Q(s - \Delta/4) + \Delta/4 = \Delta \left\lfloor \frac{s - \Delta/4}{\Delta} \right\rfloor + \Delta/4 \quad (10)$$

where,  $\lfloor \cdot \rfloor$  denotes rounding to the nearest integer.

Modify  $Q(i, k)$  to  $\text{mod}Q(i, k)$  using QIM. According to  $Q(i, k)$  and watermark bit to be embedded, we modified  $\text{Sum}(i, k)$  to  $\text{mod}S(i, k)$  as per (11) and (12),

$$\begin{aligned} \text{Sum}(i, k) &= \text{Sum}(i, k) \quad \text{when } Q(i, k) \text{ is even and } W_m = 0 \\ &= \text{Sum}(i, k) \quad \text{when } Q(i, k) \text{ is odd and } W_m = 1 \\ \text{Sum}(i, k) &= \text{Sum}'(i, k) \quad \text{when } Q(i, k) \text{ is even and } W_m = 1 \\ &= \text{Sum}'(i, k) \quad \text{when } Q(i, k) \text{ is odd and } W_m = 0 \end{aligned} \quad (11)$$

Applying weights to modulate  $\text{Sum}(i, k)$ ,

$$\text{mod}S(i, k) = \sum_l W_s(i, k, l) |AC(i, k, l)| + W_e(i, k, l) D(i, k) \quad (12)$$

where,  $D(i, k)$  denotes the difference between  $\text{mod}Q(i, k)$  and  $Q(i, k)$ .  $AC(i, k, l)$  and  $W_e(i, k, l)$  denote the  $l$ th AC value of the  $k$ th block in the frames of  $i$ th group and the corresponding embedding weight respectively. By repeating the above procedure until all watermark bits are embedded, the embedding process is completed.

### C. Watermark Extraction:

The extraction process is inverse of the embedding process. The watermarked video sequence is divided into several groups consisting of four frames in a group. Apply pseudo 3-D DCT to obtain a new DC and three AC values as done while embedding. Calculate the sum of AC values as  $\text{Sumex}(i, k)$ . Then, we computed the quotient  $Q_{ex}(i, k)$  using (13) as,

$$Q_{ex}(i, k) = \frac{\text{Sumex}(i, k)}{t} \quad (13)$$

After computing  $Q_{ex}(i, k)$ , watermark bits are extracted using inverse quantization index modulation. The detector used for QIM is Minimum Distance Detector. The signal detected is given by (14),

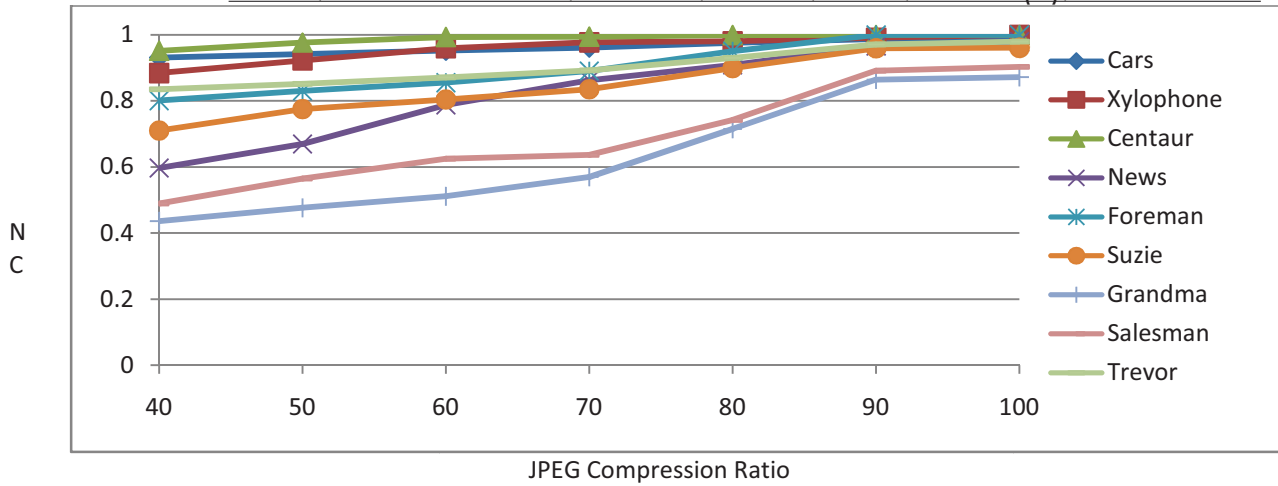


Fig. 7. Results for JPEG Compression

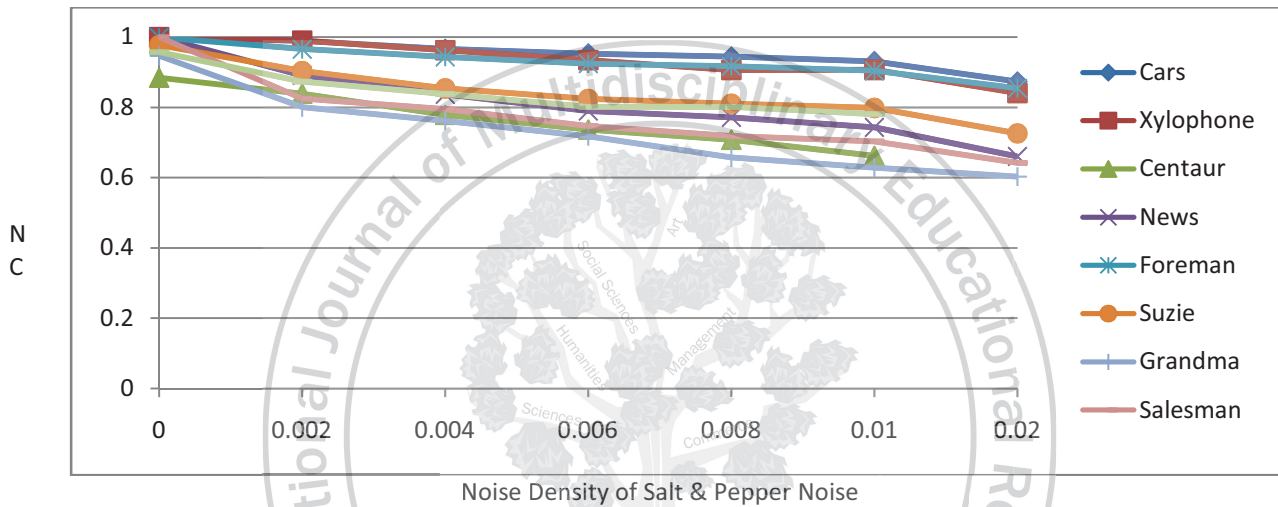


Fig. 8. Results for Salt & Pepper Noise attack

$$y' = y + n = (x + m) + n \quad (14)$$

The decoder needs to choose the nearest reconstruction point to the detected signal  $y'$ . This is implemented according to minimum Euclidean distance rule. The decoded message bit is defined using (15),

$$\hat{m} = \underset{0,1}{\operatorname{argmin}} \|y' - Qm(y')\| \quad (15)$$

By repeating the above procedure until all the bits are extracted, the extraction process is completed.

The extracted watermark bit stream gives the first four bit planes of binary watermark. We have to recombine them with the remaining four Least Significant Bit planes to reconstruct the watermark completely.

#### IV. RESULTS

##### A. Experimental Setup:

We have used six videos (frame size  $176 \times 144$ ) of QCIF type in uncompressed domain and three videos (frame size  $320 \times 240$ ) in compressed domain for evaluating the performance. The 32 frames of each video for uncompressed domain and 16 frames of each video for compressed domain are used for embedding and extraction of watermark.

The binary watermarks used are of size  $36 \times 22$  and  $30 \times 40$  for uncompressed and compressed domain testing. The original binary watermarks (watermark 1 and 2) and extracted bit planes are shown in Fig. 3 and Fig. 4 where (a) shows

original binary watermark and (b), (c), (d), (e) are extracted bit planes 7, 6, 5 and 4 respectively.

##### B. Parameter Setting:

In (1), the weights  $Ws(i, k, l), l = 1, 2, 3$  are set to 1. The embedding weights  $We(i, k, l), l = 1, 2, 3$  in (12) are set to  $1/6, 1/3, 1/2$  respectively. We have embedded 396 bits in each group in uncompressed domain and 1200 bits in compressed domain.

##### C. Performance Measure:

For measuring the performance, quality of watermarked frames and extracted watermark, we have used the following measures:

The PSNR and MSE judge the quality of watermarked video frames as compared to original video frames. The PSNR and MSE are calculated using (16) and (17),

$$PSNR = 10 \log \frac{255 \times 255}{MSE} \quad (16)$$

$$MSE = \frac{\sum_{H,W} |f(i,j) - g(i,j)|^2}{H \times W} \quad (17)$$

where,  $f(i, j)$  and  $g(i, j)$  denote the original frame and watermarked frame.  $H$  and  $W$  denote the height and width of frames.

For measuring the robustness of extracted watermark, we have used normalized correlation given by (18),

$$NC = \frac{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} W(i,j)\widehat{W}(i,j)}{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} [W(i,j)]^2} \quad (18)$$















where,  $W(i,j)$  denotes the original watermark and  $\widehat{W}(i,j)$  denotes the extracted watermark.

D. Experimental Results:

TABLE I Performance of Proposed Algorithm

Video Name	Execution Time (seconds)	Average PSNR of original frames and watermarked frames (dB)	Average MSE of original frames and watermarked frames	Normalized Correlation
News	9.8692	39.4824	10.9713	0.9540
Foreman	10.3239	35.1123	29.0377	1.00
Salesman	9.9398	43.8452	4.1361	0.9047
Suzie	9.7809	44.6283	2.8485	0.9493
Grandma	10.0473	49.4568	0.9758	0.8809
Trevor	9.9880	36.4466	28.5314	0.9211
Cars	14.9532	35.3601	23.3747	0.9935
Xylophone	15.9251	34.3666	31.7358	1.00
Centaur	15.0993	49.3053	0.8936	1.00

TABLE II Performance of Proposed Algorithm for various attacks

Type of Attack	Uncompressed Domain		Compressed Domain	
	Normalized Correlation of Extracted Watermark	Extracted Binary Watermark	Normalized Correlation of Extracted Watermark	Extracted Binary Watermark
Luminance Modification (Brightening)	0.9421		0.9639	
Luminance Modification (Darkening)	0.9542		0.9935	
Median Filtering	0.9613		1.00	
Wiener Filtering	0.9752		0.9867	
Rotation and Scaling	0.9621		0.9742	
Average Filtering	0.9814		0.9865	
Gaussian Filtering	0.9689		1.00	

We have implemented the proposed algorithm on MATLAB version R2012a.

Fig. 5 and Fig. 6 shows the result of proposed algorithm in uncompressed and compressed domains where (a) and (b) shows original video frames and original watermark respectively, (c) is watermarked video frames; (d) is extracted watermark. The basic performance of the algorithm is shown in Table I indicating the watermark embedding and extraction time, average PSNR and MSE of frames and Normalized Correlation for extracted watermark. It is seen that average PSNR and NC values are well acceptable. The proposed method works satisfactorily for uncompressed as well as compressed domain videos with the exception that the execution time required for compressed domain videos is more due to the applied compression schemes like JPEG, MPEG to the frames. In the proposed method, we have used probabilistic method for threshold estimation which works acceptably well.

The quality of extracted watermark also depends on the number of watermark bits embedded in each group. If we embed less number of bits in each group, the quality can be enhanced further. In that case, if we embed first two MSB planes in the video frames, the quality of extracted watermark improves. The quality of extracted watermark can also be improved by reducing the block size. If the block size is reduced to 4x4, the watermark can be extracted more efficiently but the time required for execution of algorithm increases considerably.

To evaluate the performance of proposed method against robustness, we have used attacks like luminance modification i.e. brightening and darkening of pixels, filtering, noising, rotation and scaling (Table II). We also have tested the performance against the JPEG compression attack. Fig. 7 shows the normalized correlations of extracted watermarks for JPEG compression ratios ranging from 40 to 100. From this figure we can conclude that the proposed watermarking scheme withstands with the JPEG compression attack. For luminance modification of frames, either brightening or darkening, the quality of extracted watermark is well acceptable.

For filtering attacks like median, wiener and average filtering, watermark is satisfactorily extracted for the mask size of 3x3 and 5x5, but the quality degrades for mask size of 9x9. If gaussian filtering is applied to video frames with neighborhood of 3x3 and sigma value varying from 0.1 to 3, then it is seen that watermark is extracted up to sigma value of 2. Salt & Pepper noise is added to the video frames with intensity varying from 0.001 to 0.05. It can be seen from Fig. 8 that NC values of extracted watermark are good for the intensity of noise up to 0.03 but degrades after that. For all attacks, increasing the intensity of attack significantly affect the extracted watermark reducing its quality drastically. For the attacks like frame averaging, rotation and cropping, the algorithm is vulnerable.

V. CONCLUSION

In this paper, we have proposed a watermarking technique for video authentication based on pseudo 3-D DCT and quantization index modulation. It is found that the proposed algorithm works acceptably well in uncompressed and compressed domains. If robustness of the system is considered, it can withstand with attacks like luminance modification, JPEG compression, filtering and certain amount of noise but fails for geometric attacks like frame cropping, tilting and frame averaging.

Future work aims to improve robustness of system for different attacks like frame averaging, cropping and tilting.

## VI. REFERENCES

- [1] T. Jayamalar and Dr. V. Radha, "Survey on Digital Video Watermarking Techniques and Attacks on Watermarks", *International Journal on Engineering Science and Technology*, vol. 2 (12), pp.6963-6967, 2010.
- [2] C. I. Podilchuk and E. J. Delp, "Digital Watermarking: Algorithms and Applications," *IEEE Signal Processing Magazine*, vol. 18, no. 4, pp. 33-46, July 2001.
- [3] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoan, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Transactions on Image Processing*, vol.6, no.12, pp. 1673-1687, December 1997.
- [4] Chiou-Ting Hsu and J. L. Wu, "DCT based Watermarking for Video," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 1, pp. 206-216, February 1998.
- [5] M. A. Suhail and M. S. Obaidat, "Digital Watermarking-based DCT and JPEG Model," *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 5, pp. 1640-1647, October 2003.
- [6] B. Chen and G. W. Warnell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423-1443, May 2001.
- [7] Hui-Yu Huang, Cheng-Han Yang, and Wen-Hsing Hsu, "A Video Watermarking Technique based on Pseudo 3-D DCT and Quantization Index Modulation," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 625-637, December 2010.
- [8] B. Chen and G. W. Warnell, "Digital Watermarking and Information Embedding using Dither Modulation," *Multimedia Signal Processing, 1998 IEEE Second Workshop*, pp. 273-278, December 1998.

