Discrete Multiwavelet–Based Video Watermarking Scheme Using SURF

Leelavathy Narkedamilly, Venkateswara Prasad Evani, and Srinivas Kumar Samayamantula

This paper proposes a robust, imperceptible block-based digital video watermarking algorithm that makes use of the Speeded Up Robust Feature (SURF) technique. The SURF technique is used to extract the most important features of a video. A discrete multiwavelet transform (DMWT) domain in conjunction with a discrete cosine transform is used for embedding a watermark into feature blocks. The watermark used is a binary image. The proposed algorithm is further improved for robustness by an error-correction code to protect the watermark against bit errors. The same watermark is embedded temporally for every set of frames of an input video to improve the decoded watermark correlation. Extensive experimental results demonstrate that the proposed DMWT domain video watermarking using SURF features is robust against common image processing attacks, motion JPEG2000 compression, frame averaging, and frame swapping attacks. The quality of a watermarked video under the proposed algorithm is high, demonstrating the imperceptibility of an embedded watermark.

Keywords: Discrete multiwavelet transform, Speeded Up Robust Feature, motion JPEG2000 compression, quantization index modulation, watermarking, digital cinema.

Manuscript received Jan. 3, 2014; revised Feb. 10, 2015; accepted Mar. 19, 2015.

Leelavathy Narkedamilly (corresponding author, nleelavathy@gmail.com) and Venkateswara Prasad Evani (profevprasad@yahoo.com) are with the Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Andhra Pradesh, India.

Srinivas Kumar Samayamantula (samay_ssk2@yahoo.com) is with the Department of Electronics and Communication Engineering, Jawaharlal Nehru Technological University, Andhra Pradesh, India.

I. Introduction

Nowadays, the ease with which one can easily reproduce digital media in its exact original form is an important issue and one that is strongly allied to copyright infringement. With recent advances in Internet broadcasting technology, large data files, such as video files, can be transmitted with relative ease and without compromising the quality of the data. Hence, there is a concomitant requirement for the designing of video watermarking algorithms, which are an adopted solution to the problem of copyright infringement and one that is gaining more importance among researchers.

Thus far, the technology that has been used for video watermarking is nothing more than an extension of that used for image watermarking. However, in video, a large amount of redundant data is present between adjacent frames. This redundant data is usually removed while compressing the data in perceptual coding. Moreover, regions in video frames inherently have an imbalance between motion and motionless regions.

Additional care needs to be taken against pirate attacks, such as frame averaging, statistical analysis, code conversions, and so on, when designing video watermarking techniques. Digital video watermarking algorithms can be classified into the following two categories based on cover media: watermarking in compressed video stream, and uncompressed or frame-by-frame video stream. Compressed-domain methods use motion vectors [1]; I-, P-, and B-frames in a group of pictures [2]; and the coefficients of the transform domain in the encoder to embed copyright information.

As there are very few motion vectors and transform coefficients available in a compressed-domain method, the embedding capacity is quite low. Moreover, compressed-

domain methods, such as H.264, MPEG2, and MPEG4, are domain specific and do not retain a watermark if code conversion or transcoding is done on a video file using powerful tools such as video code converters. In comparison, uncompressed-domain or frame-based watermarking methods [2]–[16] are more robust and can withstand compression and common image processing attacks.

Many wavelet-based video watermarking methods have been presented in recent research literatures [4]–[6]. In [5], O.S. Faragallah proposed a discrete wavelet transform (DWT) with singular value decomposition (SVD) video watermarking, but it was not robust enough against scaling, rotation, and cropping attacks. R.O. Preda and N.D. Vizireanu [6] explored spatial localization and an embedded watermark in selected wavelet coefficients. Their algorithm was robust to common image processing attacks but not to cropping or geometrical attacks. In [7], a watermark logo was embedded in the DWT coefficients of every video frame. The DWT coefficients were replaced with the maximum/minimum value of neighboring coefficients. This method was shown to be robust against small geometric attacks and compression; however, the design of a perceptual model is not possible. Reference [8] proposes a semi-blind image watermarking algorithm that embeds information in selected points using Speeded Up Robust Feature (SURF) descriptors and SVD. SURF descriptors are extracted from a watermarked image and used to estimate the quantity or value of changes in the original video due to an attack and recover the copyright information whenever needed for proving ownership without fail; however, they are not suitable for watermarking video as large data has to be processed. A combination of a DWT-based watermarking algorithm and a DCT-based watermarking algorithm outperforms DWT-based watermarking algorithms [9].

Moreover, Xiong and Xiao [10] have proposed multiwavelet-based blind watermarking using just-noticeable difference with texture characteristics. The results show that the process of detection of embedded copyright information was improved by properly locating the embedding position. C. Serdean and others [11] have proved that a multiwavelet transform outperforms a wavelet transform in offering better visual quality at equivalent peak signal-to-noise ratio (PSNR) values, as well as showing that it is superior against attacks such as cropping, scaling, and low-pass filtering. Thus, in this paper, a comparative study is carried out on both the wavelet domain and the multiwavelet domain.

This paper presents a robust, imperceptible high-quality video watermarking method that is based on a discrete multiwavelet transform (DMWT) domain. This method jointly exploits both the spatial and the temporal multi-resolution of a video signal using DMWT. SURF descriptors are used to select a block for embedding one bit of binary watermark, $W_i \in \{0, 1\}$, using quantization index modulation (QIM) [12]. In the design of the proposed robust video watermarking algorithm, to make the algorithm computationally efficient, DMWT is performed on selected blocks of size $b \times b$ rather than on a whole frame, as the computational complexity of DMWT is $O(l \cdot w)$, where l is the length and w is the width of a video frame such that $b < w \le l$. Embedding using QIM in the DMWT domain has given a better performance than recent works. The proposed algorithm is an uncompressed-domain method, where larger-sized coefficients are available for embedding and are robust against common image processing attacks.

Experimental results show that the video watermarking algorithm combined with the DMWT domain and SURF features outperforms those methods found in other related works, in most cases. In particular, the method can sustain motion JPEG2000 (MJP2K) compression and common image processing attacks, giving higher quality video. The remaining part of this paper is organized as follows. The design aspects of the algorithm are discussed in Section II. The proposed video watermarking scheme is discussed in Section III. The experimental results and performance of the proposed scheme are shown in Section IV. Finally, conclusions about the proposed algorithm are given in Section V.

II. Design Aspects of Algorithm

The major applications of video watermarking algorithms demand the following requirements.

1. Invisibility of Watermark

The choice of blocks to embed a bit of a watermark logo plays a vital role to meet the invisibility criteria in high-quality digital videos. DMWT, which is a proven technique, is used for embedding a watermark so as to meet the requirement of invisibility.

2. Payload of Watermark

An important aspect of watermark payload is the granularity of a watermark. Granularity describes the size of the information needed to embed one bit of a watermark into video. This factor determines the number of frames (NoF) required for embedding a complete watermark. To maintain a high quality of video and reduced bandwidth requirement, a series of frames (F_m) are chosen to embed a single watermark logo instead of embedding the watermark logo in every frame.

3. Secret Keys for Watermarking

The security of a watermarking technique should rely on secret keys rather than on a watermarking algorithm. A secret key is used for scrambling coefficients or watermark bits. This key should be difficult to predict for efficient watermarking techniques.

The proposed algorithm employs three secret keys — one for scrambling the watermark before embedding, one to choose F_p frames randomly out of a sequence of F_m frames in a video, and one to choose a set of DCT coefficients for embedding.

4. Robustness Criteria

When designing a watermarking technique, one of the major concerns is that of its robustness against intentional and unintentional malicious attacks. In the proposed algorithm, SURF features are used to select blocks in a frame to meet the requirement of robustness. Resilience to MJP2K compression is considered; and to improve the robustness against MJP2K compression, the watermark is spatially and temporally embedded.

5. Embedding and Detection in Real Time

The processes of embedding and detecting a watermark must be done in real time. Many of the current video watermarking techniques for embedding a watermark have been motivated by video coding and compression. Frame-by-frame watermarking is found to be more robust and is not specific to a particular compression standard. A number of approaches for embedding a watermark in video are based on the idea of spread spectrum communications. In these approaches, the PSNR is too low as there is more interference with the host signal. In view of the limitation of the low PSNR discussed here, techniques that perform coefficient quantization to embed a watermark using a human visual system (HVS) model are investigated in this paper.

This paper presents QIM for embedding a watermark. Quantization-based watermarking techniques have a higher payload capacity and are preferable for a blind extraction process. A QIM method is a class of nonlinear methods that rejects host-signal interference. A well-designed quantizer will produce a discrete output signal with low distortion. Significant coefficients are quantized using scalar quantization as given in (1), where x is a transform coefficient to be quantized with a quantization step size, q.

$$Q(x,q) = \text{round}\left(\frac{x}{q}\right)\cdot q. \quad (1)$$

QIM, a quantization technique, has a more favorable performance characteristic for watermarking in terms of its ability to achieve trade-offs among the robustness of the embedding, the degradation to the host signal, and the amount of data embedded.

A block-based DMWT-domain embedding algorithm for video is proposed. The criterion to select a block ($b \times b$ pixels) from a frame for embedding is vital. The blocks are usually selected according to certain features present, such as blocks containing high border areas, edges, texture [13], energy, entropy [10], mean, and so on. Sometimes, there may be no significant edges or texture features in a frame, as in the case of a still background frame. Moreover, the same embedding position cannot be chosen for every frame. The blocks may be located in plain areas in some frames; thus, there would be a discernible watermark problem. Scene-based video watermarking [14] methods or a chosen sequence of frames will solve some of the aforementioned problems.

The proposed algorithm considers the SURF features of blocks to determine which blocks should be used for embedding purposes. SURF features are invariant to translation, rotation, and scaling [17]. Furthermore, SURF has the added advantage of fast computation of feature points. SURF features have been used for object tracking and recognition; camera calibration; image retrieval; and so on. Our proposed algorithm means that it is possible to detect a watermark without the need for the original video.

The aim of the proposed algorithm is to develop high-quality watermarked video that is robust to common image processing attacks, PSNR values higher than 45 dB, and compression attacks (so as to be compatible for digital cinema). To obtain an overall performance that is of high quality and robust against compression attacks, we propose an algorithm that integrates the following schemes.

- The SURF technique is used to determine the characteristic featured blocks for embedding so as to meet the requirement of invisibility and so as to be able to withstand malicious attacks.
- Embedding using QIM in a DMWT domain has given better performance than recent works, as DMWT produces coefficients that are orthogonal and has high energy compaction.
- A simple Hamming code is used as the error-correction code. This error-correction code helps improve the robustness of the algorithm.
- Requirement of high-quality video used in medical and digital cinema editing, archive, and distribution.

III. Proposed Video Watermarking Scheme

The proposed development of a block-based robust video

watermarking algorithm uses two transform domain techniques — DMWT and DCT. The DCT technique is also a proven method for watermarking and compression in image processing. Most of the significant information of an image is concentrated in a few low-frequency coefficients of the DCT. High-frequency coefficients of the image are usually removed during compression and are prone to various noise attacks. The energy compaction property of DCT helps to identify the best coefficients for watermark embedding.

The DCT technique is combined with DMWT, which is a vectored wavelet transform. In particular, the hierarchical property of the DMWT technique gives scope to analyse the signal at different levels and orientations. It is compatible with HVS, avoiding blocking artefacts. In the DMWT domain, the essential properties of the filter design, such as orthogonality, symmetry, short support, and higher vanishing moments, are achieved simultaneously, which is not possible in scalar wavelets. Moreover, energy compaction is higher in DMWT than in DWT; hence, coefficients with a high level of energy compaction are chosen as significant coefficients for embedding. Utilizing the aforementioned properties of DCT and DMWT, we propose a combined DCT-DMWT-based video watermarking algorithm. Symmetric keys, k_1, k_2, and k_3, are generated using a Rijndael algorithm [18] and are based on a secret code. Out of a series of F_m frames in a video sequence, a number of frames, F_p, is chosen using the key k_2 to embed the watermark. Then the SURF technique is used to detect the location of points of interest. The SURF technique is reliable at finding these same points of interest under various viewing conditions. The most important properties of a points-of-interest detector include its repeatability and its ability to find the best possible matching of feature vectors with as low a dimension as possible and within the minimum time possible; that is, it is to have a fast computation time without sacrificing the performance of an embedding algorithm. SURF features are produced for the same block of information even after applying various attacks, such as scaling, rotation, and affine transformation (see Fig. 1). As the characteristic features are preserved even after attacks, these located blocks are used for watermark extraction. Hence, a block of size $b \times b$ with points of interest is selected for the embedding of a watermark bit. The luminance, Y, of every color block is transformed to the second level of the multiwavelet domain. The choice of the second level of decomposition is a trade-off between invisibility and robustness of the watermark to attacks. The low-frequency "LowLow" sub-band (see Fig. 2) is chosen because of its higher energy compaction. These coefficients are further applied to DCT. The secret key k_3 chooses a set of c DCT coefficients using a technique that generates a pseudo random number for the given secret key k_3 to embed the

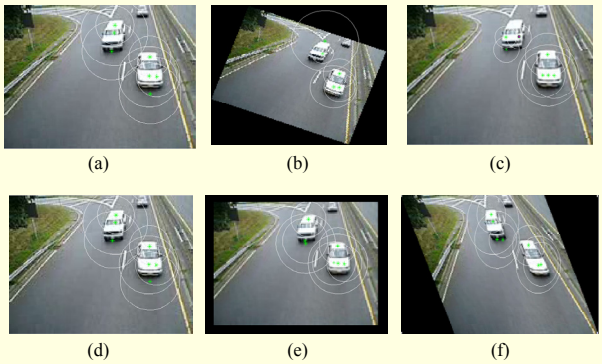

Fig. 1. Detection of SURF features under various attacks: (a) original, (b) rotation 20°, (c) resize 120%, (d) JPEG 80%, (e) resize 80%, and (f) affine transform.

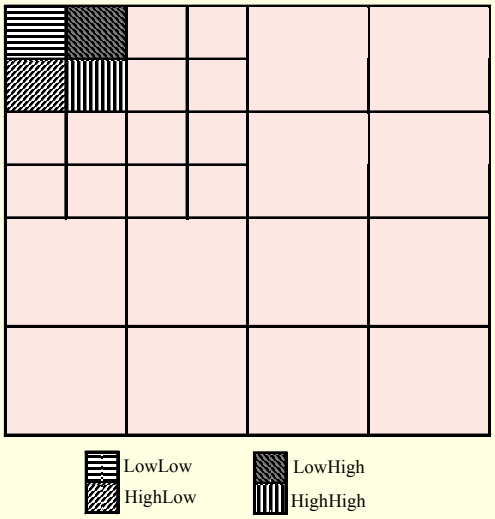

Fig. 2. GHM multiwavelet second-level sub-band classification (with repeated row preprocessing).

watermark. The watermark logo is scrambled by symmetric key k_1 to add more security. Even if the watermark is retrieved without knowledge of k_1, the original logo cannot be regained. The watermark logo is applied to an error-correction code and then scrambled.

A block diagram of the proposed video watermarking algorithm is shown in Fig. 3. The watermark embedding algorithm is as follows:

Step 1. A total of F_m frames from a test video are chosen. Randomly, F_p frames, $F_p \le F_m$, are selected using symmetric key k_2. The selected frames must be sufficiently large to embed a watermark logo of size $m \times n$, and each frame is divided into blocks of size $b \times b$.

Step 2. The SURF technique is applied on a selected frame. The frame is divided into blocks of size $b \times b$. The blocks containing maximum significant and strongest points of

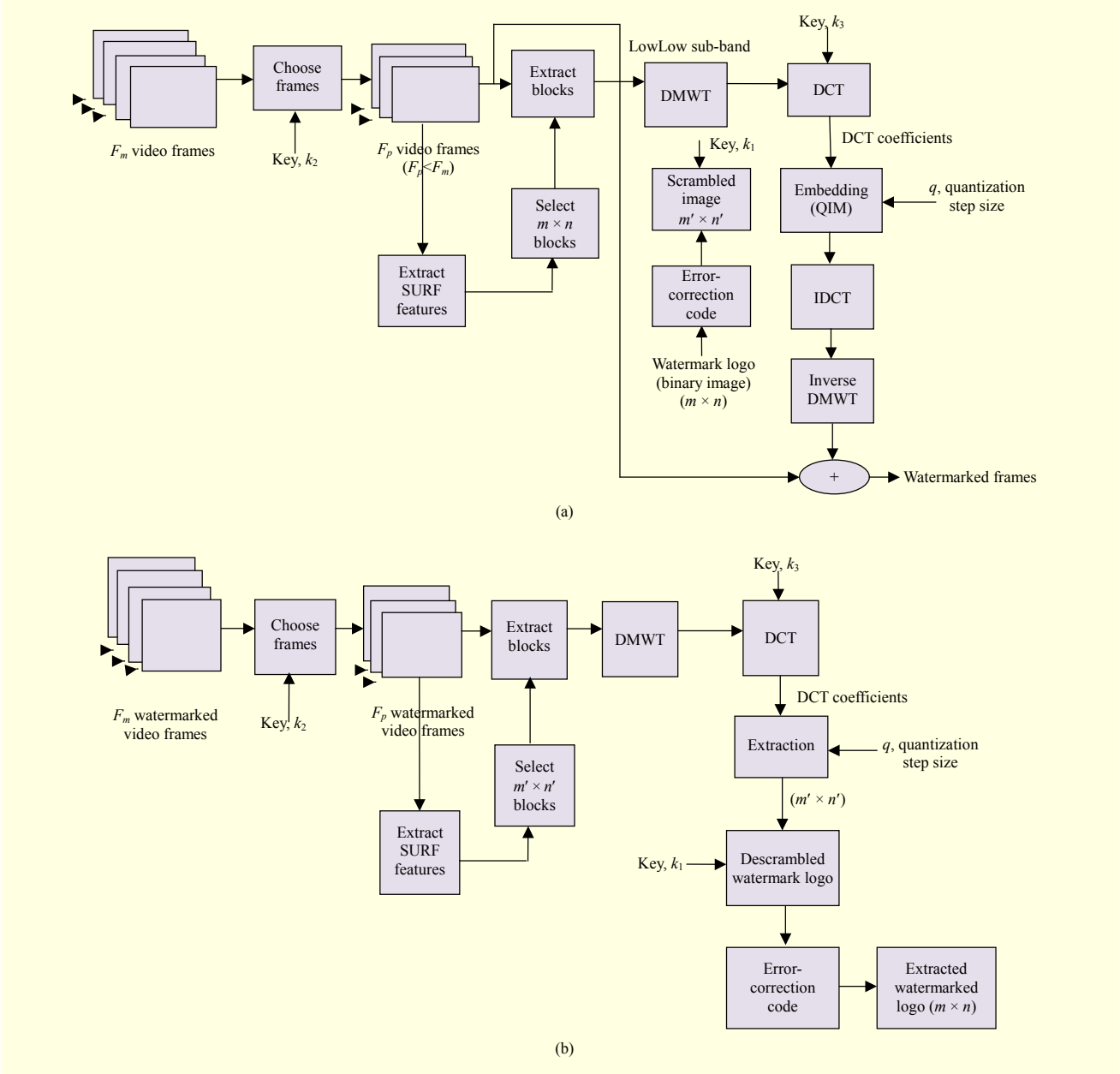

Fig. 3. Block diagram of proposed video watermarking algorithm: (a) embedding system and (b) extraction system.

interest are chosen. The number of blocks chosen must be equal to the modified size of the watermark logo; namely, $m' \times n'$ after error correction.

Step 3. The luminance component Y from an extracted RGB block, denoted as $L_{(ip\times jp)}$, in frame F_p is obtained.

Step 4. The second level of decomposition of DMWT (GHM multiwavelet, with repeated row (RR) preprocessing) is applied on the selected luminance component of the block, $L_{(ip\times jp)}$. The obtained LowLow sub-band is transformed using DCT.

Step 5. A set of c coefficients of DCT; namely, x_i, are selected by symmetric key k_3. These coefficients are quantized according to the watermark logo bits with a quantization step size, q.

Step 6. The quantization step size q is optimized by using a genetic algorithm (see (6)).

$$\begin{gathered}\text{if}(w==1)\\ k_3 = 0.75\\ \text{else}\quad k_3 = 0.25\\ \text{if}\ x(i) > 0\\ x'(i) = x(i) - x(i)\ \bmod\ (q) + k_3 \times q\\ \text{else}\\ x'(i) = x(i) - \text{sign}\left(x(i)\right) \times \text{abs}\left(x(i) \bmod q\right) - k_3 \times q\end{gathered} \tag{2}$$

for $i = 1, 2, 3, \ldots, c$, where $x(i)$ and $x'(i)$ are DCT coefficients and watermarked DCT coefficients, respectively.

Step 7. An inverse DCT is carried out on the selected block.

Step 8. An inverse DMWT that includes other sub-bands is carried out to produce the watermarked block. These selected blocks are transformed back to RGB blocks.

Step 9. Steps 3 through 7 are repeated for the remaining selected blocks to obtain the watermarked frames.

Step 10. The average PSNR, PSNR_{avg}, is calculated over F_m frames as follows:

$$\text{PSNR}_{\text{avg}} = \frac{\sum_{j=1}^{F_m} \text{PSNR}(j)}{F_m}. \quad (3)$$

Equation (2) is a QIM method of non-linear approximation of selected coefficients $x(i)$ according to the watermark. The remainder obtained after dividing by q of the selected coefficients $x(i)$ is adjusted to be 3/4 or 1/4 of the value of q. The average PSNR given in (3) determines the quality of the watermarked video over m frames.

The watermark extraction algorithm is summarized as follows:

Step 1. A total of F_m frames from a watermarked video are selected. Randomly, F_p frames are selected using the same symmetric key as in the embedding algorithm, k_2, and the frame is divided into blocks, each of which is $b \times b$ in size.

Step 2. The SURF technique is used to detect the points of interest in the selected F_p frames. Blocks of size $b \times b$ containing maximum and strongest significant points of interest are chosen.

Step 3. The luminance component Y from an extracted RGB block, denoted as $L'_{(ip \times jp)}$, in frame F_p is obtained.

Step 4. The second level of decomposition of DMWT (GHM multiwavelet, with RR preprocessing) is applied on the selected watermarked luminance block, $L'_{(ip \times jp)}$. The obtained LowLow sub-band is transformed into DCT coefficients.

Step 5. A set of c coefficients selected by symmetric key k_3 is used to detect the watermark logo bit.

$$\begin{gathered} \text{Initialize count} = 0 \\ \text{if } \text{abs}\left((x'(i)) \times \text{mod}(q)\right) \geq 0.5 \times q \\ \text{count} = \text{count} + 1 \text{ for } i = 1, 2, 3, \ldots, c \\ \text{if } \text{count} \geq \frac{c}{2}, w = 1 \\ \text{else } w = 0. \end{gathered} \quad (4)$$

Step 6. Steps 3 through 5 are repeated for the remaining selected blocks to obtain the watermark logo. The obtained logo is descrambled with the secret key k_1 and applied to an error-correction code.

Step 7. After extracting the watermark, normal cross correlation (NCC) between the original watermark, P_{mn}, and the extracted one, Q_{mn}, is performed to quantify the similarity.

$$\text{NCC} = \frac{\sum_m \sum_n (P_{mn} - P')(Q_{mn} - Q')}{\sqrt{\left(\sum_m \sum_n (P_{mn} - P')^2\right)\left(\sum_m \sum_n (Q_{mn} - Q')^2\right)}},$$

where

$$P' = \text{mean}(P) \text{ and } Q' = \text{mean}(Q). \quad (5)$$

Equation (4) checks the remainder of the selected coefficients, $x'(i)$, and if more than 50% of them are above half of the value of q, then the watermark bit is marked as 1; that is, $w = 1$; else $w = 0$. Equation (5) calculates NCC to evaluate the performance measure of the algorithm.

IV. Experimental Results and Performance Evaluation

Experimentation is performed on four standard test video sequences in RGB uncompressed AVI format (see Fig. 4). The resolution is 120 × 160, and the frame rate is 30 frames per second. The video sequences chosen are for simulation of results. Higher resolutions and larger frame sizes can be implemented. More information can be embedded with larger frame sizes. The proposed algorithm is tested using the following two methods: (i) combined DWT with DCT and (ii) combined DMWT with DCT. This is so as to be able to compare the results of the proposed algorithm in wavelet and DMWT domains. The coefficients obtained in the DMWT domain are much larger in size than those obtained in the wavelet transform domain. Hence, quantization-based embedding is best suited and higher noise immunity is observed in DMWT than DWT. The objective perceptual quality of the watermarked videos is measured in terms of average PSNR. The average PSNR of watermarked video frames and NCC of the watermark logo to be extracted is computed for test video sequences (see Table 1). The average

(a)

(b)

Fig. 4. (a) Original and (b) watermarked video sequences (viptraffic, carphone, foreman, and Suzie, respectively).

PSNR after an attack and efficient retrieval of a watermark (despite such degradation) is observed in both the DWT domain and the DMWT domain.

The proposed algorithm transforms only small-sized blocks as opposed to a complete frame. So, the computational complexity of a multiwavelet transform is comparable to that of a wavelet transform; and using a multiwavelet transformation, the watermarked video quality is found to be higher than that from a wavelet transformation. The sustainability against a compression attack is tested for the four given video sequences. The method using DMWT with DCT outperformed with higher average PSNR values under an MJP2K compression

Table 1. Average PSNR of watermarked video sequences and NCC of extracted watermark with different quantization step sizes (q), with and without MJP2K compression.

Video			Viptraffic		Carphone		Foreman		Suzie	
Transform methods			Wavelet + DCT	Multiwavelet + DCT	Wavelet + DCT	Multiwavelet + DCT	Wavelet + DCT	Multiwavelet + DCT	Wavelet + DCT	Multiwavelet + DCT
Attack	Q									
No attack	Q=16	PSNR	50.55	51.27	52.14	52.75	52.38	52.89	52.45	53.00
		CORR	1.00	1.00	1.00	1.00	0.90	0.93	1.00	1.00
	Q=40	PSNR	48.19	49.67	49.71	51.08	50.35	51.50	50.49	51.60
		CORR	1.00	1.00	1.00	1.00	0.93	0.93	1.00	1.00
	Q=64	PSNR	46.45	48.22	48.02	49.63	48.98	50.34	49.03	50.33
		CORR	1.00	1.00	1.00	1.00	0.86	0.86	1.00	1.00
	Q=88	PSNR	44.89	46.83	46.87	48.51	47.86	49.35	48.07	49.41
		CORR	1.00	1.00	1.00	1.00	0.85	0.86	0.99	0.99
Compression ratio 2 (50%)	Q=16	PSNR	48.36	48.84	52.14	52.75	46.64	48.67	48.51	48.79
		CORR	1.00	1.00	1.00	1.00	0.90	0.90	1.00	1.00
	Q=40	PSNR	46.49	47.72	46.61	47.60	46.64	47.91	47.17	48.01
		CORR	1.00	1.00	1.00	1.00	0.91	0.93	1.00	1.00
	Q=64	PSNR	44.93	46.51	45.19	46.55	46.64	47.07	45.90	47.04
		CORR	1.00	1.00	1.00	1.00	0.86	0.86	1.00	1.00
	Q=88	PSNR	43.43	45.27	44.14	45.63	46.64	46.25	45.00	46.24
		CORR	1.00	1.00	1.00	1.00	0.85	0.86	0.99	0.99
Compression ratio 5 (80%)	Q=16	PSNR	46.96	47.35	45.78	45.95	43.84	46.64	47.42	47.64
		CORR	1.00	1.00	1.00	1.00	0.86	0.89	1.00	1.00
	Q=40	PSNR	45.32	46.41	44.65	45.39	43.50	44.65	46.25	47.00
		CORR	1.00	1.00	1.00	1.00	0.88	0.91	1.00	1.00
	Q=64	PSNR	43.84	45.34	43.46	44.61	43.03	43.69	45.06	46.13
		CORR	1.00	1.00	1.00	1.00	0.86	0.86	1.00	1.00
	Q=88	PSNR	42.40	44.18	42.51	43.85	42.47	43.14	44.19	45.39
		CORR	1.00	1.00	1.00	1.00	0.85	0.86	0.99	0.99
Compression ratio 10 (90%)	Q=16	PSNR	40.89	40.99	41.16	41.21	37.71	42.96	42.96	43.05
		CORR	0.92	0.64	0.97	0.76	0.45	0.81	1.00	0.93
	Q=40	PSNR	40.26	40.72	40.65	40.96	37.61	46.64	42.36	42.78
		CORR	1.00	0.94	1.00	0.97	0.85	0.87	1.00	1.00
	Q=64	PSNR	39.42	40.26	39.92	40.57	37.47	46.64	41.54	42.31
		CORR	1.00	1.00	1.00	1.00	0.86	0.86	1.00	1.00
	Q=88	PSNR	38.44	39.66	39.24	40.11	37.22	46.64	40.83	41.80
		CORR	1.00	1.00	1.00	1.00	0.85	0.86	0.99	0.99

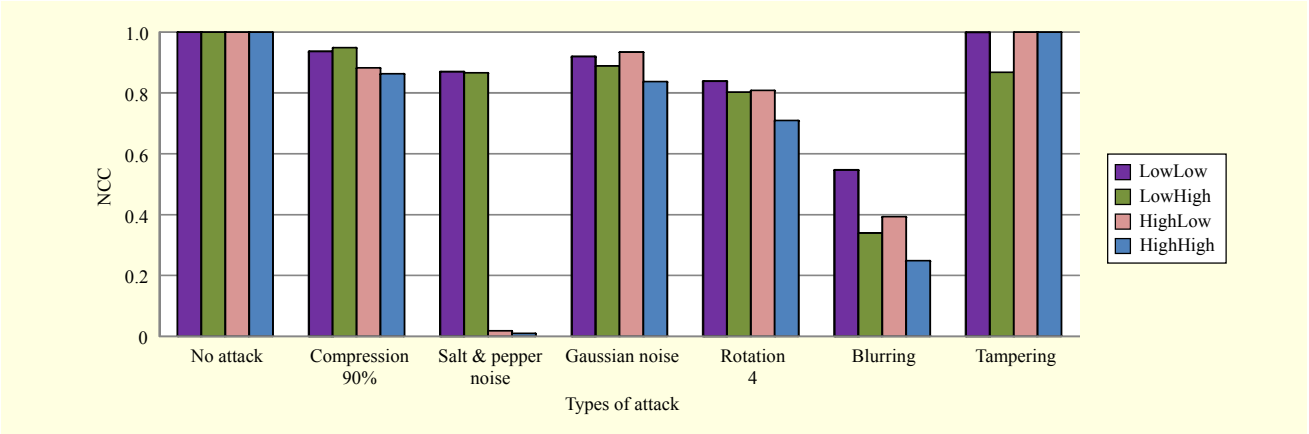

Fig. 5. NCC for various attacks in different sub-bands used for embedding.

attack and better extraction of the watermark logo (shown in Table 1).

Hence, the proposed algorithm using the DMWT domain is analyzed for performance. The embedding capacity can be increased by selecting a greater number of feature points in the SURF feature extraction method. The strongest feature threshold to select SURF points is taken as 1,000. The size of the watermark logo used is 22 × 15 for experimentation. Out of 30 frames, 20 frames were sufficient to satisfy this condition. Four DCT coefficients including a DC component are chosen to obtain spatial redundancy. These four DCT coefficients are quantized as per one bit of the watermark logo.

The quantization step q is optimized by a genetic algorithm. The fitness function is as follows:

$$f(c) = \text{avgPSNR} + \sum_{m=1}^{P} (\text{NCC}_{c,m} \times \alpha_{c,m}), \quad (6)$$

where p represents the number of attacks, $\alpha_{c,m}$ is a weighting factor to NCC, avgPSNR is the average PSNR of a video stream, and $f(c)$ is the fitness value. The average PSNR is important for the quality of the video, whereas NCC determines the robustness of the algorithm. The value of q, using a genetic algorithm, is found to be 40. The experimentation is performed in various DMWT sub-bands; namely, LowLow, LowHigh, HighLow, and HighHigh, for various attacks (see Fig. 5).

We can infer from the experiment that the quantization of DCT coefficients of the LowLow DMWT sub-band produced better NCC than other sub-bands. If embedding is performed in those coefficients with higher energy compaction, then any modification done to these coefficients due to an attack will destroy the original video. Hence, any possible attack to modify or remove the embedded watermark may destroy the original video. Moreover, the embedded watermark can sustain attacks that can modify the coefficients by up to half of the quantization step size given in (4). Hence, the QIM technique proposed in this work gives good performance if embedding is done in larger sized coefficients and with a higher quantization step size. However, the use of optimization techniques, such as a genetic algorithm, would be impractical for real-time video applications. In such cases, an appropriate quantization step can be fixed experimentally. The watermark is embedded with temporal redundancy once in every 30 frames so as to improve the correlation of the retrieved watermark logo. The error-correction code used is a Hamming code with three different codeword lengths — (31, 26), (15, 11), and (7, 4). It is a simple code to correct one bit error. The extra bits added for error correction are 65, 120, and 249, respectively, for a watermark logo of size 22 × 15. There is always a trade-off between the selection of codeword lengths and quality of the video. Experimentally, the codeword length of 15 bits and data word length of 11 bits performed well under various attacks (see Table 2).

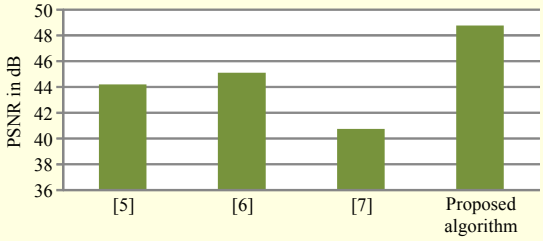

Fig. 6. Average PSNR values of authenticated and recovered frames.

The proposed algorithm is compared with [5]–[7]. The perceptual quality of the watermarked video in the proposed algorithm is noticeably improved compared to when other algorithms are used (see Fig. 6). Table 3 gives a comparison of different algorithms including the proposed algorithm in terms of calculated NCC between the authenticated watermark and

Table 2. Experimental results and comparison under various attacks.

Attack	Proposed method							
	No error correction		Error correction (31, 26)		Error correction (15, 11)		Error correction (7, 4)	
	PSNR	NCC	PSNR	NCC	PSNR	NCC	PSNR	NCC
No attack	49.67	1	49.26	1	48.76	1	48.04	1
Motion JPEG2000 compression ratio 2 (50%)	47.72	1	47.44	1	47.15	1	46.70	1
Motion JPEG2000 compression ratio 5 (80%)	46.41	1	46.21	1	46.04	1	45.74	1
Motion JPEG2000 compression ratio 10 (90%)	40.72	0.9371	40.66	1	40.64	1	40.56	1
Motion JPEG2000 compression ratio 15 (93%)	37.01	NS	36.99	0.6765	36.97	0.9226	36.95	0.8357
Frame swapping	48.29	0.9454	48.03	0.8934	47.51	1	46.77	1
Frame averaging	46.98	0.8357	46.58	0.7940	46.07	1	45.33	1
Salt & pepper 0.002	32.91	0.9111	32.93	0.8023	32.49	0.8964	32.83	0.9346
Salt & pepper 0.003	30.99	0.8081	31.09	0.7636	31.41	0.9111	31.05	0.9346
Rotation 2 degree	23.89	0.8388	23.89	0.8850	23.89	1	23.89	0.9256
Rotation 4 degree	20.55	0.8388	20.55	0.8740	20.55	1	20.55	0.9111
Rotation 8 degree	17.57	0.2134	17.57	0.5254	17.57	0.5038	17.57	0.5266
Gaussian noise 0.003	34.85	0.9371	34.88	1	34.86	0.9736	34.87	1
Motion blur (2, 7)	30.71	0.5470	30.71	0.4390	30.71	0.3951	30.71	0.5161
Blurring (0.65)	37.68	0.7566	37.68	0.7638	37.68	0.7865	37.68	0.9256
Tampering	12.21	1	12.21	1	12.21	1	12.21	1

the recovered watermark under various attacks. The quality of the retrieved watermark has an acceptable NCC value; that is, greater than 0.80. The proposed algorithm shows better sustainability against rotational attacks and motion blurring attacks in comparison with other algorithms. R.O. Preda and N.D. Vizireanu's [6] algorithm used a complete frame for a wavelet transformation, whereas our proposed algorithm uses the significantly smaller sized characteristic SURF blocks for DMWT. The DMWT is applied on a smaller block size rather than the whole frame. Hence, the computational complexity of the proposed algorithm is almost comparable to that of R.O. Preda and N.D. Vizireanu [6], although the computational complexity of our DMWT is higher than that of a wavelet transform. The algorithms proposed in [5]–[7] have shown a poor sustainability against attacks and have a low PSNR value compared to that of our proposed algorithm.

In [8], a novel digital watermarking method was proposed using SURF against RST attacks. The SURF features, which are extracted from a group of frames, can be saved for attack-type estimation and image correction after an attack. But, this method needs to store all of the SURF features related to every video, which is impractical. Moreover, the method becomes non-blind.

The proposed algorithm has outperformed sustainability against attacks, particularly up to four degrees of rotational attack, blurring with a disk kernel of 0.65, tampering in a 40 × 40 pixel area, and noise addition attacks with zero-mean and a Gaussian white noise of 0.003 local variance. Also, MJP2K compression up to 80% of the original frame size is achieved by maintaining the video quality to be above 45 dB. In the absence of an attack, the quality of the video is above 48 dB, which appears visually identical to the original video. As the embedding is done using the references of SURF feature points, the blocks used for embedding are correctly identified for extraction even after geometrical attacks. Our watermarking scheme is robust against frame averaging, frame swapping, and MJP2K compression, as well as having good resilience against common image processing attacks such as salt & pepper noise, Gaussian noise, rotation, blurring, tampering, and motion blur.

V. Conclusion

This paper presented a robust, imperceptible high-quality video watermarking method based on DCT in the DMWT domain. The SURF technique, which is invariant to translation,

Table 3. NCC between authenticated watermark and recovered watermark.

Attacks	NCC										
	No attack	Motion JPEG2000 compression ratio 5 (80%)	Gaussian noise (0.003)	Salt & pepper noise (0.003)	Rotation 4 degree	Motion blur (2, 7)	Blurring (0.65)	Tampering	Frame swapping	Frame averaging	Frame dropping
[5]	1	0.8923	0.9346	0.4365	–0.0125	0.0221	0.3265	0.8357	0.9765	0.9256	0.9884
[6]	1	0.7863	0.9764	0.9882	0.0024	0.3484	0.6395	1	1	0.9478	0.9444
[7]	1	0.6255	0.8762	0.7863	–0.0004	–0.0012	0.4356	0.7982	0.6554	0.9221	0.8774
Proposed algorithm	1	1	0.9736	0.9111	1	0.3951	0.7865	1	1	1	1

rotation, and scaling, is used for the selection of blocks in a video frame for embedding. The secret to the algorithm lies not in the procedure but in the secret keys used and number of strongest SURF features extracted to locate the blocks for embedding.

The faster computation of SURF features without sacrificing the performance makes the algorithm simple, and it is suited for real-time applications. This proposed algorithm has good resilience against MJP2K compression, is robust against common image processing attacks, and has a high PSNR value when no attacks are used. These factors result in a high-quality video; thus, the algorithm is suitable for use in medical imaging and digital cinema applications. As per the specifications of Digital Cinema Initiatives (DCI), the forensic mark data payload should contain the following information about a movie: location of the movie played (19 bits) and time stamp on information (16 bits). All 35 bits are required to be included in each five minute segment. Under the given video frame sizes and due to the limited number of SURF descriptors, the proposed algorithm can insert a maximum of 500 bits of watermark in a video segment of 1 s, which is more than the DCI requirement (35 bits). Hence, the proposed algorithm is apt for adding watermark to digital cinema.

Machine learning methods can be used to find appropriate blocks for embedding that can sustain an embedded bit under various attacks. Further improvements to the algorithm can be made for a synchronization misalignment; for example, SURF descriptors could be stored and retrieved in a detection algorithm. We are also considering extending the proposed algorithm so as to be resistant to other standard compression techniques.

References

[1] W. Pei, Z. Zhendong, and L. Li, "A Video Watermarking Scheme Based on Motion Vectors and Mode Selection," *IEEE Int. Conf. Comput. Sci. Softw. Eng.*, Wuhan, China, vol. 5, Dec. 12–14, 2008, pp. 233–237.

[2] C.-X. Wang et al., "A Blind Video Watermarking Scheme Based on DWT," *Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Kyoto, Japan, Sept. 12–14, 2009, pp. 434–437.

[3] C. Cruz-Ramos et al., "A Blind Video Watermarking Scheme Robust to Frame Attacks Combined with MPEG2 Compression," *J. Appl. Res. Technol.*, vol. 8, no. 3, 2010, pp. 323–337.

[4] L. Coria et al., "A Video Watermarking Scheme Based on the Dual-Tree Complex Wavelet Transform," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 3, Sept. 2008, pp. 466–474.

[5] O.S. Faragallah, "Efficient Video Watermarking Based on Singular Value Decomposition in the Discrete Wavelet Transform Domain," *AEU-Int. J. Electron. Commun.*, vol. 67, no. 3, Mar. 2013, pp. 189–196.

[6] R.O. Preda and N.D. Vizireanu, "Quantisation-Based Video Watermarking in the Wavelet Domain with Spatial and Temporal Redundancy," *Int. J. Electron.*, vol. 98, no. 3, 2011, pp. 393–405.

[7] P.W. Chan, M.R. Lyu, and R.T. Chin, "A Novel Scheme for Hybrid Digital Video Watermarking: Approach, Evaluation and Experimentation," *IEEE Trans. Syst. Video Technol.*, vol. 15, no. 12, Dec. 2005, pp. 1638–1649.

[8] B. Zhang et al., "A Novel Digital Watermark against RST Distortion Based on SURF," *IEEE Int. Conf. Inf. Theory Inf. Security*, Beijing, China, Dec. 17–19, 2010, pp. 130–133.

[9] A. Al-Haj, "Combined DWT-DCT Digital Image Watermarking," *J. Comput. Sci.*, vol. 3, no. 9, 2007, pp. 740–746.

[10] N.I. Yassin, N.M. Salem, and M.I.E. Adawy, "Entropy Based Video Watermarking Scheme Using Wavelet Transform and Principle Component Analysis," *Int. Conf. Eng. Technol.*, Cairo, Egypt, Oct. 10–11, 2012, pp. 1–5.

[11] C.V. Serdean et al., "Wavelet and Multiwavelet Watermarking," *IET Image Process.*, vol. 1, no. 2, June 2007, pp. 223–230.

[12] B. Chen and G.W. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," *IEEE Trans. Inf. Theory*, vol. 47, no. 4,

May 2001, pp. 1423–1443.
[13] X.-B. Zheng and X.-W. Zhang, "A Multiwavelet Based Digital Watermarking Algorithm Using Texture Measures," *Int. Conf. Wavelet Anal. Pattern Recogn.*, Qingdao, China, July 11–14, 2010, pp. 260–265.
[14] C. Pik-Wah, "Digital Video Watermarking Techniques for Secure Multimedia Creation and Delivery," Ph.D. dissertation, CiteSeer, 2004.
[15] S. Joo et al., "A New Robust Watermark Embedding into Wavelet DC Components," *ETRI J.*, vol. 24, no. 5, Oct. 2002, pp. 401–404.
[16] C.-Y. Yang and C.-H. Lin, "High-Quality and Robust Reversible Data Hiding by Coefficient Shifting Algorithm," *ETRI J.*, vol. 34, no. 3, June 2012, pp. 429–438.
[17] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," *European. Conf. Comput. Vis.*, Graz, Austria, May 7–13, 2006, pp. 404–417.
[18] J. Daemen and V. Rijmen, "AES Proposal: Rijndael," Sept. 3, 1999.

Leelavathy Narkedamilly received her BE degree in electronics and communication engineering from Vasavi College of Engineering, Osmania University, Telangana, India, in 1992 and her MTech degree in computer science from Jawaharlal Nehru Technological University, Telangana, India, in 2003. She is currently working as a professor and is head of the Department of Computer Science and Engineering, Pragati Engineering College, Andhra Pradesh, India. She is working towards her PhD degree in image processing at Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, India. She has fifteen years of experience in teaching undergraduate and post-graduate students. Her research interests are in the area of digital image processing, image watermarking, cryptography, Hadoop, big data, and network security.

Venkateswara Prasad Evani is currently the director at Lakkireddy Bali Reddy College of Engineering, Andhra Pradesh, India. Prior to his current assignment, he was the Rector at Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, India. He received his PhD degree in computer science and engineering from the University of Roorkee, Uttarakhand, India, in 1990. He received his ME degree in computer science from Madras University, Chennai, India, in 1978 and his BE degree in electronics and communication engineering from Sri Venkateswara University, Andhra Pradesh, India, in 1975. He has thirty-five years of experience in teaching undergraduate and post-graduate students. He held different positions in his career, such as vice principal, principal, director, registrar, and chairman of the board of studies. He has guided seven students toward PhD degrees, co-authored six books, and published 102 research publications to date in national and international journals and conferences. He was the recipient of the "State Best Teacher" award in 2008 — an award that is given to meritorious teachers by the government of Andhra Pradesh, India. His research interests include parallel computing, data mining, image processing, and information security.

Srinivas Kumar Samayamantula is currently working as a professor of electronics and communication engineering and is a director the Sponsored Research and Development, Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, India. He received his PhD degree in digital image processing from the Electronics and Electrical Communication Engineering Department, IIT Kharagpur, India, in 2002. He received his MTech degree in electronics and communication engineering from Jawaharlal Nehru Technological University, Telangana, India. He has twenty-one years of experience teaching undergraduate and post-graduate students and has guided a number of post-graduate theses. He has published 30 research papers in national and international journals. Presently, he is guiding ten PhD students in the area of image processing. His research interests are in the areas of digital image processing; computer vision; and application of artificial neural networks and fuzzy logic to engineering problems.