Video Watermarking Scheme based on Candidates I-Frames for Copyright Protection

Rakesh Ahuja\textsuperscript{1}, S. S. Bedi\textsuperscript{2}
\textsuperscript{1} Department of Computer Science & Engineering, Moradabad Institute of Technology, Moradabad, India
\textsuperscript{2} Department of Computer Science & Information Technology, MJP Rohilkhand University, Bareilly, India
Corresponding author, e-mail: ahuja2305@gmail.com\textsuperscript{1}, dearbedi@gmail.com\textsuperscript{2}

Abstract

The proposed scheme embedded the watermark during the differential pulse code modulation process and extracted through decoding the entropy details. This technique utilizes the moving picture expert groups standard (MPEG-2) in which discrete cosine transform coefficients are adjusted from selected instantaneous decoder refresh frames for watermarking purpose. The subsets of frames as candidate I-frames are chosen to achieve better perceptibility and robustness. A secret key based cryptographic technique is used to select the candidate frames. Three more keys are required to extract the watermark whereas one of the key is used to stop the extraction process and the remaining two are used to display the scrambled watermark. The toughness is evaluated by testing spatial and temporal synchronization attacks. High sturdiness is achieved against video specific attacks frequently occurs in the real world. Even a single frame can accommodate thousand of watermark bits which reflect that high watermark capacity can be obtained.

Keywords: discrete cosine transforms, information security, MPEG-2 standard, signal processing, video watermarking

Copyright © 2017 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

The success of internet and better bandwidth has made it easier to transfer, edit and copy digital multimedia data from one system to other irrespective of their geographic location. This ability raises issues like creation of illegal redistribution, piracy, tempering, ownership and false copyright of digital multimedia data. Encryption [1] is the technique used earlier for protecting the contents. Once decrypted, this technique fails as the decrypted contents are exactly same as original one. Steganography [2] is also used to hide the secret information into the cover object. The limitation of this technique is that no relationship exists between cover and hidden information. Therefore its applications are restricted and therefore never been used for copyright protection of multimedia contents. Digital watermarking scheme is used to address the above said issues. Digital watermarking schemes embeds the owner specific information as watermark into the multimedia object in order to protect the copyright in a robust and imperceptible manner. The benefit associated with this technique is that the watermark can be extracted at any time from the watermarked object to claim the ownership of the related multimedia contents.

Digital video watermarking techniques are exploited for wide variety of applications [3] in addition to copyright protection. For example, one of the application is used to prevent creating illegal multiple copies in copy control applications by implementing the fingerprinting issue in which customer information is embedded as watermark into the video to trace the user responsible for creating the illegal multiple copies called destination based watermarking. The other relevance is to support the ownership protection also named as source based watermarking in which the owner’s data is inserted into the host signal and can be extracted whenever needed for proof of ownership. The most significance purpose of watermarking is to test the integrity of the video contents. The watermarking process also verifies the video content during broadcast monitoring. In this scheme, the video contents are broadcast along with identification information which decoded straightforward reliably in order to interpret correctly.

The video watermarking schemes are based on three types of video signal: original video, compressed video and during encoding the video. The video is original in two ways;
either it never been compressed before or decompress before embedding. In principle both are same and considered as the sequence of still images. The image watermarking techniques [4-5] extended for video multimedia objects also. However, such schemes always undergo because video objects consist additional features that do not present in the image. One of the distinguishable features of video is the existence of high temporal redundancies between or among the frames unlike image. Another feature which segregates the video from image is that video consist huge data. Therefore the primarily necessary requirement of video multimedia data is to compress it before transmitting via internet to reduce the storage requirement and to save the bandwidth. Another way of watermarking is to consider the video in compressed domain. The benefit with this scheme is that the perceptibility doesn’t affect but the shortcoming is that the limited capacity of watermark bits is allowed for embedding purpose and it is complex to design. The last category to implement the watermarking scheme is to consider the video during encoding it. In this approach, the original compression standard is partially modified in order to accommodate the watermark signal. This approach is more realistic and practical because compression and watermarking, two necessary operations, runs atomically to support real time video watermarking. Thus advance video watermarking scheme [6] preferred to perform watermarking utilizing compression standard like MPEG-2, H.264 etc because of achieving watermarking features at the cost of marginable increase in computational time.

In this paper, a new MPEG-2 based video watermarking scheme is illustrated that perform partial decoding proposed by Yoshito Ueno [7]. The proposed technique is entirely based on selection of I-frames, known as candidate frames for embedding the scrambled watermark. Four secret cryptographic keys are used to design the scheme more secure. One key is used for selecting the candidate I-frames and another three for extracting the watermark image. It makes the watermarking system as complex as possible for malicious user so that the probability of temporal synchronization attacks must be minimized in order to achieve better robustness. P-frames, B-frames and even I-frames also excluding candidate I-frames are untouched during the watermarking process illustrate that high perceptibility also obtained.

This paper is structured into five following sections. A concise literature review of digital video watermarking in compressed domain have been elaborated in section 2. Proposed digital video watermark scheme is explained in section 3. The experimental results with analyses are illustrated in section 4. Section 5 concludes the work with its limitations and future prospects.

2. Previous Research Work

In the past two decades, a number of video watermarking schemes exploiting various compression standards have been proposed by Ahuja et al. [8]. A brief about all these delivered techniques with their limitations are illustrated below.

Chiou-Ting Hsu [9] embedded the watermark into the quantized middle frequency coefficients of DCT to survive the MPEG structure. As the suggested scheme passes through the quantization operation for achieving the higher level of adaptation, therefore the limitation associated with this approach is that the embedding process always depends upon the quantization factor. Larger the quantization providing the larger survival of watermark against compression ratio otherwise poor perceptibility would be obtained. Bijan G. Mobasseri [10] simulated the experiments for extracting the embedded data from the MPEG-2 decoded bit stream. The drawback in this scheme is the limited amount of watermark capacity would be allowed. You-Ru Lin [11] inserted the watermark image according to the direction of motion vector by using the block matching algorithm. A satisfactory perceptibility is obtained however, the scheme lost approx. twenty percent of the watermark bits while extracting. Robustness is certainly affected sharply when any further video editing operations were applied on watermarked.

Satyen Biswas [12] described the video watermarking scheme based on MPEG-2 structure too. In this scheme, the gray scale image is converted into multiple bit-plane images and DCT is applied to all the bit-plane images. Each partitioned watermarking image is embedded into each different scene of the movie. In this way, this scheme inserted the watermark into the entire video. Robustness carried out by simulating the various attacks like collusion, filtering, frame dropping, rotation, blurring, scaling and temporal shift. The constraint factor is that the high computational complexity is obtained.
Lu Jianfeng [13] described the video watermarking scheme based on DCT coefficient. The DCT coefficients are adjusted for inserting the watermark image and extracted successfully without the need of original video or watermark. However, the scheme requires the careful adjustment of the control parameter as it directly affects on the quality of watermarked video. Therefore, the limitation of the scheme is that the little increasing the control parameter may drastically degrade the perceptual quality of watermarked video. Another limitation is that the robustness has been tested only against Salt & Pepper noise even though a range of image processing attack as well as video specific attacks could be apply to judge the robustness of the scheme.

Ahuja, et al., [14] further extended this work based on MPEG-2 structure to cover the above limitations. The scheme tested the robustness by covering more number of attacks also focused on improving the elapsed time to embed the watermark. The shortcoming with this scheme is that robustness has not been evaluated against unintentional data compression attack. Yoshito Ueno [8] proposed the watermarking scheme by first finding the macro block of Y or Cr or Cb component of I-frame when P-frame has a motion vector and then these blocks have been used for embedding the binary watermark bits. The positive aspect of the scheme is that neither original video nor original watermark is needed during the retrieval of watermark.

Yuk Ying Chung [15] developed the MPEG-2 video watermarking scheme based on quantized DCT coefficients. The watermark bits are embedded into the LSB of the DC coefficient obtained from DCT blocks of I frames. The common restriction with this approach is that the attacker can randomized all the LSBs in order to completely destroy the watermark. Daniel Cross, et al., [16] proposed the MPEG-2 based video watermarking method for evaluation of authentication and temper detection. The algorithm is entirely based on the compressed bit streams for identifying the watermark carrying VLCs. These VLCs are used to embed the watermark bits. The scheme compared the LSB of Ic-VLC with the watermark bit. If both the bits are same, the VLC remains unchanged otherwise LSB is replaced by the watermark bit. The constraint is obvious that the watermark capacity is directly proportional to the number of Ic-VLC bits. As illustrated above, LSB based technique is always vulnerable to attack as illustrated above. Anil Kumar Sharma [17] applied quantization index modulation (QIM) technique to embed the watermark bits into P frames. Again, the constraint with such schemes are the limited amount of watermark information is allowed to insert watermark in these types of frames because they are already highly compressed and always have less space to accommodate additional information into it.

In the summarized way, some common drawback have been identified in compressed domain based delivered techniques. One of them is exploiting entire I-frames or P-frames for watermarking purpose. Certainly, such methods are responsible for the degradation of perceptibility. Another familiar issue is to use the watermark as plain text. In this case, if the nature of the watermarking algorithm is known to attacker then there is a higher probability to detect and collapse the embedded watermark.

The objective of this paper is to design the robust video watermarking system to overcome the above limitations. In view of these issues, a novel watermarking scheme is proposed in which few I-frames are selected for embedding purpose. This selection is carried out by using the secret cryptographic key. Three more conventional keys are used to extract the watermark. Another security aspect introduces in this scheme is to ciphered the watermark before embedding it. The watermark can never be obtained without knowing these entire four key, even though the nature of watermarking algorithm is known to attacker. This process provides more strengthen to the proposed watermarking system. P-frames and B-frames are never being a part of this scheme. In addition to that, other key technologies as generation of motion vectors, motion compensated image are untouched from watermarking process. The entire resultant affect is that the better perceptibility is maintained as far as possible. Another positive aspect of this scheme is that neither original video nor original watermark is required during extraction of watermark, which supports public watermarking.

3. Proposed Video Watermarking Method

In this paper, the scheme implemented the motion estimation based video watermarking by exploiting the quantized DCT coefficients generated from I-frames during the intra-coding process through MPEG-2 style as shown in Figure 1. The binary watermark bits are
embedded by adjusting the DC components lies at the intersection of first row and first column of quantized 8x8 DCT blocks only from those I-frames which are categorized as candidates I-frames. The mapping between inserting bit and each DCT block is one to one. This process not only maintains the quality of watermarked video but also minimizes the computation cost. The design of video watermarking is started with some preprocessing steps like selection of candidate-I frame from the set of I-frames, generating the cryptographic keys used for extracting the watermark defined in the following sub-sections.

CanFr(I) : Candidate I-Frame  
$E_k(W_s)$ : Watermark ($W_s$) is Encrypted with Key $K$  
$Q$ : Quantization  
$B_w$ : Encrypted Watermark bit

3.1. Generation of Watermark Keys

A two dimensional binary image Copyright.tif bit depth 1 is used as a watermark. Three keys ($r$, $c$, size) are extracted from the original watermark itself in the following manner. The number of rows and columns are denoted by $r$ and $c$ respectively used to extract the watermark. The size of the watermark Size treated as third key used to stop the extraction process while recovering the watermark bits. Double columnar transposition method [18] is used to encrypt the watermark object ($W$) to further augmenting the security of the watermark itself. The encrypted watermark is converted into 1-Dimensional array as $B_w = \{W_s\}$, $s=1$, 2……N; $W_s \in \{1, 0\}$ to single out one bit at a instance.

3.2. Selection of Candidate I-frames

The selection of candidate I-frames from a group of all I-frames is the inevitable process of the proposed scheme. There are two major reasons for choosing such frames. Firstly, it is an elementary video frame holding the complete information about the frame and has adjustment compatibility in DC coefficients doesn’t degrade the quality of watermarked video. Bit rate of watermarked video does not increases significantly is another reason to choose IDR frames. On the contrary, there is always a less space for embedding in P-frames and B-frames as they are highly compressed frames by motion estimation techniques during encoding through MPEG-2 style. The scheme buffered entire IDR frames extracted from a video clip into two-dimensional array with the constraint that number of rows always exceeding by one than the number of columns. The watermarking process exploits only those IDR frames that are inside the oval shape as shown in the Figure 2. A simple way to pick these frames is to traverse the matrix row by row but choosing the elements (potential I-frames) from every alternate column.
3.3. Watermark Embedding Algorithm

The MPEG-2 structure is partially modified to accommodate the binary watermark. During the encoding process, each I-frame from GOP is tested for candidate I-frame. If the current I-frame is candidate I-frame then it passes through the following watermarking process. Each candidate I-frame is divided into three channels, one as luminance component (Y) and other two Chroma channels known as CB and CR respectively. The luminance component (Y) is divided into the number of equal sized (8 x 8) of blocks for applying DCT followed by quantization process. The insertion process of single watermark bit is defined by the following clause:

\[
DC(1,1) = \begin{cases} 
B_w = 0 \\
DC((1,1) \mod 2 = 0) \\
B_w = 1 \\
DC((1,1) \mod 2 = 1)
\end{cases}
\]

For this purpose, only one DC coefficient is required to insert the scrambled single watermark bit \(B_w\). If the selected bit of watermark is 1 and the value of DC component is odd then no change is required in this component otherwise make it even. Otherwise, if the selected bit of watermark bit is 0 and the DC value is even then, again, no change in DC component is required otherwise make it even. In this way, entire watermark bits inserted in equal number of DCT blocks. If still some DCT blocks are not utilized from the untouched candidate I-frames then the embedding process will not stop. However, it continues to pick the watermark bits from the scratch and inserting into the remaining DCT blocks. This step is necessary because the assailant cannot estimate the statistical relation between the number of DCT blocks and watermark bits to be used to extract the watermark information. Watermarked I-frame is regenerated by applying the inverse quantization followed by inverse DCT operation. Each watermarked frame is stored into the buffer to create the motion compensated image for B-frame or P-frame. In this way, all frames are combined in the same sequence in which they processed to get watermarked compressed video.

3.4. Watermark Extraction Process

The proposed approach is semi blind detection in which neither original video nor original watermark are required. The only requirement is to know about the three cryptographic keys \((r, c, \text{size})\). The estimation of watermark bits are based on analyzing the DC coefficients of the 8x8 DCT blocks from each candidate I-frames. The extraction process is started by picking the candidate I-frame from watermarked video and dividing it into three channels, one as luminance component Y and other two Chroma channels known as CB and CR respectively. Choose Y component and further dividing it into the number of equal size (8 x 8) of blocks for applying DCT and then further quantization operation. An extraction process of one watermark bit is based on testing the quantized DC coefficient to odd or even in the following manner.
\[ W_s = \begin{cases} 
0, & DC(1,1) \mod 2 = 0 \\
1, & \text{else} 
\end{cases} \quad (2) \]

If the \( DC(1,1) \) is odd then it is considered that 1 otherwise 0 as watermark bit is extracted and the extracted bit is added into the buffer. The buffer is incremented by 1 to extract another watermark bit from another adjacent DCT block in the same manner. Before incrementing the buffer, one more condition is to check whether the buffer size is equal to the value of third cryptography key (\textit{size}). If this clause is true then the extraction process will stop otherwise continue to extract the remaining watermark bits. In this way, if entire DCT blocks for the currently selected candidate I-frame are processed and still the buffer size is not equal to the key size then select next candidate I-frame to collect remaining watermark bits. After extracting all watermark bits, convert one dimensional buffer into two dimensional by using two keys \( r \) and \( s \). However, the recovered watermark is in encrypted form which is decrypted by using the columnar transposition algorithm [18] to get the plain text form of watermark image.

4. Experimental Results and Analysis

Experimental results have been evaluated by considering two static videos named as \textit{Akiyo.avi}, \textit{mother.avi} and one dynamic video as \textit{Foreman.avi}. The two videos are static since in \textit{Akiyo}, a newsreader is having only lips movements and rest part of the video including background is immovable throughout the playing of video and the same reason also applicable for another video \textit{Mother} in which a mother is describing about her child. The \textit{Foreman} video is dynamic in nature since foreground as well as background part both are movable at different point of time. The length of all these video sequences is 300 frames with a frame rate 25 fps and the size of each video frame is 176 x 144. The sample original videos are shown in Figure 3(a) and Figure 4(a) and their respective watermarked videos are shown in the Figure 3(c) and Figure 4(c) respectively. Binary watermarks considered for embedding are shown in the Figure 3(b) and Figure 4(b) and corresponding extracted watermarks are shown in the Figure 3(d) and Figure 4(d) respectively. NC obtained is 0.99343 and 0.97457 from the watermarked video \textit{Akiyo} and \textit{Foreman} respectively without any attack indicates that the scheme is highly robust for both types of video.

![Figure 3(a). 90th Original frame](image1)

![Figure 3(b). Original watermark](image2)

![Figure 3(c). 90th Watermarked frame](image3)

![Figure 3(d). Extracted Watermark](image4)

![Figure 4(a). 110th Original frame](image5)

![Figure 4(b). Original Watermark](image6)

![Figure 4(c). 110th Watermarked frame](image7)

![Figure 4(d). Extracted Watermark](image8)
4.1. Testing of Watermark Detection and Robustness Assessment

The robustness parameters calculate the degree of resemblance between original and extracted watermark. The normalized coefficient (NC) also named as Correlation Coefficient (CC):

$$ NC / CC = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} W(i, j)W'(i, j)}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} W(i, j)^2} $$

(3)

Where $W$ and $W'(i, j)$ are the original and extracted watermarks of size $M \times N$ each. If NC $> \mu$, it is alleged that the watermark is detected in the watermarked video otherwise it signals the false negative alarm indicates that this scheme fails to perceive the watermark in the video sequence. The value of $\mu = 0.70$ is chosen experimentally and considered to be perfect by performing exhaustive experiments consisting different watermarks and videos.

Another way to judge the robustness is to calculate the Bit Error Rate (BER). It reflects the degree of dissimilarities between the extracted and original watermark image. The proposed algorithm also tested the bit error rate (BER) as per the following formula:

$$ BER = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} W(i, j) \oplus W'(i, j)}{MN \times 100\%} $$

(4)

Where $W(i, j)$ and $W'(i, j)$ are the bits of the original and extracted watermark, respectively. $M \times N$ is the size of watermark image. NC and BER both the parameters evaluate the robustness but former find the similarities between the original and extracted watermark and later described the degree of dissimilarities. The numerical relationship between these two is that NC is inversely proportional to BER i.e. higher the NC value and lower the BER indicates the higher robustness is obtained. In a simple way, when NC=1 and BER=0 reflects that there is no viable difference between the original and the extracted watermark.

The proposed algorithm measured the robustness by applying different categories of attacks to Akiyo watermarked video sequence as shown in Table 1. These attacks are broadly classified in two categories: Spatial Synchronization and Temporal Synchronization Attacks. Spatial synchronization attacks further classified as geometric attacks includes rotation and cropping attacks; Noise attacks contains speckle, salt & pepper and Gaussian noise. Filter attacks covers median and wiener filter and data compression attack. Temporal synchronization attacks are the most noteworthy attacks based on video characteristics. It comprises frame insertion, frame deletion, frame replacement, frame swapping and frame averaging.

4.1.1. Spatial Synchronization Attacks

Spatial synchronization issue refers to the alteration of the coordinates of the embedded watermark by applying image processing attacks particularly geometric attacks like rotation and cropping in order to distort the watermarked signal by malicious attacker. Since realistic video watermarking applications always suggested that the toughness must be verified against these attacks therefore such issues cannot be disregard. The scheme tested the robustness by rotating each frame anticlockwise by 1°, 2°, 5° and 10°. It is observed that the proposed video watermarking scheme is able to survive up to 5° rotation of watermarked frames as 85% of the watermark is extracted. Other part of experiment is cropping each watermarked frames. The attacker cropped the frame size in such a manner that size of frame can never be altered and it can be done by substituting zeros in the cropped part of video frame. The frames are cropped by 10%, 20%, 30%, 40% and 50% in the corresponding first 15, 29, 44 and 58 and last 72 columns of each watermarked frame. Satisfactory perceptual qualities measures are obtained up to 40% frames are cropped. As robustness is concerned the watermark has been successfully extracted (70%) when the watermarked frames are cropped up to 45% as shown in the Figure 5.
The robustness is also tested by adding the different categories of noises accountable for the deformation and deprivation of the watermarked video. As mentioned in Table 1, these attacks reduces the perceptual quality of the watermarked video up to 11dB but the robustness results are acceptable for salt & pepper, speckle and Gaussian noise by adding the noise up to 0.003 with zero mean and 10% variance as shown in the Figure 6.

Table 1. Best Robustness results for static and dynamic videos

<table>
<thead>
<tr>
<th>Attack Type</th>
<th>Parameters Applied</th>
<th>Akiyo</th>
<th>Foreman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CC</td>
<td>BER</td>
<td>CC</td>
</tr>
<tr>
<td>Rotation</td>
<td>At 5°</td>
<td>0.85419</td>
<td>15.8419</td>
</tr>
<tr>
<td>Cropping 20%</td>
<td>20% frames</td>
<td>0.86561</td>
<td>12.8849</td>
</tr>
<tr>
<td>Speckle Noise</td>
<td>0.003</td>
<td>0.94040</td>
<td>5.78370</td>
</tr>
<tr>
<td>Salt &amp; Pepper</td>
<td>0.003</td>
<td>0.93446</td>
<td>6.34290</td>
</tr>
<tr>
<td>Gaussian Noise</td>
<td>Default</td>
<td>0.81850</td>
<td>17.2974</td>
</tr>
<tr>
<td>Frame Replacement</td>
<td>13 frames replaced (13-24) by Mother</td>
<td>0.94187</td>
<td>5.6534</td>
</tr>
<tr>
<td>Frame Insertion</td>
<td>13 frames inserted (13-24) by Mother</td>
<td>0.69888</td>
<td>29.8682</td>
</tr>
<tr>
<td>Frame Deletion</td>
<td>13 frames deleted (13-24)</td>
<td>0.90238</td>
<td>0.90238</td>
</tr>
</tbody>
</table>

4.1.2. Temporal Synchronization Attacks

The most essential characteristic of video sequence is the existence of temporal redundancy which do not likely to be present for still images. The proposed method takes advantage of this feature to spread the watermark repeatedly in selected I-frames. But due to this features, another issues arises in the form of temporal synchronization attacks like frame insertion, frame replacement, frame dropping, frame swapping and most crucial is the frame averaging also named as temporal collusion attacks. It is obvious that the dynamic composition of embedded watermark in watermarked video sequence may remove completely by applying the averaging of multiples frames together. These entire attacks are done on 10%, 15%, 20% and 25% frames on watermarked video as shown in the Figure 7 indicates that the proposed scheme carries the good robustness against all temporal synchronization attacks.
4.1.3. Data Compression Attack

The sturdiness obtained against different compression bit rate for all three video sequences is exposed in the Figure 8. It is obvious that the relation between NC and compression bit rate is almost linear for both static and dynamic types of videos. The average toughness (NC) obtained is 0.7985, 0.7257 and 0.7957 at their corresponding average compression bit rate are 1178, 1622 and 1285 respectively for all these video sequences. The watermarks are successfully extracted from all three videos because NC obtained is greater than the threshold value. The simulation results demonstrate that the sturdiness obtained for static video is higher than dynamic video.

The summarization of Table 1 indicates that the former is much finer as a carrier to conceal the watermark.

![Figure 8. NC Vs Compression bit rate](image)

5. Conclusion

In this paper, a new MPEG-2 based robust, imperceptible and semi-blind digital video watermarking is proposed. The key idea in this algorithm is to select the candidate I-frames for embedding the scrambled watermark. P-frames and B-frames and even motion vectors are untouched during the watermarking process illustrated that high perceptibility is obtained. Robustness is successfully verified by simulating various intentional and non-intentional attacks. This approach also focused on the payload capacity. Due to the availability of large amount of DCT blocks therefore thousand of watermark bits can be inserted into the single I-frame. In a nutshell, simulation results have been evaluated covering all three main contradictory issues as robustness, imperceptibility and payload capacity. The future work can also be extended to test the robustness by applying more attacks like ambiguity attacks, collusion attacks and joint attacks.

References


