Digital Video Watermarking : A Retrospective

Dixa Patel, Nehal Chitaliya, Manoj Pandya, Parth Trivedi, M.B.Potdar, Parul Thakkar

Abstract—Due to the ease of availability for multimedia content over the Internet, it is too easy for the owners to transfer the data (audio or video) over the internet, and hence it can be copied and stored easily without any loss in fidelity. These leads to serious problem of Copyright Infringement, Ownership Protection. Digital Watermarking is the technique to address such issues. This paper incorporates the detail study of Video watermarking techniques, various type of watermarking Attacks & Application.

Index Terms— Digital Multimedia, Fidelity, Spatial Domain, DCT, DWT, Robust

1 INTRODUCTION

In the past several years there has been an explosive growth in multimedia and network technology which leads to the necessity of developing some secure methods for legal distribution of the digital content. As digital multimedia works (video, audio and images) become available for retransmission, reproduction, and publishing over the Internet, a real need for protection against Lossless and unlimited duplication, arbitrarily modification, unauthorized uploads, illegal redistribution is increasing day by day. Digital watermarking is the technology that provides and ensures security, data authentication and copyright protection to the digital media. Digital watermarking is the embedding of secret information (i.e. Watermark) into the digital media (Image, audio or video) in such a way that any attacker can’t remove or replace it. Application for Digital watermarking involves Copyright protection, Fingerprinting, Broadcast monitoring, Video Authentication etc. For the protection of confidential and sensitive data from unauthorized access and tampering various methods for data hiding is used. Cryptography technique includes the encryption & decryption process. Encryption Process converts message into unreadable format which is known as Cipher text. Transmitted Cipher text can be decrypt into original message by only those for whom it is intended. So, Cryptography mainly provides security to message [18] [30] While steganography is mainly used to hide any secret information within the original message. Steganography is a combination of two words, which are taken from Greek Language: ‘stegano’ means cover and graphie means writing, so called as Cover writing. Secret information can only identify and extract by intended recipient [28]. Watermarking is the direct embedding of additional information (Watermark) into the original content or host signal in a way that any attacker cannot remove or replace it [19][4]. Steganography is useful when the protection of secret information is the required while watermarking is more useful to protect the original content from illegal use. The concept of cryptography is totally different then these approaches of data security. Digital content is encrypted at transmitter using a key and can be decrypted at receiver if and only if the correct key is available. Cryptography gives advantage only through the channel. Once encrypted content is decrypted using a key at receiver, no means of security is available for protecting digital content from copyright. Therefore, encryption must be replaced by some method which protects digital content after decryption and there concept of watermarking comes [18].

2 CLASSIFICATION OF WATERMARKING

Watermark can be classified according to host signal, perceptivity, robustness, and watermark type, Domain of watermark embedding, application and according to data required for watermark extraction. Fig.1 gives overall idea about watermark classification.

![Watermarking Classification Diagram](image-url)
2.1 Classification According To Host Signal
Watermarking technique can be classified depending on the media into which watermark is embedded. Text watermarking inserts a watermark in the font shape and the space between characters and line spaces. This has some disadvantages that one cannot detect the watermark in case of modulating the fonts.[7] Image watermarking embeds a special Information (Logo, Stamp, Label etc.) to an image and detects or extracts it later for ownership confirmation. This approach is most widely used and it utilizes the characteristics of human visual system.[19][4][25]. Video watermarking is an extension of image watermarking. This method requires real time extraction and robustness for compression. Video is a three dimensional signal having 2D in space (Image) and 1D in time. Time synchronization is one of the most important parameter which differs Video Watermarking from Image Watermarking. Interest in Audio watermarking increases because of the internet music, MP3. This approach needs watermark robustness and inaudibility like other cases.[18][2][16]

2.2 Classification According To Perceptivity
According to human visual perception, Watermark can be classified as Visible or Invisible. In visible watermarking, Watermark is transparent. As the watermark is visible to viewer, it cannot remove the copyright illegally. It confirms the authentication and mainly used in logo or trademark label. Disadvantage of visible watermark is its fragility to attack. Also, watermark embedding process degrades quality of original content in some cases. Another type of watermarking is Invisible which is mainly used as a proof of ownership. Invisible watermark is very robust as compared to visible.[18]

2.3 Classification According To Robustness
Watermarks need robustness to protect the ownership from various attacks. There are mainly three type of watermark: Robust, Semi-Fragile and Fragile. Robust watermark is very difficult to extract and so cannot remove or replace easily. Semi-fragile watermark is capable of tolerating some degree of the change to a watermarked frame, such as the addition of quantization noise from lossy compression. Fragile watermark is designed to be easily destroyed if a watermarked frame is manipulated in the slightest manner. This watermarking method is mainly used for the protection and the verification of original contents.[18]

2.4 Classification according to Watermark type
The watermark which is embedded into original multimedia content are of various type. Embedded Watermark can be Noise or Image. Noise can be randomly generated PN sequence, Gaussian Sequence or can be Chaotic sequence. PN sequence is widely used watermark having a good feature of self-correlation function. It is robust against cryptographic attack but weak for image processing, noise & compression attack. Gaussian sequence is a random generation with Gaussian distribution N(0,1). It is mainly used for extracting multiple inserted watermark. Chaotic sequence is very easy to generate. New chaotic sequence can be generate by simply modifying initial values. Image watermark can be Binary Image, logo, Stamp or Label. Visual degradation of original watermark is possible to observe once it is extracted. [18]

3 GENERAL FRAMEWORK OF VIDEO WATERMARKING

As shown in Fig 2, The general framework of watermarking contains three main components: Embedding, attacker, and Extraction. The embedding algorithm accepts the multimedia signal I, and the watermark, and generates a watermarked signal Iwm. The watermarked signal is then communicated to another person via the communication channel. If third person or the communicated one, makes any changes whether malicious or not, then it is termed as an attack. But the term is mainly used when intruder attempts to remove the watermark by changing the contents of multimedia by using common signal processing operations or some intentional attacks like adding noise to the multimedia content. Detecting/extracting algorithm is used by detector/extractor which is then applied over the attacked signal I’wm to detect the presence or extract the watermark from it. If the signal does not undergo any change during transmission, then the watermark will be present and can be extracted easily. The extracted signal is represented as I’[25]. Attacks can be intentional or unintentional. The unintentional attacks involve lossy copying, Compression of video, Change of frame rate or Change of Resolution. The intentional watermark attack include various frame attacks like filtering attacks, Geometrical attack and noise adding attack.[16] Table 1. Shows various type of intentional attack with examples [15][20].

Table 1: Various type of attack.

<table>
<thead>
<tr>
<th>Types Of Attack</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise Attack</td>
<td>Salt &amp; Pepper Noise</td>
</tr>
<tr>
<td></td>
<td>Gaussian Noise</td>
</tr>
<tr>
<td></td>
<td>Speckle Noise</td>
</tr>
<tr>
<td>Geometrical Attack</td>
<td>Rotation</td>
</tr>
<tr>
<td></td>
<td>Scaling</td>
</tr>
<tr>
<td></td>
<td>Translation</td>
</tr>
<tr>
<td></td>
<td>Resizing</td>
</tr>
<tr>
<td></td>
<td>Cropping</td>
</tr>
<tr>
<td>Attacks Specific to Video</td>
<td>Frame Dropping</td>
</tr>
<tr>
<td></td>
<td>Frame Averaging</td>
</tr>
<tr>
<td></td>
<td>Frame Swapping</td>
</tr>
<tr>
<td></td>
<td>Frame Insertion</td>
</tr>
<tr>
<td></td>
<td>Compression</td>
</tr>
</tbody>
</table>
3.1 Properties Of Digital Watermark

These are some parameters needs to consider for implementation of watermarking techniques are :

1. Transparency
2. Security
3. Ease of embedding and retrieval
4. Robustness
5. Effect on bandwidth
6. Interoperability

1. Transparency:
The most fundamental requirement for any Watermarking method shall be such that it is transparent to the end user. The watermarked content should be consumable at the intended user device without giving annoyance to the user. Watermark only shows up a watermark-detector device[18][30].

2. Security:
Watermarked information shall only be accessible to only authorized parties. They only have the right to alter the Watermark content. Encryption can be used to prevent unauthorized access of the watermarked data[18][30].

3. Ease of embedding and retrieval:
Ideally, Watermarking on digital media should be possible to be performed on the fly. The computation needed for the selected algorithm should be least[18].

4. Robustness:
Watermarking must be robust enough to withstand all kinds for signal processing operations attacks or unauthorized access. Any attempt, whether intentionally or unintentionally, that has a potential to alter the data content is considered as an attack. Robustness against attack is a key requirement for Watermarking and the success of this Technology for copyright protection depends on its stability against attacks[18].

5. Effect on bandwidth:
Watermarking should be done in such a way that it does not increase the bandwidth required for transmission. If Watermarking becomes a burden for the available bandwidth, the method fails.[18]

6. Interoperability:
Digitally watermarked content shall still be interoperable so that it can be seamlessly accessed through heterogeneous networks and can be played on various plays out devices that may be aware or unaware of watermarking techniques.[18]

4 VARIOUS VIDEO WATERMARKING TECHNIQUE

Current video watermarking techniques can be classified into two major types: spatial-domain watermarking techniques and Frequency domain watermarking techniques.

As shown in Fig.3, In Spatial-domain techniques watermark is embedded in the frames of given video by modifying its pixels directly. These techniques are easy to implement and require few computational resources but not much robust against attacks. In case of transform-domain techniques, first step is to implement any transform (Fourier, Cosine, Wavelet etc.) on extracted Video Frames and then watermark can be embed by modifying the coefficients of the transformed video frames. Compared to spatial-domain techniques, frequency-domain watermarking techniques proved to be more effective with respect to achieving the imperceptibility and robustness requirements of digital watermarking algorithms.[16]

4.1 Spatial Domain Techniques

Spatial domain digital watermarking algorithms directly load the raw data into the original video frame [13]. Spatial watermarking technique embed the watermark in only one of the colour band (either Red, Green or Blue Channel) of video frame using colour separation. So, it is difficult to detect embedded watermark under regular viewing [9]. Some of its main algorithms are as discussed below:

1. Least Significant Bit (LSB)
The type of inserted watermark is Image for LSB method. For embedding the watermark, a subset of image pixels is selected and then Least Significant Bit (LSB) of each of the selected pixels is substituted by watermark bits. The watermark may be spread throughout the frame or in the selected locations of the frame. Advantage of this method is that it is easy to implement and does not generate serious distortion to the video; however, it is vulnerable to attacks and watermark can be easily destroyed.
2. **Patchwork Algorithm**

Patchwork is a data hiding technique developed by Bender et alii and published on IBM Systems Journal, 1996[3]. It is based on a pseudorandom, statistical model. It works by invisibly embedding a specific statistic, with a Gaussian distribution, into the host image. Two sets of pixel, or patches, of the image are chosen, the first A and the second B. Then the algorithm works by slightly brightening points in A, while darkening of the same factor those in B. One of the most important characteristics of patchwork is its resistance to cropping and to gamma and tone scale corrections [27]

3. **Correlation-Based Technique**

In Correlation based technique, type of watermark used is Noise and Image both. A pseudo-random noise (PN) pattern is added to the extracted video frame. In this technique, there is a major trade-off between robustness and quality of watermarked image. Watermark can be extract by generating the same pseudo-random noise pattern at the receiver end, and using correlation technique, similarity between the extracted watermark and original watermark is computed. [16]

4.2 **Frequency domain Technique**

Frequency-domain methods are more widely used as compared to spatial-domain methods. Here, the watermark is embedded in the spectral coefficients of the original video frame. the Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), and Discrete Wavelet Transform (DWT) are widely used transforms. The reason for watermarking in the frequency domain is that the characteristics of the Human Visual System (HVS) are better captured by the spectral coefficients [10][12][15]. Some of its main algorithms are discussed below:

1. **Discrete Cosine Transforms (DCT):**
   DCT like a Fourier Transform, represents data in terms of frequency space rather than an amplitude space. DCT based watermarking techniques are robust compared to spatial domain techniques. Such algorithms are robust against simple image processing operations like low pass filtering, brightness and contrast adjustment, blurring etc. However, they are difficult to implement and are computationally more expensive. At the same time they are weak against geometrical attacks like rotation, scaling, cropping, translation etc. DCT domain watermarking can be classified into Global DCT watermarking and Block based DCT watermarking. Embedding in the perceptually significant portion of the image has its own advantages because most compression schemes remove the perceptually insignificant portion of the image [16][31].

2. **Discrete wavelet transforms (DWT):**
   Wavelet Transform is a modern technique frequently used in digital image processing, compression, watermarking etc. The transforms are based on small waves, called wavelet, of varying frequency and limited duration. The wavelet transform decomposes the image into three spatial directions, i.e. horizontal, vertical and diagonal. Hence wavelets reflect the anisotropic properties of HVS more precisely. Magnitude of DWT coefficients is larger in the lowest bands (LL) at each level of decomposition and is smaller for other bands (HH, LH, and HL) [9]. Robustness of embedded watermark can be increase at the cost of visible distortion [9]. However, DWT is much preferred because it provides both a simultaneous spatial localization and a frequency spread of the watermark within the host image. [5].DWT is more advantageous than DCT as it is multi-resolution transform and hence provides resolution from lower level to higher level which facilitates various levels to embed the watermark as compared to DCT. Also, Wavelet transform provides better HVS than the DCT. [8] But Computational complexity & hence time complexity of DWT is more compared to DCT and it is more time consuming for video watermarking. As Feig (1990) pointed out it only takes 54 multiplications to compute DCT for a block of 8x8, unlike wavelet calculation depends upon the length of the filter used, which is at least 1 multiplication per coefficient [8]

3. **Discrete Fourier transform (DFT):**
   Transforms a continuous function into its frequency components. It has robustness against geometric attacks like rotation, scaling, cropping, translation etc. DFT shows translation invariance. Spatial shifts in the image affects the phase representation of the image but not the magnitude representation, or circular shifts in the spatial domain don't affect the magnitude of the Fourier transform. DFT is more robust against geometrical attacks as compared to DWT and DCT because DFT is rotation, scaling and translation (RST) invariant. The spatial domain techniques, DCT and the DWT are not RST invariant and hence it is difficult to overcome from geometric distortions. [8]

4. **SVD Domain Video Watermarking Technique:**
   Singular value decomposition is the multivariable and numerical technique which decomposes the input data into three sub matrices, Left singular vector matrix, Right singular vector matrix, and Diagonal matrix. Diagonal elements of last matrix are singular values of the given signal known as Eigenvalues. Eigenvalues represent the energy of the signal. Because of translation, scaling properties of SVD it can be used as a tool to develop watermarking schemes [31].
5. **Contourlet Transform:**

Contourlet transform uses Pyramidal Directional Filter Bank structure, which represents an image, multi-resolutionally and multi-directionally to capture the contours and fine details for image analysis. This Pyramidal Directional Filter comprises of Laplacian Pyramid (LP) and Directional Filter Bank (DFB) Structure. Laplacian Pyramid used to capture the discontinues at edge point and Directional Filter Bank links these point discontinues into linear structure and results in a contour segments and thus this process called as Contourlet Transform. This can be carried up to multiple levels [24]. The two-dimensional singularities found in images cannot be captured easily by wavelets. Therefore, contourlet transforms is used to achieve directionality and multi resolution for image signals. The contourlet transform efficiently represents contours and textures are in an images [24].

5 **APPLICATION SCENARIO OF DIGITAL VIDEO WATERMARKING**

Figure 4 sketches the basic scenario for which the watermarking algorithm was designed. A video stream acquired using a digital device is subsequently processed in the TV production chain and finally broadcast as an MPEG-2 bit stream. The digital data is marked directly after recording and before entering the production chain. The data is not necessarily recorded in a closed studio environment; it can be recorded in the field and subsequently transmitted to the broadcaster’s headquarters. The raw material is cut and assembled during the production process. Other parties can digitally record the data and broadcast it again instantly, or with some delay, as a digital video stream without any loss of quality. The watermark monitoring process can check for the watermark and automatically detect unauthorized use of copyrighted material.[2]

### Table 2. Video watermarking: Application & Associated purpose[7]

<table>
<thead>
<tr>
<th>Application</th>
<th>Purpose of the Embedded watermark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy Control</td>
<td>Prevent unauthorized copying</td>
</tr>
<tr>
<td>Broadcast Monitoring</td>
<td>Identify the video item being broadcasted</td>
</tr>
<tr>
<td>Fingerprinting</td>
<td>Trace back a malicious user</td>
</tr>
<tr>
<td>Video Authentication</td>
<td>Insure that the original content has not been altered</td>
</tr>
<tr>
<td>Copyright Protection</td>
<td>Prove Ownership</td>
</tr>
<tr>
<td>Enhanced Video Coding</td>
<td>Bring additional information E.g. for error correction</td>
</tr>
</tbody>
</table>

6 **PERFORMANCE EVALUATION PARAMETERS**

In order to evaluate the performance of the watermarked frames, there are some quality measures such as SNR, PSNR, MSE, and NC.

The MSE (mean square error) is defined as average squared difference between a reference video frame and watermarked frame. It is calculated by the formula given below

\[
MSE = \frac{1}{XY} \sum_{i=1}^{X} \sum_{j=1}^{Y} (I(i,j) - I'(i,j))^2
\]

In eq.1, X and Y are height and width respectively of the video frame, I (i, j) is the pixel value of the watermarked frame and I’(i, j) is the pixel value of the original frame. [22]. Ideally, Value of MSE should be minimum.

SNR (Signal to Noise ratio) measures the sensitivity of the imaging. It provides the strength of signal against the background noise. It is calculated by the formula given below, [9]

\[
SNR_{dB} = 10 \log \left( \frac{P_{signal}}{P_{noise}} \right)
\]

The PSNR (peak signal to noise ratio) is used to determine the degradation in the watermarked video. It is calculated by the formula as

\[
PSNR = 10 \log \left( \frac{L^2}{MSE} \right)
\]

In eq. 3 , L is the peak signal value of the watermark image which is equal to 255 for 8 bit images. [3] Ideally, PSNR value should be in between 40 to 60. Higher value of PSNR indicates that quality of watermarked video is maintained as original one.
NC (Normalized correlation) is the similarity measurement of the extracted and referenced watermarks which is used for objective judgment of extraction fidelity [12]. Ideally, Value of NC should be 1 which indicates exact similarity between extracted and referenced watermark.

7 CONCLUSION

The purpose of this paper is to evaluate different techniques for video watermarking. Watermark in the video can be embed using spatial or frequency domain technique. Frequency-domain watermarking techniques are more efficient for imperceptibility and robustness requirements but time consuming while Spatial-domain techniques are easy to implement and less time complex. Digital watermarking for real time application like video broadcasting and monitoring demands less time consuming techniques. So, Spatial Domain techniques are more useful for such application.

ACKNOWLEDGMENT

Firstly, I would like to thank Dr. Nehal Chitaliya from Sardar VallabhBhai Patel Institute of Technology, Vasad for her aspiring guidance; constructive criticism and valuable Suggestion for my ongoing research work, as well as to publish this review paper. She kindly read this review paper and offered valuable detailed suggestion on grammar, organization, and the theme of the paper.

Secondly, I would like to thank Dr. M.B. Potdar who allowed me to carry out the research work in Bhaskaracharya Institute For Space Applications and Geo-Informatics (BISAG), Gandhinagar. I would also like to give my sincere gratitude to Dr. Manoj Padya and Mr. Parth Trivedi from BISAG for their exemplary guidance, valuable feedback, constant encouragement and for sharing their truthful and illuminating views on a number of issues related to the research work.

Finally, I sincerely thank to my parents, family, and friends for their continuous support. The product of this research paper would not be possible without all of them.

REFERENCES


