Video Watermarking Algorithm Based on Relative Relationship of DCT Coefficients

Cheng Mingzhi, Du Yanping, Wang Yan
Beijing Institute of Graphic Communication, Beijing, China
Email: chengmz@bigc.edu.cn, duyanping@bigc.edu.cn, wangyanzi@bigc.edu.cn

Xi Minchao, Yuan Kaiguo
School of Computing, Beijing University of Posts and Telecommunications, Beijing, China
Email: minchao@gmail.com, flyingdreaming@gmail.com

Abstract—A video watermarking algorithm based on relative relationship of DCT coefficients is proposed in this paper. By modification of coefficients on chosen positions, the robustness of watermark embedding is ensured. Meanwhile, as the relationship of coefficients remains unchanged, the quality of video is not affected. The analysis results shows that the performance of proposed algorithm is better than the typical algorithms.

Index Terms—Video Watermark; Relative Relationship; DCT; BCH; Arnold Transformation

I. INTRODUCTION

In recent decade, as multi-media information become more and more digitalized, technology of video digitalization becomes rapidly advanced in way of information inquiry, presentation, storage and transfer. However, issues have been raised simultaneously with the most phenomenon one as intellectual property theft. The cause of this issue should be primarily blamed on the low limitation of duplication.

Watermarking technology become more often used in the area of copyright protection, with this method, original video owner could insert the property identification information into the video. The video file could still be used or transferred normally. The originality of video is always traceable, intellectual property is as a result being protected.

Although watermarking technology cannot prevent illegal duplication of video data, authors and publishers could still track the source of unauthorized video, and stop the illegal behavior as soon as possible.

Similar with information hiding techniques, watermark algorithms work by embedding a encoding signal into the video file. The signal could generally include the owner identification, publishing dates and serial number. All the information is used to track the status of publishing and usage of protected videos [1] [2].

Video watermarks can be either visible or invisible. The logo of authors, photocopy of library books etc. always showed as the visible watermark. Invisible watermarks in the other way, does not affect the video appearance, copyright information is coded inside without other influence. According to the preservation of video quality, invisible watermarks are more commonly used [2].

Video watermarking algorithms could also be classified into two categories: watermark of spatial domain and watermark of frequency domain. Spatial domain watermarking algorithms embed the watermark by directly modifying the pixel values of the host video, while frequency domain watermarking algorithms embed watermark by modulating the frequency coefficients. The latter one is more robust when suffering ordinary video manipulations. Furthermore, among the several frequency domains-DCT, DWT, DFT, algorithms based on DCT domain are more capability in digital video codec standards, so it draws majority attentions of researchers [3].

In this paper, the algorithm we are presenting modifies coefficients without changing their relative relationship, while the former algorithms tends to change coefficients arbitrarily, causing problem in the quality of video. In fact, video quality could be even better, as the characteristics of the image data have been taken advantage.

The main contributions of this paper are as follows. We reveal the principle of common watermarking algorithms: storing binary information in the host data using binary relationship. And we proposed a new video watermarking algorithm storing information by a different binary relationship which improves the video quality.

The rest of this paper is organized as follows. Definitions of DCT and BCH code are introduced in section II. Principle and procedures of the proposed watermark algorithm are discussed in detail in section III. In section IV, performances of our algorithm are evaluated and analyzed. Finally, conclusions and future works are raised in section V.

II. RELATED WORK

A. DCT and Video Codec Standards

DCT, short for Discrete Cosine Transform, is widely used in popular video compression standards, such as
H.264 and MPEG4. The purpose of DCT transform in video codec is to convert image or motion-compensated residual data into another domain, the frequency, or more commonly called, transform domain. The choice of transform depends on a number of criteria:

1) Data in the transform domain should be decorrelated, i.e. separated into components within minimal inter-dependence, and compact, i.e. most of the energy in the transformed data should be concentrated into a small number of values.

2) The transform should be reversible.

3) The transform should be computationally tractable, e.g. less memory requirement, achievable using limited-precision arithmetic, fewer arithmetic operations, etc.

There are plenty of video compression algorithms and the most popular transforms tend to fall into two categories, block-based and image-based. Examples of block-based transforms include the KLT (Karhunen-Loeve Transform), SVD (Singular Value Decomposition), and the ever-popular DCT (Discrete Cosine Transform). Each of these operates on blocks of M * N image or residual samples and hence the image is processed in units of a block. Block transforms have less memory requirements and are well suited to compression of block-based motion compensation residuals while suffering from artifacts at block edges, however. Image-based transforms operate on an entire image or frame or a large section of the image known as a ‘tile’. The most popular image transform is the DWT (Discrete Wavelet Transform), or ‘wavelet’. Image transforms such as the DWT have been shown to out-perform block transforms for still image compression but they tend to have higher memory requirements because the whole image or tile is processed as a unit and they do not necessarily fit well with block-based motion compensation which is the most important and effective operation of video compression. The approximations to the DCT are featured in H.264 for still image compression but they tend to have higher precision arithmetic, fewer arithmetic operations, etc.

The Discrete Cosine Transform operates on X, a block of M * N samples, typically image samples or residual values after prediction, to create Y, an M * N block of coefficients. The action of the forward DCT of a sample block is given by [4]:

\[ Y(u,v) = a(u,v) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} X(i,j) \cos \left( \frac{(2i+1)\pi}{2M} \right) \cos \left( \frac{(2j+1)\pi}{2N} \right) \]

Where \( Y \) stands for the DCT coefficient matrix of \( X \), in which \( Y(0,0) \) is called DC coefficient, and others are mentioned as AC coefficients.

The inverse DCT of a sample block is given by:

\[ X(i,j) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} a(u,v) Y(u,v) \cos \left( \frac{(2i+1)\pi}{2M} \right) \cos \left( \frac{(2j+1)\pi}{2N} \right) \]

\( i = 0,1,2,...,M-1, \quad j = 0,1,2,...,N-1. \)

B. BCH Code

In coding theory, BCH code is one of the cyclic error-correcting codes that are constructed using finite fields. BCH code is invented in 1959 by Hocquenghem, Bose, Chaudhuri independently. One of the key features of BCH codes is that during code design, there is a precise control over the number of symbol errors correctable by the code. In particular, it is possible to design binary BCH codes that can correct multiple bit errors. Another advantage of BCH codes is the ease of use as they can be decoded, namely, via an algebraic method known as syndrome decoding. This simplifies the design of the decoder for these codes, using small low-power electronic hardware. With all these advantages, BCH code is widely used in modern communication systems.

Given \( t \) as the number of errors designs to correct, \( m(x) \) be the minimal polynomial, \( d \) stands for the code minimal distance, where \( d \geq d_0 = 2t + 1 \), BCH code has the length of \( n = 2^m - 1 \), whose generator polynomial is [

\[ g(x) = LCM \left( m_1(x), m_2(x),...,m_{2t+1}(x) \right) \]

III. PROPOSED METHOD

A. Binary Relative Relationship

The storage of watermarking information in video frames usually depends on the relationships of coefficients, and most of which are binary relationship. Either watermarking algorithms based on energy modulation [4], or the ones depend on correlation detection and spread spectrum, the most basic principle could be expressed as following [1][6-10].

As the basic information unit in digital storage, binary relationship is the most basic relationship, as the formula below:

\[ W = R(x_1, x_2) \]

Where \( x_1, x_2 \) are the values based on image contents, or const values which are pre-defined.

For instance, a kind of watermarking algorithm based on energy modulation is proposed in paper [4]. In which the watermark information are kept on the binary relationship of the coefficient energy in each block and the average energy of the whole image. This could be represented as this formula, which has represented a kind of binary relationship:

\[ W_i = R(p_i, \overline{p}) = \begin{cases} 1, & \text{if } p_i \geq \overline{p} \\ 0, & \text{if } p_i < \overline{p} \end{cases} \]

Where \( W_i \) stands for the \( i \) th bit of watermark information, \( p_i \) expresses the energy of \( i \) th block, and \( \overline{p} \) is the average energy of blocks of the whole image.

Another kind of example algorithms comes from articles [1][6-10], which proposed watermark algorithms based on spread spectrum and correlation detection. Within these algorithms, watermark information are stored by this formula, which also represents a kind of binary relationship:

\[ W = R(x_1, x_2) \]
\[ W_i = R(S, S_w) = \begin{cases} 1, & \text{if } \text{Corr}(S, S_w) \geq \text{Threshold} \\ 0, & \text{if } \text{Corr}(S, S_w) < \text{Threshold} \end{cases} \] (7)

Where \( S \) stands for the coefficient sequence, \( S_w \) is the detection sequence generated by watermark information.

Based on the principles mentioned before, a watermark algorithm based on binary relationship of three coefficients is proposed, which stores watermark information by modulating the distances between coefficients. Compared with other algorithms which modified the coefficients arbitrarily, this algorithm changes the coefficients when the relationship of which is kept, and takes advantage of the characteristics of image data, which could improves the video quality.

The proposed algorithm could be expressed in the following formula:

\[ W_i = R[D_{left}, D_{right}] = \begin{cases} 0, & \text{if } D_{left} \leq D_{right} \\ 1, & \text{if } D_{left} > D_{right} \end{cases} \] (8)

\[ D_{left} = \text{Coef}_{left} - \text{Coef}_{mid} \]
\[ D_{right} = \text{Coef}_{right} - \text{Coef}_{mid} \]

It should be emphasized that \( \text{Coef}_{left} \), \( \text{Coef}_{mid} \), and \( \text{Coef}_{right} \) are not coefficients whose positions are fixed. They are a ordered coefficient sequence from three predefined positions, where \( \text{Coef}_{left} \leq \text{Coef}_{mid} \leq \text{Coef}_{right} \).

The proposed algorithm keeps the relationship of three coefficients unchanged while embedding the watermarks, which means the characteristic of video image data are sustained, and excess modifications are avoid.

**B. Watermark Embedding**

Given a video of size \( M \times N \). The video frame are divided into blocks of 8*8, which are expressed as \( D_i(x,y) \), \( x, y = 1,2,3,...,8 \), \( i = 1,2,3,...,L \), \( L = \lceil M/8 \rceil \lceil N/8 \rceil \). For each block, a block of DCT coefficients is generated, marked as \( C_i(u,v) \), \( u, v = 1,2,3,...,8 \), \( i = 1,2,3,...,L \).

To generate watermark information, a \( H \times H \) binary bitmap is converted to bit stream \( W \). Then, encode \( W \) with a BCH encoder by the parameters \( N = 63, K = 18 \), a bit stream of \( W' \) is generated.

Embed \( W' \) into each 8*8 DCT coefficient block \( C_i(u,v) \) bit by bit using the following steps.

a) Select three coefficients from \( C_i \), sort them into \( \text{Coef}_{left} \leq \text{Coef}_{mid} \leq \text{Coef}_{right} \).

b) Calculate the distance between \( \text{Coef}_{left} \) and \( \text{Coef}_{right} \), \( D_{mid} = \text{Coef}_{right} - \text{Coef}_{mid} \). To ensure the robustness of algorithm, a threshold \( \text{Threshold} \) is defined to indicate the minimal distance of \( \text{Coef}_{left} \) and \( \text{Coef}_{right} \).

If \( D_{mid} < \text{Threshold} \), then the algorithm modifies the values of \( \text{Coef}_{left} \) and \( \text{Coef}_{right} \), to increase the distance of them to \( \text{Threshold} \), as following:

\[ \text{Coef}_{left} = \text{Coef}_{left} + \frac{\text{Threshold} - D_{mid}}{2} \]
\[ \text{Coef}_{right} = \text{Coef}_{right} + \frac{\text{Threshold} - D_{mid}}{2} \] (9)

If \( D_{mid} \geq \text{Threshold} \), then will not be modified, that is :
\[ \text{Coef}_{left} = \text{Coef}_{left} \]
\[ \text{Coef}_{right} = \text{Coef}_{right} \]

b) Embed \( W' \) into \( C_i \), to get \( C'_i \), where \( W'_i \in \{0,1\} \).

If \( W'_i = 0 \), then do the following operation:
\[ \text{Coef}_{mid} = \text{Coef}_{left} + \frac{D_{mid}}{\text{Scale}} \] (10)

If \( W'_i = 1 \), then do the following operation:
\[ \text{Coef}_{mid} = \text{Coef}_{right} - \frac{D_{mid}}{\text{Scale}} \] (11)

Where \( \text{Scale} \geq 3 \) is a predefined const. This const number stands for the distance of \( \text{Coef}_{mid} \) and \( \text{Coef}_{left} + \text{Coef}_{right} \). The bigger the value is, the stronger the watermark signals are embedded.

d) Convert each embedded coefficient block \( C'_i \) to spatial domain- a embedded pixel block \( D'_i \).

Compose all the embedded blocks into a video frame \( I' \).

**C. Watermark Detection**

Just like the embedding procedure, given a suspicious video of size \( M \times N \). The video frame are divided into 8*8 blocks \( D'_i(x,y) \), \( x, y = 1,2,3,...,8 \), \( i = 1,2,3,...,L \), \( L = \lceil M/8 \rceil \lceil N/8 \rceil \). For each block, a block of DCT coefficients is generated, marked as \( C'_i(u,v) \), \( u, v = 1,2,3,...,8 \), \( i = 1,2,3,...,L \).

Detect watermark information \( W'E \) from each 8*8 DCT coefficient block \( C'_i(u,v) \) bit by bit using the following steps.

a) Select the three coefficients used in the embedding procedure. Then sort them to generate the coefficient sequence \( \text{Coef}_{left} \leq \text{Coef}_{mid} \leq \text{Coef}_{right} \).

b) Calculate \( D_{left} = \text{Coef}_{left} - \text{Coef}_{mid} \) and \( D_{right} = \text{Coef}_{right} - \text{Coef}_{mid} \), then each bit of \( W'E \) is generated.

If \( D_{left} \leq D_{right} \), then \( W'E' = 0 \).

If \( D_{left} > D_{right} \), then \( W'E' = 1 \). (12)

To generate watermark information, a \( H \times H \) binary bitmap is converted to bit stream \( W' \). Then, encode \( W' \) with a BCH encoder by the parameters \( N = 63, K = 18 \), a bit stream of \( W' \) is generated.

Decode \( W'E \) with a BCH decoder by the parameters \( N = 63, K = 18 \), a bit stream \( W'E \) is generated. Then, \( W'E \) is converted to a \( H \times H \) binary bitmap, as the extracted watermark image.
IV. EXPERIMENTAL RESULTS

A. Algorithm Parameters

Parameters $\text{Threshold}=40$ and $\text{Scale}=6$ are selected to evaluate the performances of the algorithm.

Four test videos are chosen to evaluate the algorithm: flower (352*288), foreman (352*288), mother (352*288), waterfall (352*288).

A 16*16 binary bitmap is used as the watermark image, as following:

B. Embed Performances

Some of the embedded video frames are showed in Figure 1 to Figure 8.

Figure 1. Original frames in flower_cif.yuv. PSNR= 35.6507.

Figure 2. Watermarked frames in flower_cif.yuv. PSNR= 35.6507.

Figure 3. Original frames in foreman_cif.yuv. PSNR= 38.5698.

Figure 4. Watermarked frames in foreman_cif.yuv. PSNR= 38.5698.

Figure 5. Original frames in mother_cif.yuv. PSNR= 39.0091.

Figure 6. Watermarked frames in mother_cif.yuv. PSNR= 39.0091.

Figure 7. Original frames in waterfall_cif.yuv. PSNR= 39.6570.

Figure 8. Watermarked frames in waterfall_cif.yuv. PSNR= 39.6570.
Figure 8. Watermarked frames in waterfall_cif.yuv. PSNR~39.6570.

C. Extract Performances

This paper evaluates the algorithm robustness against JPEG compression (With the compression strengths of 35, 40, 50, 60, 70, 80, 90, 100) and video zooming attacks (With the scales of 50%, 75%, 90%, 110%, 150%, 200%). The PSNR between embedded video frames and original frames are calculated, and NC values between extracted watermark image and original one are also evaluated.

PSNR (Peak Signal to Noise Ratio) is used to evaluate the visual quality after watermark embedding, larger value means the better visual quality. Definition is listed below:

$$\text{PSNR} = 10 \cdot \log \left( \frac{255^2}{M \cdot N \cdot \sum_{i=1}^{M} \sum_{j=1}^{N} (I(i,j) - I'(i,j))^2} \right)$$  \hspace{1cm} (13)

Where $I$ is the original video frame, and $I'$ means embedded frame.

NC (Normalized Mutual Correlation Coefficient) is the standard to evaluate the similarity between extracted watermark image and original one, the larger the better.

$$\text{NC} = \frac{\sum_{i=1}^{M} W(i)WE(i)}{\sqrt{\sum_{i=1}^{M} W(i)^2} \cdot \sqrt{\sum_{i=1}^{M} WE(i)^2}}$$  \hspace{1cm} (14)

The evaluate results are given in table I.

<table>
<thead>
<tr>
<th>TABLE I. EXTRACTED WATERMARK RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>flower</td>
</tr>
<tr>
<td>PSNR</td>
</tr>
<tr>
<td>JPEG_30</td>
</tr>
<tr>
<td>JPEG_40</td>
</tr>
<tr>
<td>JPEG_50</td>
</tr>
<tr>
<td>JPEG_60</td>
</tr>
<tr>
<td>JPEG_70</td>
</tr>
<tr>
<td>JPEG_80</td>
</tr>
<tr>
<td>JPEG_90</td>
</tr>
<tr>
<td>JPEG_100</td>
</tr>
<tr>
<td>RESC_50</td>
</tr>
<tr>
<td>RESC_75</td>
</tr>
<tr>
<td>RESC_90</td>
</tr>
<tr>
<td>RESC_110</td>
</tr>
<tr>
<td>RESC_150</td>
</tr>
<tr>
<td>RESC_200</td>
</tr>
</tbody>
</table>

D. Algorithm Comparison

In [11], a typical DCT watermark algorithm mentioned before is evaluated. With the proposed algorithm, results are showed in Figure 9 and Figure 10.

The comparison results between the typical algorithm and the proposed algorithm are given in table II.

Compared with the typical algorithm, the robust of proposed algorithm against JPEG compression and rescaling is improved, and the visual quality of embedded video is better.
Figure 9. Original image of Lena.bmp. PSNR= 38.8547

Figure 10. Watermarked image of Lena.bmp. PSNR= 38.8547.

### TABLE II. COMPARISON OF THE PROPOSED METHOD WITH THE TYPICAL ALGORITHM

<table>
<thead>
<tr>
<th>Lena</th>
<th>Proposed Alg.</th>
<th>Typical Alg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>38.8547</td>
<td>35.761</td>
</tr>
<tr>
<td>JPEG_30</td>
<td>0.8565</td>
<td>0.7792</td>
</tr>
<tr>
<td>JPEG_40</td>
<td>0.9031</td>
<td>0.9115</td>
</tr>
<tr>
<td>JPEG_50</td>
<td>0.9967</td>
<td>0.9868</td>
</tr>
</tbody>
</table>

### V. CONCLUSION

In this paper, a video watermark algorithm in DCT domain based on binary relationship of three coefficients is proposed, which stores watermark information by modulating the distances between coefficients. The processes of this watermarking scheme, including embedding detection, are described in detail. The evaluation has proved that our method can survive compression and rescale attacks, and further improvements could be made to get better performance. Improvements could be made from following aspects:

1. Increasing the number of embedding positions.
2. Scrambling the watermark image before embedding to improve the robustness.
3. Evaluating the robustness of coefficients in different positions.

### ACKNOWLEDGMENT

This work was supported in part by a grant from Beijing Natural Science Foundation under Grant No. 4122026, Scientific Research Common Program of Beijing Municipal Commission of Education under Grant KMZ01210015007, KM201210015006 and KZ201210015015, BJOC Project No. E-a-2012-25, E-6-2012-28, E-a-2012-27 and E-b-2013-26, Beijing University of Posts and Telecommunications under Grant 2012RC0217.

### REFERENCES


Mingzhi Cheng was born in 1974 in Hubei China. He received his B.S. degree in Communication Engineering from Chongqing University of Posts and Telecommunications, Chongqing, China, in 1996, and received his Ph.D. degree in Cryptography from Beijing University of Posts and Telecommunications, Beijing, China, in 2010. From 2010, Dr. Cheng is working in Packaging Anti-counterfeiting and Digital Content Security in Beijing Institute of Graphic Communication as a lecturer.

© 2013 ACADEMY PUBLISHER
Yanping Du was born in 1971 in Shanxi China. She received her B.S. degree in computer science and technology from Shanxi University, Shanxi, China, in 1994, and received her Ph.D. degree in control theory and engineering from Beijing Jiaotong University, Beijing, China, in 2005.

From 2006, Dr. Du is working in Packaging Anti-counterfeiting and Digital Content Security in Beijing Institute of Graphic Communication as a professor.

Yan Wang was born in 1975 in Hubei China. She received her B.S. degree in high voltage technology from Huazhong University of Science and Technology, Hubei, China, in 1996, and received her Ph.D. degree in weapon system and application engineering from Beijing Institute of Technology, Beijing, China, in 2005. Her research topic is electronic information and signal processing.

From 2006, Dr. Wang is working in Beijing Institute of Graphic Communication as a lecturer.

Minchao Xi was born 1987 in Shanxi China. He received his B.S. degree in computer science and technology from Shanxi University, Shanxi, China, in 2009, and received his Ph.D. degree in computer science and technology from Beijing University of Posts and Telecommunications, Beijing, China, in 2013.

Dr. Xi is working in Image and Video Watermarking Theory as a lecturer in Beijing University of Posts and Telecommunications.

Kaiguo Yuan was born in 1982 in Guizhou China. He received his Ph.D degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2009.

Dr. Yuan is working in Information Security as a lecturer in Beijing University of Posts and Telecommunications.