A ROBUST VIDEO WATERMARKING METHOD

*Dug-Ryung Kim, **Sung-Han Park

*Department of Computer Science, Ansan College, 1 Dong, Ansan, Gyung-Gi Do, 425-701, Korea
**School of Electronics and Computer Engineering, Hanyang University 4-1 Dong, Ansan, GyungGi Do, 425-791, Korea

ABSTRACT

We propose a new watermarking method to embed a label in a video which is robust against the change of the group of picture. The proposed method embeds labels in the pixel domain, but detects the label in the DCT frequency domain. A lookup table having the pixel patterns and the sequences of a sign of eight low DCT coefficients is used. To keep a quality of videos in embedding a label, the size of watermark is determined based on the human visual system. In this paper, we analyze bit error rates for labels of videos compressed by MPEG2 using the central limit theorem and compare the simulation results with previous methods.

1. INTRODUCTION

To protect the illegal use of multimedia data, many watermarking methods have been proposed. For video proposed each watermarking method which uses uncompressed video[1]-[3]. Hartung proposed a method to embed a watermark in MPEG2 bit stream and detect it in the pixel domain[4]. However, these methods require the preprocessing of Inverse DCT and motion compensation if a watermark is detected in compressed video with MPEG1, MPEG2, H.261,H.263.

On the other hand, Langelaar proposed a method to embed a label into the I frame in DCT domain[5]. Linnarz proposed a method to embed data into frame sequence with a predefined order for picture coding type[6]. These methods are weak to a change of the GOP structure of watermarked video on re-encoding MPEG2 to remove an embedded data.

To solve these problems, in this paper, a new method is proposed, where a label is embedded in pixel domain of each frame and detected from DCT domain of I frame in MPEG bit stream. For this purpose, a lookup table having the pixel patterns and the sign sequences of these corresponding low DCT coefficients is exploited. The proposed method is robust against the change of GOP structure. In addition, DCT, IDCT and motion compensation are not required in a label embedding and detecting.

The rest of the paper is organized as follows. In section2 and Section 3, label embedding and label detection algorithms are described, respectively. Section 4 provides simulation results and conclusions follows in section 5.

2. LABEL EMBEDDING

To embed a label, an image is divided into label blocks, whose number is equal to the number of label bits to be embedded, and each label block consists of several basic blocks. In this paper, the sign pattern of the low DCT frequencies of a basic block is employed to embed watermark. Figure 1(a) shows a 8x8 basic block, where “1” and “0” represent the pixel whose value is larger and smaller, respectively, than the average pixel value of the basic block. For the pixel pattern block of Figure 1 (a), the sign sequence of eight DCT low frequency coefficients of the basic block has a unique pattern of “+++++” as shown in Figure 1(b). In this way, we can have 256 different sign sequence pattern, and their corresponding pixel patterns. For convenience of representation, a look up table relating pixel patterns and sign patterns of DCT coefficients can be made as shown in Table1, where each row of figure 1 (a) is represented as 2-byte octal numbers, and the sign “+” and “−” of the coefficient sign are replaced with ‘1” & ‘0”, respectively.
Figure 2 shows the block diagram of the proposed label embedding procedures. Luminance planes are founded from each frame sequence of a video data. A basic block is defined in the luminance plane. The 256 sign patterns of DCT frequency coefficients are divided into two groups, one example is shown in Table 2. One group is used for embedding the bit “1” and the other group “0”. In this way, the possible number of look-up table will be $2^{128} \times 128!$. One of 127 Group ID is assigned to each basic block by using a PN-sequence. The pixel value represented as “1” and “0” in its corresponding pixel pattern is increased and decreased by a watermark value, respectively. The watermark values of each pixel in a luminance plane are determined using a human visual system based on the Peterson’s model[7].

Figure 3 shows the video label embedding algorithm, where $L_m$ is the $m$’th bit to be embedded into the $m$’th label block, $p_{ij}$ is the pixel value of position $(i,j)$ in a basic block, $\hat{p}_{ij}$ is the watermarked pixel value, and $t_{ij}$ is the watermark value.

![Table 1. Lookup table of pixel pattern and sign patterns of DCT coefficients.](image)

<table>
<thead>
<tr>
<th>Pixel patterns</th>
<th>DCT frequency coefficients sign sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1C, 1E, 1E, FF, FF, FF, 7F</td>
<td>00</td>
</tr>
<tr>
<td>07, 0E, 3C, 3C, 7C, FE, C7, 83</td>
<td>01</td>
</tr>
<tr>
<td>1C, 1C, 1F, 1F, 1F, FF, FC, F8</td>
<td>02</td>
</tr>
<tr>
<td>7F, 3F, 1F, 1E, 1C, 00, C0, E0</td>
<td>27</td>
</tr>
<tr>
<td>E3, E3, C7, 07, 07, 07, 07, 07</td>
<td>7F</td>
</tr>
<tr>
<td>1C, 1C, 38, F8, F8, F8, F8, F8</td>
<td>80</td>
</tr>
<tr>
<td>80, C0, E0, E1, E3, FF, 3F, 1F</td>
<td>D8</td>
</tr>
<tr>
<td>E3, E1, E1, 00, 00, 00, 80, E0</td>
<td>FF</td>
</tr>
</tbody>
</table>

![Table 2. An example of lookup table for label embedding](image)

<table>
<thead>
<tr>
<th>Group ID</th>
<th>DCT frequency coefficients sign sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group A</td>
</tr>
<tr>
<td>0</td>
<td>00</td>
</tr>
<tr>
<td>1</td>
<td>01</td>
</tr>
<tr>
<td>2</td>
<td>02</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>126</td>
<td>7E</td>
</tr>
<tr>
<td>127</td>
<td>7F</td>
</tr>
</tbody>
</table>

![Figure 1. A pixel pattern and a corresponding DCT frequency coefficients](image)

![Figure 2. The Diagram of watermarking algorithm](image)
3. LABEL DETECTION

To get a DCT coefficients from MPEG2 bit stream without IDCT and motion compensation procedure, only I frames are used. After variable length decoding and de-quantization of the watermarked DCT coefficients for I frames, a sign pattern of each DCT coefficient is assigned to each basic block of the watermarked DCT coefficients with the used PN-sequence.

The watermarked DCT coefficients corresponding to “1” of the sign sequence of DCT coefficients of a group are defined as a set X, and the others as a set Y. To make the watermarked DCT coefficients smaller than a threshold T, every watermarked DCT frequency coefficients are mapped using equation (1) to increase the label detection rates.

\[
\begin{cases}
    w_{ij} = c_{ij} & \text{if } \langle c_{ij} \rangle > T \\
    w_{ij} = \text{sign} \left( c_{ij} \right) T & \text{if } \langle c_{ij} \rangle \leq T
\end{cases}
\]  

(1)

where \( c_{ij} \) is the watermarked DCT coefficients of frequency (i,j) and \( w_{ij} \) is the mapped DCT coefficients used in a label detection. If a bit “1” is embedded in a label block, the mean of the mapped DCT coefficients in the set X is greater than that in the set Y. If a bit “0” is embedded, vise versa. Hence, one bit embedded in a label block can be detected by the equation (2).

\[
\begin{align*}
    l_m &= 0 & \text{if } & \bar{X}_m < \bar{Y}_m \\
    l_m &= 1 & \text{if } & \bar{X}_m \geq \bar{Y}_m
\end{align*}
\]  

(2)

where \( l_m \) is the detected value of the m’th label block of a frame, \( \bar{X}_m \) and \( \bar{Y}_m \) is the mean of mapped DCT coefficients in the set X and Y.

For the performance analysis of the proposed system, the label detection bit error rate is derived based on the central limit theorem as shown in equation(3)[8].

\[
BER = \frac{1}{2} \text{erfc} \left( \frac{-K}{\sqrt{2(N\sigma^2 + \sigma_k^2)}} \right)
\]

(3)

where \( K \) and \( \sigma_k^2 \) is the mean and variance of an added watermark size in DCT domain, \( N \) is the number of the used DCT coefficients in watermarking, and \( \sigma^2 \) is variant of the mapped DCT frequency coefficients. In the equation (3), the larger \( K \) and \( N \) is and the smaller \( \sigma_k^2 \) and \( \sigma^2 \) is, the smaller the label detection bit error rate is.

4. SIMULATION RESULTS AND DISCUSSIONS

Three video sequences, Football, Pingpong and Miss America, are used. The size of each frame is 352 x 240. The PSNR of watermarked Football and Pingpong video frames are larger than 35 dB and Miss America video frame is larger than 43 dB. The watermarked video appears visually identical to original.

Table 3 shows the label bit error rates of the watermarked videos, where the symbol '*' indicates no error. The label bit error rate of Miss America video is less than one of Football or Pingpong video, because a variant of DCT coefficients of Miss America video is less than one of Football or Pingpong video.

To compare the robustness of the proposed method and the previous methods against MPEG2 compression, watermarked videos embedded with 60 bits per one frame are MPEG2 encoded with 25 frames per second and bit rates such as 4Mbps, 2Mbps, 1Mbps, 0.8Mbps. Figure 4
and 5 show the label bit error rates of Darmstaedter’s method, Hartung’s method, and the proposed method. Label bit error rates of previous methods are better than the proposed method for the uncompressed video. However, for MPEG bit stream, the label bit error rate of the proposed method is better than that of the previous methods.

In this experiment, to be robust against change of GOP, the same label is repeatedly embedded in every frame. If a preprocessing finds shots, a different label can be embedded in each shot. So a watermark propagation does not happen. Under the condition of keeping a good quality of video, a frame order change must be happened in a shot. Hence this method may also be robust against the modification of frame sequence.

Table 3. Label bit error rates

<table>
<thead>
<tr>
<th>N</th>
<th>Football</th>
<th>Pingpong</th>
<th>Miss America</th>
</tr>
</thead>
<tbody>
<tr>
<td>176</td>
<td>5.59×10^-4</td>
<td>4.85×10^-3</td>
<td>2.67×10^-3</td>
</tr>
<tr>
<td>352</td>
<td>6.15×10^-4</td>
<td>1.25×10^-4</td>
<td>*</td>
</tr>
<tr>
<td>704</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

5. CONCLUSIONS

In the proposed method, a label is embedded in pixel domain of every frame using a lookup table having an information between the pixel block pattern and the sign sequences of 8 low DCT coefficients. The embedded label is detected in DCT domain of MPEG bit stream without the IDCT and the motion compensation process. The simulation results show that the proposed method is better than previous methods in terms of label bit error rate of MPEG2 encoded video.

6. REFERENCES