New Video Watermarking Scheme by Using Dither Modulation to Local Energy

Xiao-Yan Liu, Li-Hong Ma, Hu-yun Yin
School of Electronic Engineering, South China University Technology
xyli1982@163.com

Abstract

This paper proposes the differential energy modulation (DEM) algorithm for video watermarking. In DEM scheme, the watermark bits are embedded by dithering differential energy in Quantization Index Modulation (QIM). The energy difference between horizontal edge components and the vertical ones in a quantized block is modulated by modifying some edge components to make the differential energy consistent with a dithered quantizer. Besides, the watermarked positions are carefully selected to preserve video quality. Experimental results demonstrate that our DEM scheme performs better on visual quality impact, capacity, robustness and rate maintenance than the DEW (Differential Energy watermarking) method.

1. Introduction

Watermarking technique has been considered as an efficient approach in data protection, such as images, video and audio signals. The most significant video watermarking scheme is embedding the watermark signal directly in the compressed stream for this way demands less computational effort. The DEW scheme is a typical video watermarking method [1,2]. It selectively discarded AC components to embed the watermark bits. The scheme has low complexity in embedding but has several limitations. Firstly, it discarded some coefficients without considering the texture and edge effects, thus this action will result in blurring of edge regions. Secondly, a high percentage of coefficients been changed will cause image degradation, error propagation and difficulty in bit rate control. The third, its embedding and extraction energy threshold were chosen empirically and kept invariant in all cases of attack, thus the extraction are erroneous. Another typical watermarking method was emphasized on bit rate control[3]. In [3], a watermark bit was embedded in a nonzero DCT coefficient only when the codeword was longer than its substituted watermarked one. But it caused problems of transmitting embedded positions in blind detection. Drift compensation was used to balance the bit amount of the watermarked signal, but the computational complexity was increased.

To improve the drawbacks mentioned above, we propose a blind method for compressed video watermarking. Dither modulation is used to quantize the energy difference between two sub-regions of a quantized block to embed a watermark bit. Simulation results demonstrate that our DEM scheme is superior to the DEW scheme in transparency, capacity, robustness and rate maintenance.

2. Differential energy watermarking and quantization index modulation

2.1 Differential energy watermarking (DEW)

Fig.1 illustrated the implementation of the DEW method[1], and the block number of a label carrying region (lc-region) was 16. We could briefly describe the DEW scheme by the following procedure.

Step 1 Differential energy calculation. (1) Computed energy $E_A$ of the top half lc-region $A$ and $E_B$ of the bottom one $B$ with the high frequency DCT coefficients (shown as the white triangular areas in Fig.1, denoted as subset $S(c)$). The subset $S(c)$ was defined as:

$$S(c) = \{u \in \{0, ..., 63\} \mid (u > c)\}$$

(1)

where $u$ represented coefficient locations after a cutoff point $c$ in the zig-zag scanning, and $c$ should be bigger than the minimal cutoff index $C_{\text{min}}$ decided by optimization. Energy $E_A$ was computed by:

$$E_A = \sum_{b=0}^{n-1} \sum_{u \in S(c)} (DCTcoeff(u,b)^2)$$

(2)

Where $b$ denoted the number of blocks of sub-region $A$. And energy $E_B$ of $B$ could be defined similarly.

(2) Obtained differential energy between $A$ and $B$
The larger the $D$, the more robust the watermark, but the less DCT coefficients preserved, and the worse visual quality.

**Step 2** Label bit embedding and decoding.

Employed the sign of $D$ to represent a label bit. If $D>0$, ‘0’ was represented. If $D<0$, ‘1’ was defined. Then, adapted $E_A$ or $E_B$ to manipulate the energy difference $D$ to embed a label bit:

If the label bit was ‘0’, all DCT components after $c$ in region $B$ were eliminated to make $E_B$ to zero.

$$D = E_A - E_B = +E_A$$  \hspace{1cm} (4)

If the label bit was “1”, set $E_A$ to zero,

$$D = E_A - E_B = -E_B$$  \hspace{1cm} (5)

A similar process was performed to extract a label bit. Calculated $E_A$ and $E_B$ for each $c \in \{0, \ldots, 63\}$ to find an cutoff point $c'$, then the label bit was detected by comparing the detection energy difference with 0:

$$T = E_A(c') - E_B(c')$$  \hspace{1cm} (6)

If $T>0$, ‘0’ was decoded. If $T<0$, ‘1’ was decoded.

### 2.2 Quantization index modulation (QIM)

QIM embedded a watermark by modulating indices [4,5]. Its embedding function was:

$$S(X, W) = Q_W(X)$$  \hspace{1cm} (7)

Where $X$ was a host signal, $W$ was a watermarking signal, $Q_W()$ was a quantizer, and $S$ was a composite watermarked signal. Dither modulation is a typical realization of QIM. Its embedding function was:

$$S(X, W) = Q_W(X + D(W)) - D(W)$$  \hspace{1cm} (8)

Where $D(W)$ was a dither vector of $W$.

Fig. 2 illustrated a composite dither modulation quantizer with uniform distribution. Two quantizers $Q(, 0)$ and $Q(, 1)$ were used to embed a watermark signal $w \in \{0,1\}$. Positions marked ⊙ and ⊙ belonged to $Q(, 0)$ and $Q(, 1)$ respectively. If $w=0$, $X$ was quantized to its nearest ⊙ using $Q(, 0)$, and quantized to its nearest ⊙ by $Q(, 1)$ while $w=1$. The final signal $S$ consisted of the output sets of two quantizers.

At receiving end, a watermark bit could be estimated by a noisy watermarked $S'$.

$$w = \arg\min\{S', \wedge w\}, \quad w \in \{0,1\}$$  \hspace{1cm} (9)

If $S'$ was an output of $Q(, 0)$, ‘0’ was decoded. Otherwise, ‘1’ was decoded.

### 3. Selected embedding positions and energy dither signal

In this section, we will suggest the differential energy act as a dither host signal to embed a label bit.

#### 3.1 Embedded positions and differential energy definition

A certain change of DC value usually brings a variation of luminance and results in block effects. And the artifacts may propagate to other blocks and subsequent frames. So we abandon the DC option and select 10 AC coefficients labeled as $VF_j$ and $HF_j$ (as shown in Fig. 3) ($j=1,2,3,4,5$) respectively to embed a watermark bit.

In particular, we define a simple vertical factor ($VF$) and a horizontal one ($HF$) to substitute energies of $VF_j$ and $HF_j$ regions for calculation convenience.

**Fig. 1** 1 bit embedding in DEW [1].

**Fig. 2** Quantizers for dither modulation embedding.

**Fig. 3** $VF_j$ and $HF_j$ positions in a quantized block.
\[ VF = \sum_{j=1}^{5} |VF_j| \quad (10) \quad HF = \sum_{j=1}^{5} |HF_j| \quad (11) \]

Actually, \( VF \) and \( HF \) represent the distributions of low frequency energy in vertical and horizontal directions, and their difference \( DF \) could approximate the differential energy between vertical texture and horizontal one of the block to some extent.

\[ DF = VF - HF \quad (12) \]

### 3.2 Dither modulation to differential energy

By applying modulation procedure, we dither \( DF \) to \( DF' \) to indicate a label bit.

(1) If the embedded bit \( w=1 \), \( DF' \) is acquired by

\[
DF' = \left\lfloor \frac{DF}{\Delta} \right\rfloor \times \Delta \text{ if } \left\lfloor \frac{DF}{\Delta} \right\rfloor \text{ is even} \quad (13)
\]

(2) If the embedded bit \( w=0 \), \( DF' \) is acquired by

\[
DF' = \left\lfloor \frac{DF}{\Delta} \right\rfloor + 1 \times \Delta \text{ if } \left\lfloor \frac{DF}{\Delta} \right\rfloor \text{ is odd} \quad (14)
\]

where \( \lfloor \cdot \rfloor \) is a round operator, and \( \Delta \) denotes the quantization step size. And \( DF \) and \( DF' \) satisfy:

\[
|DF' - DF| \leq \Delta \quad (15)
\]

Then we modify some components of \( VF_j \) or \( HF_j \) to make the energy difference satisfy with \( DF' \).

(1) Calculate the difference \( diff \) between \( DF \) and \( DF' \).

\[
diff = DF' - DF \quad (16)
\]

(2) Split \( diff \) into two values \( val1 \) and \( val2 \).

\[
val1 = \text{int}(\frac{diff}{2}) + 1 \quad (17)
\]

\[
val2 = \text{diff} \mod val1 \quad (18)
\]

(3) Let \( HF_{fmax} \) and \( HF_{smax} \) be the biggest value and the second biggest one of \( |HF_j| \). \( VF_{fmax} \) and \( VF_{smax} \) are similar defined on \( |VF_j| \), they are modified referring to the sign of \( diff \) :

- If \( diff > 0 \), decrease \( HF \) when \( |HF_{smax}| \geq val2 \) or \( |HF_{fmax}| \geq diff \), otherwise increase \( VF \). \( DF \) will finally increase to its nearest quantized value \( DF' \).
- If \( diff < 0 \), decrease \( VF \) when \( |VF_{smax}| \geq val2 \) or \( |VF_{fmax}| \geq -diff \), otherwise increase \( HF \). \( DF \) will decrease to its nearest quantized value \( DF' \).
- If \( diff = 0 \), no values needs to be changed.

Fig.4 gives the dithering sketch when \( diff > 0 \). We could find that our scheme varies energy difference in a small range and spreads their error over some coefficients. And the video could be well protected.

### 3.3 Watermarking scheme

#### Label bit embedding.

When a watermark sequence \( w_j \) \( (j=0,1,\ldots,L-1) \) is embedded in I-frames by DEM, each \( w_j \) is hidden in a quantized DCT block. Fig.5 gives the dithering sketch of embedding a watermark bit.

#### Watermark extraction.

The embedded watermark signal could be decoded without original video content. Compute the \( DF \) of a watermarked block by equ.(12), and obtain the rounded value \( \left\lfloor \frac{DF}{\Delta} \right\rfloor \). If it is even, ‘1’ was decoded, otherwise ‘0’ was decoded.

### 4. Experiments results

The test sequence "forman" contains 100 frames, and each frame contains 440 luminance blocks. In the DEM scheme, the maximum capacity will totally be 440 watermark bits (1bit/block). Its step size \( \Delta \) takes values of 3, 4 and 5 respectively. For DEW, a watermark bit is represented by 16 \( 8 \times 8 \) blocks, i.e.
each I frame can embed 27 bits at most. The enforced energy difference in embedding is $D=500$, the detection threshold is $T=50$, and the minimum cutoff point is $C_{\min}=6$. To compare DEM with DEW under the same condition, we both embed 25 bits in each I frame.

**Transparence.** The visual quality is evaluated by the average degradation of PSNR in I pictures. Fig.6 shows the PSNR curves of DEM and DEW when the sequence is encoded at 0.4Mbps. In this example, the average PSNR degradation of DEM is 0.5111dB ($\Delta=3$), 0.7667dB ($\Delta=4$) and 1.1778dB ($\Delta=5$), respectively. While for DEW, it is 2.6333dB. Thus the DEM scheme is superior to DEW in visual quality when the same label bits are embedded.

**Capacity.** The capacity of a block-based watermarking method is determined by the number of blocks used to embed one label bit. For DEW, one bit is usually embedded in 16 $8 \times 8$ blocks. While for DEM, one bit is embedded in one block, thus DEM scheme has 16 times capacity than DEW.

**Robustness.** To test the watermark robustness, the watermarked sequence is encoded at 10 Mbps firstly. Hereafter, the sequence is transcoded at different lower bit-rates. With the bit-rate is decreased from 10 Mbits/s to 2 Mbit/s (as shown in Fig.7), label bit errors introduced by DEW are increased from 1.78% to 13.78%, while introduced by DEM algorithm are increased from 0 to 0.89% ($\Delta=3$), to 1.78% ($\Delta=4$) and to 1.33% ($\Delta=5$), respectively. Thus our DEM scheme performs better on robustness than DEW scheme.

**Rate Control.** The ability of rate maintenance is determined by the codeword substation strategy. The amount of bits variation between the watermarked codeword and its original one is shown in Fig.8. It appears that the average variation of bits in I pictures of DEW is 1127.6 bits much greater than 511.22 bits ($\Delta=3$), 289 bits ($\Delta=4$) and 305.44 bits ($\Delta=5$) of DEM scheme. It is obvious that DEM scheme is better than DEW in satisfying the requirement of rate maintenance.

5. Conclusion

In this paper, we propose a new video watermarking scheme based on energy dithering in the compressed domain. Since watermark bits are just embedded in quantized DCT blocks, the DEM scheme only requires partial decomposition of the original video sequence. The embedding watermark only slightly affects the successive frames referred to the host I frame. Simulation results demonstrate that our scheme offers a better embedding in transparence, capacity, robustness and rate maintenance than DEW scheme.

6. References