An Efficient Video Watermark Method with Adaptive Use of Block QP

DongHwan Shin and JungUk Choi

MarkAny Inc. 10F, Ssanglim Bldg, 286, Toeqye-ro, Chung-gu, Seoul, Korea
dhshin@markany.com, juchoi@markany.com

Abstract: Robustness and invisibility are all important points in video watermarking technology. In this paper, we propose the new scheme to increase the robustness and keep the invisibility as good. Through the experiment, we get the relationship linear curve between QP values and watermark strength values, and using that relationship, we get the better detection rate increased by 7%~14%. The performance increases are 6.94%, 8.03%, and 14.32% about MPEG2, MPEG4 and H.264 attacks respectively. For the subjective quality assessment, we experiment with the 23 general non-specialists. In spite of increasing the performance, we have still good invisibility which is checked by subjective test. It is difficult to obtain the distinction between the source and watermarked content statistically.

Keywords: Watermark; QP; video; robustness; invisibility

1. Introduction

Thanks to the development of present internet technology and spread of storage, it became the era when anyone can get any digital data, especially video, image and audio which they want from internet.

In contrast, because of the development of displaying device and broadcasting technology, contents maker must spend more efforts and cost compared to earlier[1,2,14].

Especially according to increase the demand of UHD video, the production costs can be much increased but leakage environments are becoming better and better. It became the times when anyone can get the good quality video only with a general video capture device[3,15,16,17,20].

As a complement to this, the contents makers demand the watermarking technology as one of copyright protection methods. The Movie Labs, which provides the technical foundation of the Hollywood studios in the United States, has issued the guideline that requires protective measures such as watermark[7]. Also as an effort to protect the unauthorized retransmission of the broadcast contents, especially premier league football, watermark technology is used[9,13].

In the watermark technology, there is the trade-off relationship between robustness and invisibility. If watermark is embedded too much strongly, the contents are apt to damage easily proportional to that[4].

In this paper, we suggest the ways to improve the detection rate but still feel the contents are good.

The remainder of this paper describes the principles that enable the adaptive strength watermarking in chapter 2. And in chapter3, linear regression curve is obtained from the experiments of actual various genre videos. Section 4 presents some extent to verify that regression curve is effective when it is applied to the actual broadcast contents.

Adaptive Watermark

Two major elements of the video watermarking, image quality and robustness, have been continuously researched as a trade-off relationship

In this study, we propose the way to set the watermark embedding strength based on QP(Quantization Parameter). If the watermark embedding strength is adaptively adjusted, there is the effect which increases the robustness and minimizes the visual artifact.

Received: December 7th, 2015. Accepted: September 20th, 2016
DOI: 10.15676/ijeei.2016.8.3.10

612
2. Quatization Parameter

Block-based hybrid video encoding schemes such as the MPEG 2, 4 and h.26* families are inherently lossy processes. They achieve compression not only by removing truly redundant information from the bit stream, but also by making small quality compromises in ways that are intended to be minimally perceptible. In particular, the quantization parameter QP regulates how much spatial detail is saved. When QP is very small, almost all that detail is retained. As QP is increased, some of that detail is aggregated so that the bit rate drops – but at the price of some increase in distortion and some loss of quality.

The Quantization Parameter determines the step size for associating the transformed coefficients with a finite set of steps. Large values of QP represent big steps that crudely approximate the spatial transform, so that most of the signal can be captured by only a few coefficients. Small values of QP more accurately approximate the block's spatial frequency spectrum, but at the cost of more bits.

In H.264, each unit increase of QP lengthens the step size by 12% and reduces the bitrate by roughly 12%. Quantization in an H.264 encoder is controlled by a quantization parameter, QP, that ranges from 0 to 51. QP is an index used to derive a scaling matrix. It is possible to calculate the equivalent quantizer step size (Qstep) for each value of QP. As QP increases, Qstep increases; in fact, Qstep doubles for every increase of 6 in QP[5].

And MPEG2 also uses a similar concept in the quantiser_scale_code. This is an unsigned integer in the range 1 to 31. Decoder shall use this value for de-quantization[6].

Figure 1 suggests that relationship for a particular input picture – if you want to lower bit rate, you can do so by lowering QP at a cost of increased distortion, which is guessed by (b)QP-bitrate relationship. (a)QP used in bitrate control in Fig.1 displays the closed loop bitrate control(CBR, Constant Bitrate Control). They use QP in order to control the bitrate in video encoding.

![Figure 1. QP-Bitrate relationship](image)

When inserting a watermark in the video, it has a little inefficient side to apply the same global constant strength to the one frame. Though this method uses the HVS to adjust the watermark strength based on the characteristics of input image, it only uses the characteristics of input image but doesn’t use the characteristics of compression of original video.

In the general video watermarking algorithm, there is the problem in which the detection rate is high in the less movement video but is low in a lot of movement video such as sports game. This is due to the compression properties. As the duplication parts between previous frames decrease in case of high-motion video, the encoder has become the status that the amount of information to compress is increasing. Therefore, if the video is compressed in the same bitrate, the encoder will increase the QP value naturally to undermine much of the image information. Because of that reason, the watermark information will be removed together. If watermark is embedded strongly considering the QP value, watermark detection rate will be improved without perceiving the image degradation.
**Embedding Watermark**

Usually, the input of video watermarking system is often the compressed video file or compressed video stream. By default, the input signals in this study are intended for the compressed bit stream such as MPEG.

Firstly, Demux part analyzes input bit stream and separates the video stream to supply to decoder. Decoder will decode the inputted the compressed video stream to give the raw video data to the Watermark embedding module with QP together.

The watermarked video raw data are sent to Encoder for compression. The compressed video stream is sent to Mux to produce the watermarked full video stream.

![Figure 2. Watermark embedding system with QP](image)

Figure 2. Watermark embedding system with QP

Figure 3 shows Watermark embedding module which has a role to embed watermark on the video raw data. This embedding algorithm uses the traditional spread spectrum method[10] and M-ary modulation method[11,12].

![Figure 3. Watermark Embedding Module](image)

Figure 3. Watermark Embedding Module

The process of embedding watermark is expressed like formula 1. If original image is inputted, HVS(Human Visual System) values are calculated considering the characteristics of human visual system, especially large values in edge parts. Previous research normally uses the constant value, $\alpha$ as a global watermark strength factor. In this study, Adaptive strength calculation part has a role to receive the QP values as macro blocks unit, and produce the $\alpha(x, y)$, adaptive watermark strength.

\[
I' = I + \alpha(x,y) \cdot HVS \cdot WM \, Sequence
\]

(1)

where $x = 0, 1, \ldots, \text{int}(M/16)$,

$y = 0, 1, \ldots, \text{int}(N/16)$,
Extracting Watermark

Figure 4 shows the watermark extracting system.

Demux receives the watermarked bit stream as a input, separates the video stream and sends the only video stream to decoder.

Watermark extracting module receives the decoded raw video data from the decoder, extracts the watermark information and outputs it.

![Figure 4. Watermark Extracting System](image)

Figure 5 shows the block diagram of Watermark extracting module. The watermarking algorithm of this study embeds and extracts the watermark information based on one frame data. That is, if there is only the image data of one frame, we can extract the complete watermark information.

![Figure 5. Watermark Extracting Module](image)

Frame scheduler receives the video raw data and the additional information such as the video format, frame resolution and frame number, and allocates the resource to the watermark extracting process.

The watermark extracting process performs various operations for separating only the watermark signal from the input image. Firstly, it can get the exact watermark starting point from the RST recovery logic. And it can get the HVS values from HVS module to estimate the watermark strength. Based on these information, the watermark extracting module tries to extract the watermark. If the watermark extraction is succeeded, Message decoder part can
translate the encoded watermark information into the decoded watermark information which we can know easily.

Watermark system may use only one frame information, but usually a combination of watermark information of multiple frames, which constitute the complete watermark information by Merge watermark part.

3. Adaptive Strength Setting

When the QP value of input video is given, we want to set the optimal watermarking strength for it.

Here, the optimal term refers to setting the best values considering the trade-off relationship between detection rate and visual quality. In order to do that, we must know the relationship among the QP value, watermark strength, and detection rate from the multiple genres of video contents.

Relationship of QP, Watermark Strength and Detection Rate

In order to know the relationship among QP, watermark strength and detection rate. We tested the 5 genres video contents(sports, drama, movie, animation and nature).

Figure 6 shows that the vertical axis in the Figures is watermark strength, which is 65 to 195. The horizontal axis is QP values. The values on the cell are the detection rate.

When the color is close to green it has good detection rate while red is bad detection rate. -1 means “Data don’t exists”
In case of high-motion video, (a) sports, the detection rate of 100% happens since the watermark strength is more than 100, which means that watermark must be embedded strongly in order to be detected.

Conversely, in case of relatively small movement image, (b) drama, we can get 100% detection rate by watermark strength of 65 in QP is 24 case.
**Linear Regression Curve**

If collecting the points in which detection rate are over 80% among the values from Fig. 6, and calculating the linear regression curve, we can get the linear line like Fig. 7. Here, after then, letting the minimum value and maximum value to be set, we need to handle in order to insert the values into the range.

![Figure 7. Linear relationship curve between QP and Strength](image)

The formula of the interval between QP is 24~32 can be written like below:

**Min2 case:**

\[ y = 5 \cdot x - 45 \]  \hspace{1cm} (2)

where y is watermark strength. [0, 200]

x is QP value, [0, 51]

You can input the base strength via API, so that you can adjust the relatively stronger or weaker strength.

Then, the API will return the average strength of the frame as well as maximum strength.

**Effect of Adaptive Watermark Strength**

Figure 8 shows the effect of adaptive watermark strength as a graph.

![Figure 8. Adaptive strength usage example](image)
For instance, the pictures on the horizontal axis display the no watermarked image (strength=0).

In case of QP 24, there is no visual difference between the constant strength embedding in one frame fully with strength 75 and the adaptive strength embedding. In addition, in case of QP 32, it shows that it doesn’t have no visual difference between inserting in fixed strength 75 and inserting in adaptive strength 95.

Adaptive strength principle is that while with a lower QP, weaker embedding strength has similar visual quality, with a higher QP, the stronger embedding can have similar visual quality.

Figure 9. Enlarged testing image

Figure 9 is enlarged pictures from Figure 8 images. The center row’s image are original images, and top row’s images are the watermarked images with fixed strength 75, and bottom row’s images are the watermarked images with adaptive strength 75, 95 respectively. The images at top and bottom show the visual quality which is difficult to visually distinguish.

Based on the principle of comparing and analysis, we subjectively picked up the strength and QP value with comparing the visual quality. Based on it, we made the strength formula. We set the base strength as minimum strength we should use. Also we set the maximum strength we should not over.

4. Experiments

In order to know the validity of adaptive strength method from chapter 3, we tested it with the real broadcast video contents.

Robustness Test

In order to check the validity of proposed algorithm, we compared the results from global constant strength and adaptive strength.

The test contents are selected randomly 6 programs from 1 minute length of Euro Sport. The video resolution of the original input image is 720x576 pixels.
In Table I, Test code TC_00 is without attack case, which means detecting immediately after embedding. TC_07_13 means compressing the video as MPEG2, 1.5Mbps after embedding watermark, TC_07_14 is the test case of MPEG4 1Mbps compression, TC_07_15 is the test case of H.264 1Mbps compression. Because the results come out from the original B frames, the detection rate is generally low in all cases.

Table 1. Compare the detection rate between the fixed and adaptive strength

<table>
<thead>
<tr>
<th>PID</th>
<th>Test Code</th>
<th>Bypass Attack</th>
<th>MPEG2 720x576 (1.5Mbps)</th>
<th>MPEG4 720x480 (1Mbps)</th>
<th>H.264 720x576 (1Mbps)</th>
<th>Previous Extraction rate</th>
<th>New Extraction rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>TC_00</td>
<td>68</td>
<td>100.00%</td>
<td>100.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_13</td>
<td>68</td>
<td>71.01%</td>
<td>77.94%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_14</td>
<td>68</td>
<td>8.70%</td>
<td>11.76%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_15</td>
<td>68</td>
<td>21.74%</td>
<td>35.29%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>201</td>
<td>TC_00</td>
<td>58</td>
<td>100.00%</td>
<td>100.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_13</td>
<td>58</td>
<td>92.86%</td>
<td>98.21%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_14</td>
<td>58</td>
<td>26.79%</td>
<td>28.57%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_15</td>
<td>58</td>
<td>41.07%</td>
<td>44.64%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>TC_00</td>
<td>73</td>
<td>97.26%</td>
<td>100.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_13</td>
<td>73</td>
<td>90.41%</td>
<td>97.26%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_14</td>
<td>73</td>
<td>26.03%</td>
<td>28.77%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_15</td>
<td>73</td>
<td>34.25%</td>
<td>50.68%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2432</td>
<td>TC_00</td>
<td>89</td>
<td>95.51%</td>
<td>100.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_13</td>
<td>89</td>
<td>87.64%</td>
<td>100.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_14</td>
<td>89</td>
<td>65.17%</td>
<td>80.68%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_15</td>
<td>89</td>
<td>47.19%</td>
<td>64.77%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>611</td>
<td>TC_00</td>
<td>69</td>
<td>100.00%</td>
<td>100.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_13</td>
<td>69</td>
<td>82.61%</td>
<td>85.51%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_14</td>
<td>69</td>
<td>23.19%</td>
<td>21.74%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_15</td>
<td>69</td>
<td>31.88%</td>
<td>47.83%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>601</td>
<td>TC_00</td>
<td>69</td>
<td>100.00%</td>
<td>100.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_13</td>
<td>69</td>
<td>89.96%</td>
<td>97.10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_14</td>
<td>69</td>
<td>7.25%</td>
<td>18.84%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC_07_15</td>
<td>69</td>
<td>40.58%</td>
<td>59.42%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 10. Compare the fixed and adaptive strength
The results of MPEG4 attacks are the worst case. Because the compression efficiency is lower than that of H.264 with same bit rate, that phenomenon happens.

Figure 10 shows the results of figure 10 graphically. The performance increases are 6.94%, 8.03%, and 14.32% about MPEG2, MPEG4 and H.264 attacks respectively.

**Invisibility Test**

In order to test the invisibility, we did two test items. One is PSNR and the other is subjective test.

1. **PSNR (Peak Signal to Noise Ratio)**

   PSNR is the usually used method to evaluate the visual quality.

   $$
   \text{PSNR} = \frac{255^2}{\text{MSE}}
   $$

   $$
   \text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (x(i, j) - y(i, j))^2
   $$

<table>
<thead>
<tr>
<th>Video</th>
<th>Fixed strength With 65</th>
<th>Original</th>
<th>Adaptive strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Quality</td>
<td>Good to see (Not much difference with original one)</td>
<td>Good to see</td>
<td>Good to see (Not much difference with original one)</td>
</tr>
<tr>
<td>Embedding Strength</td>
<td>65 (fixed)</td>
<td>N/A</td>
<td>77 (Average)</td>
</tr>
<tr>
<td>PSNR</td>
<td>41.72</td>
<td>N/A</td>
<td>39.94</td>
</tr>
</tbody>
</table>

   PSNR value of adaptive strength is naturally lower than that of the fixed strength because watermark total strength is increased.

2. **Subjective Test**

   The goal of this subjective test is to investigate whether the perceptual video quality of video sequences with watermark is worse than that of the reference video sequences.

   a. **Test environment**

   The test environment (room luminance, viewing condition, viewing distance, etc.) is set in accordance with ITU-R Recommendation BT.500[8,18,19].

   Figure 11 shows the test room in which subjective test has done.
b. Test Results
MOS (Mean Opinion Score) values are computed by averaging the scores of the 24 viewers. The correlation between the scores of a viewer and the MOS values are from 0.843 to 0.911 (average: 0.876). Figure 12 shows the MOS values of the reference and the test videos.

![Figure 12. MOS comparison with reference and watermarked videos](image)

Table 3. shows the meaning of MOS values in ITU-R.

Table 3. Meaning of MOS values in ITU-R

<table>
<thead>
<tr>
<th>Quality</th>
<th>Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Excellent</td>
</tr>
<tr>
<td>4</td>
<td>Perceptible, but not annoying</td>
</tr>
<tr>
<td>3</td>
<td>Slightly annoying</td>
</tr>
<tr>
<td>2</td>
<td>Annoying</td>
</tr>
<tr>
<td>1</td>
<td>Very annoying</td>
</tr>
</tbody>
</table>

c. Statistical Analysis
We performed the following statistical analysis (p-value).
- Mean of reference video sequences: 4.072
- Mean of test video sequences with watermarks: 4.013 (sd=0.372)
- Number of samples: 30
- test statistic: -0.880
- p-value: 0.189 (=p(z<-0.880)) > 0.05

Thus, with 95% statistical significance, one can’t reject the null hypothesis (μ = 4.072; the perceptual quality of the video sequences with watermarks is identical with that of the reference video sequences). In other words, the watermark method doesn’t produce statistical significant perceptual degradations.

5. Conclusion
Robustness and invisibility are all important points in video watermarking technology. In this paper, we propose the new scheme to increase the robustness and retain the invisibility as
Through the experiment, we get the relationship linear curve between QP values and watermark strength values, and using that relationship, we get the better detection rate increased by 7%~14%. The performance increases are 6.94%, 8.03%, and 14.32% about MPEG2, MPEG4 and H.264 attacks respectively.

In spite of increasing the performance, we have still good invisibility which is checked by subjective test. For the subjective quality assessment, we experiment with the 23 general non-specialists. With 95% statistical significance, one can’t reject the null hypothesis ($\mu = 4.072$; the perceptual quality of the video sequences with watermarks is identical with that of the reference video sequences). In other words, the watermark method doesn’t produce statistical significant perceptual degradations. It is difficult to obtain the distinction between the source and watermarked content statistically.

In the future, we need to make the QP-strength relationship curve more accurately through the analysis of more contents.

6. Acknowledgment
This research was supported by a grant from the Advanced Technology Center R&D Program funded by the Ministry of Trade, Industry& Energy of Korea (10042252).

7. References
[5]. ITU-T Recommendation H.264, also ISO/IEC 14496-10, 2005
[6]. IOS/IEC 13818-2, pp.75, 1994
[7]. MovieLabs Specification for Enhanced Content Protection, ver.1.0., 2013

DongHwan Shin received the MS degree and Ph.D. in electronics engineering from University of Seoul, Korea. From 1992 to 1994, he was a member of the LG Electronics Inc. He worked as a senior researcher in the Korea Sports Science Institute from 1996 to 2000. He has been currently a chief manager of MarkAny Inc. from 2000. Her research interests are in the areas of copyright protection, fingerprinting, watermarking and machine learning.

JongUk Choi received the Ph.D. in artificial intelligence from University of South Carolina, USA in 1989. From 1989 to 1992, he was the director of Artificial Intelligence Division. He has been currently a professor of Sangmyung Univ. from 1992. His main interests are in the areas of digital watermarking, DRM and artificial intelligence.