



Article MPEG and DA-AD Resilient DCT-Based Video Watermarking Using Adaptive Frame Selection

Jong-Uk Hou 回



Abstract: We present a robust video watermarking scheme and report the detailed robustness of the video watermarking assessed based on standard criteria obtained from Information Hiding and its Criteria (IHC) Committee. Using discrete cosine transform domain spread-spectrum watermarking, our system achieves robustness under various non-hostile video processing techniques, including MPEG compression and digital/analog–analog/digital (DA-AD) conversion. The proposed system ensures that a 16-bit embedded sequence can be extracted through adaptive frame selection in any 15-s interval, even with a long video clip. To evaluate the performance of the proposed watermarking scheme, we conducted robustness tests under a DA-AD conversion environment, based on the MPEG-4 part 10 (H.264) codec. The experiment results obtained indicate that, in addition to being robust against non-hostile video processes, the proposed method achieves invisibility. The assessment of the developed watermarking scheme also satisfies the third edition of the IHC video watermarking evaluation criteria.

Keywords: video watermark; signal processing; digital watermarking; video compression



Citation: Hou, J.-U. MPEG and DA-AD Resilient DCT-Based Video Watermarking Using Adaptive Frame Selection. *Electronics* **2021**, *10*, 2467. https://doi.org/10.3390/ electronics10202467

Academic Editor: Guido Masera

Received: 13 August 2021 Accepted: 28 September 2021 Published: 11 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

Digital technology has completely changed our lifestyles by providing several convenient digital tools; however, it has also created a plethora of problems, one of which is digital copyright infringement. Today, with the wide variety of available digital tools, digital content can be copied extremely easily, thereby facilitating the rapid spread of illegal digital content globally. Consequently, digital copyright protection has become an important issue. Digital watermarking is a process that changes a cover work invisibly to insert a secret message [1]. A major application of digital watermarking is the protection of intellectual property. In particular, digital watermarking provides robustness against analog hole piracy [2] when it is infeasible to protect original works of authorship using other protection techniques such as digital rights management (DRM).

Several studies have been conducted on digital watermarking, and there have been several benchmarking attempts at creating watermarking systems [3–7]. The Information Hiding and its Criteria (IHC) Committee has been working to improve this situation by promoting the development of digital watermarking technology [7]. Specifically, its objective is to develop standard evaluation criteria and sponsor watermark competitions based on these criteria. As a part of this objective, a benchmarking study was conducted on various types of image, video, and audio watermarking, based on standardized criteria [8–12].

This paper is a follow-up to our previous research [8], which described a robust video watermarking scheme based on discrete cosine transform (DCT) watermarking methodology. Herein, we present a robust watermarking system for video data and evaluate it based on standard criteria from IHC. The watermarking scheme achieves robustness against temporal desynchronization by adopting the adaptive selection of video frames for watermarking. In the adaptive frame selection, a secure watermark message is inserted, and the watermarked frame sequence is compressed based on the MPEG-4 part 10 codec (H.264). Subsequently, frames that are robust to lossy video compression are selected

for the watermark message embedding. Herein, we adopt a DCT based watermarking based on a spread-spectrum methodology that is robust to lossy compression such as MPEG and JPEG, the addition of Gaussian and white noise, digital/analog-analog/digital (DA-AD) conversion, and uniform scaling. Thus, the inserted watermark message and bit sequences can resist various video processes. The evaluation results indicate that our watermarking scheme achieves not only perceptual invisibility but also robustness against non-hostile processes. The assessment of the developed watermarking scheme satisfies the 3rd edition of the IHC video watermarking evaluation criteria [7].

The rest of this paper is organized as follows. In Section 2, we describe the background of this study, and in Section 3, the proposed watermarking method is provided. The experimental setup and results are presented in Section 4, and finally, in Section 5, we present some concluding remarks.

2. Background

Digital-to-analog conversion is the process of transforming a signal from the digital to the analog domain. After this process, the receiver can receive digital information through an analog-to-digital conversion. When outputting this intermediate transmission channel or video, a problem called an analog-hole might be encountered. An analog hole (also known as an analog loophole) is the perceived fundamental and inevitable vulnerability inherent in copyright protection schemes for non-interactive works in digital format. Digital multimedia can be exploited to duplicate a copy-protected work that is ultimately reproduced through analog means [2]. Once digital content has been converted into a human-perceptible format, it is relatively simple to digitally recapture an analog copy in an unrestricted format, thereby fundamentally circumventing all protections on copyrighted digital distribution and reproduction.

Digital watermarking provides robustness against analog hole piracy [2] when it is infeasible to protect original works of authorship through other protection techniques such as digital rights management (DRM). Owing to these characteristics, the application of digital watermarks as a technology in complementing DRM has been steadily receiving attention. A digital watermark is a technology that hides data in noise-resistant signals such as multimedia content. Many studies on watermarking technology applied to various contents have been conducted, such as watermarking based on neural networks [13], 3D meshes and prints [14,15], and spherical panoramic images [16].

Robust video watermarking methods based on various methodologies have been studied [17–22] Robustness against non-hostile video processing must be ensured to achieve copyright protection in video distribution cases. To provide robustness of the embedding patterns against lossy compression, various watermarking domain and primitives have been adopted, including a contourlet transform and a principal component analysis [17], gradient magnitude similarity deviation [18], and a hyperchaotic Lorentz system [19]. BhaBhardway et al. [23] proposed a significant frame selection and a quantization of the lifting wavelet transform coefficient difference based method for robust video watermarking. Huan et al. [22] demonstrated that the dual tree-complex wavelet (DT CWT) domain is effective for achieving a high robustness. For enhancing the robustness against lossy video compression, Barani et al. [21] used the coefficients of a contourlet transform to embed an encrypted watermark pattern using a singular value decomposition.

In the case of video watermarking, a pattern embedded as a secret message should be retrieved after applying various non-hostile and common video processes [24]. Video editing, photometric attacks, and spatial and temporal shifts (or desynchronization) are included in non-hostile video processes [25]. Photometric attacks are a process for modifying the pixel intensities in a frame. Noise addition and DA-AD conversions belong to this category. Another common treatment involves applying gamma correction to increase the contrast. To reduce the storage and transmission requirements, content owners re-encode their digital videos to a different compression ratio using a video compression algorithm. Video compression algorithms are based on a lossy compression on spatial and temporal domains of the media, such that the loss of information induced easily alters the performance of the watermarking algorithm. Similarly, customers can choose from standard video formats such as MPEG-1, MPEG-2, or MPEG-4. Spatial filtering and inter-frame filtering also have to be considered. A spatial shift or desynchronization can occur with changes in the display and spatial resolution. Primarily, temporal desynchronization occurs because of changes in the frame rate. In addition, video editing includes various methods from video editing tools, such as fade-and-dissolve, copy-and-paste, and splicing.

3. Proposed Method

3.1. Methodology

In our proposed scheme, we adopted DCT domain spread-spectrum watermarking [26] to work in a video frame watermarking and digital/analog–analog/digital (DA-AD) environment. A pseudo-random sequence of length N with zero-mean and unit variance was adopted as the watermark pattern. The general process of the watermarking method is summarized in Figure 1.

The frames were adaptively selected for embedding the secret message. A 1-s interval was selected for watermarking for every 7500 ms of the video content. Before embedding into the selected frames, a total of 512 pseudo-random patterns were generated to encode 16 bits of information, i.e., 256 pseudo-random patterns for encoding the front 8 bits of information, and another 256 pseudo-random patterns for the back 8 bits. The pseudo-random patterns are stored as a watermark pattern and will be used for the extraction step. Therefore, two watermark patterns are inserted into the selected video frames. Because the variations in the blue channel are less perceptible for the human visual system than those in the red and green channels [27], the pseudo-random patterns are inserted into the blue channel. In the watermark extraction stage, the primitive vectors that include the watermark pattern were extracted from the complete frames. If two watermark patterns are accurately detected, we decode 16 bits of information from these patterns using cross-correlation values with the 512 generated pseudo-random patterns.



Figure 1. Overall process of the proposed watermarking scheme: (a) Watermarking embedding step, and (b) extraction step.

3.2. Adaptive Frame Selection

With the proposed watermark embedding, pseudo-random patterns as a watermark are inserted into video frames that are robust against MPEG compression. The details of the adaptive frame selection process are summarized below.

To observe the compression effect of H.264 on a watermark extractor, we conducted a simple experiment. Using the algorithm introduced in Section 3.4, a sample pseudo-random pattern was repeatedly embedded into each frame. The modified frames were

then compressed using the MPEG-4 part 10 codec algorithm. Subsequently, the correlation responses in each video frame image were computed by adopting the extraction process described in Section 3.5. Figure 2 presents a sample of the H.264 compression effect on the correlation values as an extractor response. The watermark detector correlations are decreased in the middle frame section, which includes several video motions. Moreover, a degradation of the extractor responses becomes quite severe in a compressed video. In the case of a compressed video, the extractor responses are entirely degraded, and periodic peaks occur in the sequence of the extractor responses. These peaks generally appear in the intra-coded picture (I-frame) in a group of pictures (GOP) [28]. I-frames in GOP are more robust against lossy compression than other of video frames such as P and B-frames because they are less affected by intra-frame encoding with a motion estimation.

In our research, watermarks were efficiently embedded using this property. For example, the embedding of watermark patterns into video frames that are not robust to MPEG compression is inefficient. Thus, as shown in Figure 3, 1-s intervals with the highest correlation responses were selected from a 7.5-s interval. The selected video frames were watermarked during the embedding step, and the other frames remained unchanged to achieve a higher imperceptibility score.

A watermarked frame exists in any 15-s frame intervals in the video after the adaptive frame selection because the maximum interval of the time between two consecutive watermarked frame intervals is 13 s. Therefore, the proposed watermarking system ensures that 16 bits of secret data can be extracted from any 15-s interval in the compressed video frames.



Figure 2. H.264 compression effect on correlation values as an extractor response (test sequence #2 from ITE/ARIB video).



Figure 3. Sample of video sequence for watermarking. Embedding intervals are selected from every 7.5 s.

3.3. Data Bit Encoding Step

For the proposed watermarking scheme, two pseudo-random sequences *X* and *Y* were adopted for the watermark pattern. Here, $256(=2^8)$ unique codes were generated to represent 8 bits of information. Sequence *X* for encoding the front 8 bits of information, and sequence *Y* for encoding the back 8 bits, were used. From this, 16 bits of data were encoded using two sets, X and Y, of 256 zero-mean/unit-variance pseudo-random sequences as a watermark pattern.

Therefore, two pseudo-random sequences X_i and Y_j were selected in each sequence set for data encoding. Indexes *i* and *j* correspond to the input data bit information. For example, if the input bit code is 0xFF00, two sequences, X_i and Y_j , are selected, where i = 0xFF and j = 0x00.

3.4. Watermark Pattern Embedding

For the watermark pattern embedding, the DCT of frame *F* was computed. Subsequently, the *N* coefficients in the selected location were adopted to form watermark embedding vector *V*. The watermark embedding vector of the original DCT coefficients $V = \{v_1, v_2, v_3, ..., v_N\}$ was modified, and it became a vector of modified watermarked coefficients $V' = \{v'_1, v'_2, v'_3, ..., v'_N\}$, based on the following embedding rule:

$$v'_{i} = v_{i} + \alpha |v_{i}| \frac{x_{i} + y_{i}}{2}, \quad (1 \le i \le N),$$
 (2)

where $X = \{x_1, x_2, x_3, ..., x_N\}$ and $Y = \{y_1, y_2, y_3, ..., y_N\}$ are the selected watermarks. The watermark strength α denotes the scaling factor as the strength of the embedded pattern signal. As we described in a previous section, two watermarks X and Y were embedded in each selected video frame to encode 16 bits of information. The DCT coefficients in the mid-band spectrum were selected for the pattern embedding, such that a trade-off between perceptual invisibility and robustness against video compression was achieved. After the DCT coefficient embedding process, the watermarked frame F' was obtained using an inverse DCT.

3.5. Extraction of the Embedded Watermark

The extraction process of the embedded watermark pattern was carried out as follows. The DCT coefficients of a possibly corrupted watermarked frame F^* were computed, and vector V^* was extracted from the DCT coefficients in the pre-selected position during the embedding step. Next, the cross-correlation between the watermarks from the predefined sets X and Y and vector V^* were calculated and adopted as a measure of the presence of the watermark. The correlation *z* was defined as follows:

$$z = \frac{X \cdot V^*}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i \cdot v_i^*.$$
 (3)

The presence of watermark *X* was determined through a comparison between the threshold *T* and the calculated correlation *z*. Here, the threshold *T* was computed using

$$T = \frac{1}{20 \cdot N} \sum_{i=1}^{N} |v_i^*|.$$
(4)

To minimize the false-positive rate, the threshold value from the equation was obtained experimentally based on the evaluation results. The presence of each watermark among the watermark pattern sets X and Y in the target frame F^* was decided through thresholding. Accordingly, the presence of watermarks in all frames was checked. Subsequently, each detected watermark was matched to the proper 8 bits of data information, and all 16 bits of data were recovered. If there are more or fewer than two extracted watermarks, the data cannot be accurately recovered.

4. Experimental Results

4.1. Experimental Setup

In this section, the experimental setup used for the proposed method is described. We evaluated the proposed scheme based on standard criteria from IHC [7]. Figure 4 illustrates the general process of our evaluation for the benchmarking of our watermarking system. The evaluation included the processes of watermark pattern embedding, pattern extraction, video frame compression (MPEG-4 part 10), and digital/analog-analog/digital (DA-AD)

conversion. Finally, the average bit error rate (BER) of the extracted bit information was calculated. The proposed algorithm was implemented in a MATLAB environment (ver 2012b), and on a computer platform consisting of an Intel i7-3770 (3.40 GHz) CPU with a 16-GB main memory employed to assess the performance.



Figure 4. Overall test procedure of the proposed video watermarking system.

4.1.1. Benchmark Data Preparation

For our evaluation set, we used five benchmark video clips, i.e., ITE/ARIB Hi-Vision Test Sequence numbers 2, 8, 20, 23, and 46 [29]. First, we converted raw YPbPr data into an RGB uncompressed video using the following equation:

$$R = Y + 1.576P_r$$

$$G = Y - (0.159P_b + 0.334P_r)/0.701$$

$$B = Y + 1.826P_b.$$
(5)

The original 600 video frames of each clip had a duration of 20 s. The display resolution of the original frames is 1920×1035 . We cropped the last three rows because the MPEG-4 part 10 codec (H.264) in MATLAB automatically resizes the video resolution to 1920×1032 . To obtain a 15-s video, we removed the first 150 frames. Finally, we obtained uncompressed AVI videos with a display resolution of 1920×1032 at 30 frames per second with a 15-s length. The total size of each video is 2.49 GB (1, 430, 826 kbps).

4.1.2. Experimental Setup

The evaluation was conducted as follows. First, we embedded 16-bit data information into the original video sequences through the watermark embedding process. Next, the watermarked video clips were encoded with various compression quality factors using the MPEG-4 part 10 (H.264) codec.

Using a digital-to-analog (DA) converter, the compressed video sequences were converted into an analog signal. To process the DA conversion, a 'Fosmon HDMI to component video (YPbPr)/VGA & SPDIF output converter box' was adopted, as shown in Figure 5a. An analog component video (YPbPr) was selected as the format for the analog signal because YPbPr accommodates a full-HD display resolution (1080p).

The analog YPbPr signal was reconverted into a digital signal. In addition, the converted digital video was saved in an uncompressed YUV422 video format. For this process, we adopted 'skyHD captureX HDMI' as the AD converter, as shown in Figure 5b. The frame rate and display resolution of the converted digital video (1280×720) differed from those of the input video. Finally, we detected the embedded watermarks in the video frames using a watermark extraction algorithm. The extracted watermark patterns were

decoded into 16-bit data based on the indices of the predefined pseudo-random patterns, and the bit error rate (BER) of the bit data was computed.

To optimize the invisibility and robustness of the watermarked frames, several parameters of the proposed algorithm should be carefully set. Here, we selected a length of N = 25,000 from the fixed location of the DCT coefficient. It was empirically determined that this choice yields the maximum MPEG robustness while preserving the imperceptibility.



Figure 5. Converters for DA-AD conversion. (**a**) Fosmon HDMI to component video converter box, (**b**) skyHD captureX HDMI, and (**c**) DA-AD conversion testing environment.

4.2. Results

4.2.1. Image Quality Assessment

To evaluate the image quality of the watermarked frame, we inserted 16 bits of data into the original video clips under various scale factors α . The watermarked frames were generated to yield an average peak signal-to-noise ratio (PSNR) of more than 36 dB. Here, the PSNR was defined using the mean squared error between the $m \times n$ image *I* and watermarked image *W* as follows.

$$PSNR = 10 \cdot log_{10} \left(\frac{255^2}{MSE}\right) \tag{6}$$

where $MSE = \frac{1}{mn} \sum_{i}^{m} \sum_{j}^{n} [I(i, j) - W(i, j)]^2$. Figure 6 shows an example of the compressed original and watermarked frames.

To compensate for the limitations of the PSNR metric, regarding the watermark invisibility evaluation, we also evaluated the imperceptibility based on the SSIM metric [30].

After the watermark embedding, the video clips generated were compressed using the MPEG-4 codec (part 10 H.264), with quality factors of between 56 and 59, as depicted in Table 1.

The compressed video clips with quality factors of between 56 and 59 satisfied the bit rate of 0.01 better than the uncompressed video frame. In addition, using the same quality factor, the unwatermarked video clips were encoded. Both the watermarked video frames and the original video frames were decoded, and the PSNR was computed for each pair of luminance values as follows:

$$Y = 0.7152G + 0.0722B + 0.2126R.$$
(7)

Table 1. Quality factor to generate less than 1/100 of the original bit rate.

ITE/ARIB Test Number	Quality Factor
ITE/ARIB sequence #2	58
ITE/ARIB sequence #8	59
ITE/ARIB sequence #20	58
ITE/ARIB sequence #23	58
ITE/ARIB sequence #46	56



Figure 6. (Left) Compressed original frames based on H.264, and (Right) watermarked frames from evaluation Set-4 ($\alpha = 0.4$, SSIM 0.9945, PSNR 44.03 dB).

Table 2 and Figure 7 present the perceptual quality assessment results for various watermark strengths α , i.e., the bit error rate (BER), PNSR, and SSIM. No errors occurred in our assessment at $\alpha \ge 0.4$ according to the results. The resulting clip #2 exhibited the best detection performance with the highest perceptual score. However, the evaluation of video clip #23 showed the worst detection score because it contained a large number of motion objects throughout the entire playback. In addition, the PSNR and SSIM values of the uncompressed original and watermarked frames are shown in Table 3. These

Taat	Cat No	Cat 1	Set 2	Cat 2	Cat 1	Cat E	Set (Set 7	Cat 9
Test Set No.		Set-1	Set-2	Set-3	Set-4	Set-5	Set-6	Set-7	Set-8
WM Strength α		0.1	0.3	0.35	0.4	0.9	1.5	3	5
	PSNR(dB)	53.53	50.48	50.43	50.40	49.26	47.17	45.32	43.82
#2	SSIM	0.9994	0.9987	0.9987	0.9987	0.9983	0.9974	0.9949	0.9912
	BER	×	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	PSNR(dB)	48.35	46.62	46.59	46.58	41.66	41.22	39.40	37.80
#8	SSIM	0.9986	0.9980	0.9980	0.9979	0.9934	0.9920	0.9846	0.9710
	BER	×	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	PSNR(dB)	45.06	44.94	44.91	44.89	35.62	35.61	36.87	36.21
#20	SSIM	0.9958	0.9941	0.9954	0.9953	0.9611	0.9606	0.9626	0.9436
-	BER	×	×	0.000	0.000	0.000	0.000	0.000	0.000
	PSNR(dB)	40.75	33.7	33.7	33.69	32.28	32.31	31.93	30.82
#23	SSIM	0.9973	0.9865	0.9865	0.9864	0.9780	0.9771	0.9705	0.9346
	BER	×	×	×	0.000	0.000	0.000	0.000	0.000
	PSNR(dB)	41.79	44.66	44.63	44.58	41.06	41.22	35.65	35.41
#46	SSIM	0.9878	0.9944	0.9944	0.9943	0.9851	0.9857	0.9428	0.9360
-	BER	×	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	PSNR(dB)	45.89	44.08	44.05	44.03	39.98	39.50	37.83	36.81
average	SSIM	0.9957	0.9946	0.9946	0.9945	0.9831	0.9825	0.9710	0.9552
C	BER	×	0.400	0.200	0.000	0.000	0.000	0.000	0.000

metric scores demonstrate a visual degradation of the video sequences when applying MPEG compression.

Table 2. Evaluation result for various scale factor α .

Compression quality factor—#2: 58, #8: 59, #20: 58, #23: 58, #46: 56; ×: watermark is not detected (BER is considered as 1.000).





Figure 7. Quality assessment result for various watermark scale factor *α*. (a) PSNR, and (b) SSIM.

Test	Set No.	Set-1	Set-2	Set-3	Set-4	Set-5	Set-6	Set-7	Set-8
WM St	trength α	0.1	0.3	0.35	0.4	0.9	1.5	3	5
Average	PSNR(dB)	76.94	70.28	69.70	68.91	65.07	63.17	57.56	54.00
	SSIM	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9984

4.2.2. Robustness Assessment

A watermarking scheme must be robust, which implies that an embedded pattern should be difficult to damage through a non-hostile signal process, such as MPEG compression or digital/analog–analog/digital conversion. Thus, an embedded watermark pattern must incur little damaged from these non-hostile signal processes. Tolerance is demonstrated through the calculation of the average BER for the extracted watermark information after undergoing video processes. Evaluations were conducted with various watermarking factors, i.e., $\alpha = 0.4, 0.9$, and 1.5.

Table 4 lists the robustness results for the benchmark Set-4, where the watermark strength $\alpha = 0$, the PSNR is 44.02 dB, and the SSIM is 0.9945. With an MPEG quality factor of zero, all watermarks were damaged and no data were detected. It is important to note that video compression with a factor of zero triggers a substantially strong level of content loss, and the factor is not adopted, except under extremely unique circumstances. Regarding the test conducted with an MPEG quality factor of 10, regardless of the negligible watermark strength, we succeeded in detecting the watermark in both experimental videos #2 and #8. This is because the motion was relatively small in these videos; hence, the watermark damage owing to MPEG compression also appeared to be small. At a certain level of MPEG compression (qf > 40), all watermark bits were successfully detected. Note that a PSNR of 44.08 dB is more than 6 dB higher than 38.0 dB, which is generally regarded as the minimum distortion value for which a human can detect a watermark, and thus no distortion can be detected by the human eye.

Table 4. Robustness results for the benchmark Set-4 with the watermark strength $\alpha = 0.4$ (PSNR 44.03 dB, SSIM 0.9945, $\alpha = 0.4$).

MPEG	Quality Factor	50	40	30	20	10	0
#2	BER	0.000	0.000	0.000	0.000	0.000	×
	compressibility	1/139	1/191	1/232	1/307	1/432	1/1029
#9	BER	0.000	0.000	0.000	0.000	0.000	0.000
#0	compressibility	1/139	1/193	1/253	1/321	1/445	1/606
#20	BER	0.000	0.000	0.000	0.000	×	×
#20	compressibility	1/139	1/190	1/253	1/321	1/444	1/460
#23	BER	0.000	0.000	×	0.000	×	×
	compressibility	1/138	1/193	1/249	1/322	1/400	1/395
#16	BER	0.000	0.000	0.000	×	×	×
#46	compressibility	1/129	1/183	1/236	1/296	1/386	1/443
Average	BER	0.000	0.000	0.200	0.200	0.600	0.800
	compressibility	1/137	1/190	1/245	1/313	1/422	1/587
(1) watermank is not detected (PED is considered as 1,000)							

 \times : watermark is not detected (BER is considered as 1.000).

Tables 5 and 6 list the robustness results and compressibility of the various MPEG quality factors. In the evaluation with a watermark strength of 0.9, the embedded data bits were successfully detected from four videos, #2, #8, #20, and #23, whereas the mean compressibility was 1/604. At a watermark strength of 1.5, the embedded watermark information was accurately detected for all evaluated quality factors, as shown in Table 6.

MPEG	Quality Factor	50	40	30	20	10	0
#2	BER	0.000	0.000	0.000	0.000	0.000	0.000
	compressibility	1/139	1/188	1/236	1/310	1/444	1/1001
#8	BER	0.000	0.000	0.000	0.000	0.000	0.000
	compressibility	1/139	1/192	1/253	1/321	1/446	1/631
#20	BER	0.000	0.000	0.000	0.000	0.000	0.000
	compressibility	1/139	1/190	1/239	1/317	1/438	1/501
#23	BER	0.000	0.000	0.000	0.000	0.000	0.000
	compressibility	1/138	1/191	1/250	1/322	1/419	1/413
#16	BER	0.000	0.000	0.000	0.000	0.000	×
#46	compressibility	1/130	1/186	1/236	1/297	1/407	1/473
Average	BER	0.000	0.000	0.000	0.000	0.000	0.200
	compressibility	1/137	1/189	1/243	1/313	1/431	1/604

Table 5. Robustness results for the benchmark Set-5 with the watermark strength $\alpha = 0.9$ (PSNR 39.98 dB, SSIM 0.9831).

 \times : watermark is not detected (BER is considered as 1.000).

Table 6. Robustness results for the benchmark Set-6 with the watermark strength $\alpha = 1.5$ (PSNR of 39.50 dB, SSIM of 0.9825).

MPEG	Quality Factor	50	40	30	20	10	0
# 2	BER	0.000	0.000	0.000	0.000	0.000	0.000
#2	compressibility	1/136	1/194	1/248	1/315	1/413	1/1029
#0	BER	0.000	0.000	0.000	0.000	0.000	0.000
#0	compressibility	1/140	1/194	1/253	1/323	1/446	1/603
#20	BER	0.000	0.000	0.000	0.000	0.000	0.000
	compressibility	1/134	1/193	1/254	1/317	1/446	1/460
#22	BER	0.000	0.000	0.000	0.000	0.000	0.000
#23	compressibility	1/137	1/193	1/253	1/320	1/383	1/380
#16	BER	0.000	0.000	0.000	0.000	0.000	0.000
#46	compressibility	1/131	1/184	1/237	1/297	1/387	1/443
Average	BER	0.000	0.000	0.000	0.000	0.000	0.000
	compressibility	1/136	1/192	1/249	1/314	1/415	1/583

5. Conclusions

In this study, we proposed a robust video watermarking system, and presented detailed results based on a public video benchmark. The assessment of the developed watermarking scheme satisfied the 3rd edition of the IHC video watermarking evaluation criteria. The proposed system achieved robustness against non-hostile video processes. To evaluate the performance of the proposed watermarking, experiments based on video watermarking evaluation criteria were conducted under a DA-AD conversion setup. The experimental results showed that, in addition to being robust against non-hostile video processes, the proposed method also achieves invisibility. However, because spread spectrum watermarking is sensitive to attacks, the security of the proposed scheme is limited in a few aspects [31]. Enhancing the security of our watermarking is an important research direction for future studies. In this research, we focused solely on photometric distortions,

such as the compression of video and DA-AD conversion. Regarding the robustness of the watermark, a geometric transformation or collusion attack will also be needed in our future studies.

Author Contributions: Conceptualization, methodology, software, writing—review and editing, funding acquisition, J.-U.H. The author has read and agreed to the published version of the manuscript.

Funding: This research was supported by a Hallym University Research Fund, 2020 (HRF-202011-010).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: This research was supported by a Hallym University Research Fund, 2020 (HRF-202011-010). The author would like to express special thanks to Jinseok Park, Seung-Hun Nam, and Do-Guk Kim for their assistance with the initial setup of the experiment and the data preprocessing.

Conflicts of Interest: The author declares no conflict of interest.

References

- 1. Cox, I.J.; Miller, M.L.; Bloom, J.A.; Fridrich, J.; Kalker, T. *Digital Watermarking and Steganography*; Elsevier Science Limited: Burlington, MA, USA, 2008.
- Diehl, E.; Furon, T. Watermark: Closing the analog hole. In Proceedings of the 2003 IEEE International Conference on Consumer Electronics, Los Angeles, CA, USA, 17–19 June 2003; pp. 52–53. [CrossRef]
- Wang, K.; Lavoué, G.; Denis, F.; Baskurt, A.; He, X. A benchmark for 3D mesh watermarking. In Proceedings of the 2010 Shape Modeling International Conference, Aix-en-Provence, France, 21–23 June 2010; pp. 231–235.
- 4. van Walree, P.A.; Socheleau, F.X.; Otnes, R.; Jenserud, T. The watermark benchmark for underwater acoustic modulation schemes. *IEEE J. Ocean. Eng.* **2017**, *42*, 1007–1018. [CrossRef]
- Petitcolas, F.A.; Steinebach, M.; Raynal, F.; Dittmann, J.; Fontaine, C.; Fates, N. Public Automated Web-Based Evaluation Service for Watermarking Schemes: Stirmark Benchmark; International Society for Optics and Photonics: San Jose, CA, USA, 2001; Volume 4314, pp. 575–584.
- Steinebach, M.; Petitcolas, F.A.; Raynal, F.; Dittmann, J.; Fontaine, C.; Seibel, S.; Fates, N.; Ferri, L.C. StirMark benchmark: Audio watermarking attacks. In Proceedings of the international conference on information technology: Coding and computing, Las Vegas, NV, USA, 2–4 April 2001; pp. 49–54.
- 7. Iwamura, K.; Kawamura, M.; Kuribayashi, M.; Iwata, M.; Kang, H.; Gohshi, S.; Nishimura, A. Information hiding and its criteria for evaluation. *IEICE Trans. Inf. Syst.* 2017, 100, 2–12. [CrossRef]
- Hou, J.U.; Park, J.S.; Kim, D.G.; Nam, S.H.; Lee, H.K. Robust video watermarking for MPEG compression and DA-AD conversion. In Proceedings of the 1st International Workshop on Information Hiding and its Criteria for Evaluation, New York, NY, USA, 3 June 2014; pp. 2–8. [CrossRef]
- Tozuka, H.; Yoshida, M.; Fujiwara, T. Salt-and-pepper image watermarking system for IHC evaluation criteria. In Proceedings of the 1st International Workshop on Information Hiding and its Criteria for Evaluation, New York, NY, USA, 3 June 2014; pp. 31–36. [CrossRef]
- 10. Kang, H.; Iwamura, K. Information hiding method using best DCT and wavelet coefficients and its watermark competition. *Entropy* **2015**, *17*, 1218–1235. [CrossRef]
- Sakuma, R.; Kang, H.; Iwamura, K.; Echizen, I. Digital Watermarking Scheme Based on Machine Learning for the IHC Evaluation Criteria. In Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Matsue, Japan, 12–15 August 2017; pp. 370–380. [CrossRef]
- 12. Hamamoto, I.; Kawamura, M. Neural watermarking method including an attack simulator against rotation and compression attacks. *IEICE Trans. Inf. Syst.* 2020, 103, 33–41. [CrossRef]
- 13. Kim, W.H.; Kang, J.; Mun, S.M.; Hou, J.U. Convolutional neural network architecture for recovering watermark synchronization. *Sensors* **2020**, *20*, 5427. [CrossRef] [PubMed]
- 14. Hou, J.U.; Yu, I.J.; Lee, H.K. Collusion attack resilient 3d mesh watermarking based on anti-collusion fingerprint code. *Appl. Sci.* **2018**, *8*, 1040. [CrossRef]
- 15. Hou, J.U.; Kim, D.G.; Lee, H.K. Blind 3D mesh watermarking for 3D printed model by analyzing layering artifact. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 2712–2725. [CrossRef]
- 16. Kang, J.; Hou, J.U.; Ji, S.; Lee, H.K. Robust spherical panorama image watermarking against viewpoint desynchronization. *IEEE Access* **2020**, *8*, 127477–127490. [CrossRef]
- 17. Chen, L.; Zhao, J. Contourlet-based image and video watermarking robust to geometric attacks and compressions. *Multimed. Tools. Appl.* **2018**, *77*, 7187–7204. [CrossRef]
- 18. Himeur, Y.; Boukabou, A. A robust and secure key-frames based video watermarking system using chaotic encryption. *Multimed. Tools. Appl.* **2018**, *77*, 8603–8627. [CrossRef]
- 19. Cao, Z.; Wang, L. A secure video watermarking technique based on hyperchaotic Lorentz system. *Multimed. Tools. Appl.* **2019**, 78, 26089–26109. [CrossRef]

- 20. Mareen, H.; De Praeter, J.; Van Wallendael, G.; Lambert, P. A novel video watermarking approach based on implicit distortions. *IEEE Trans. Consum. Electron.* **2018**, *64*, 250–258. [CrossRef]
- Barani, M.J.; Ayubi, P.; Valandar, M.Y.; Irani, B.Y. A blind video watermarking algorithm robust to lossy video compression attacks based on generalized Newton complex map and contourlet transform. *Multimed. Tools. Appl.* 2020, 79, 2127–2159. [CrossRef]
- Huan, W.; Li, S.; Qian, Z.; Zhang, X. Exploring Stable Coefficients on Joint Sub-bands for Robust Video Watermarking in DT CWT Domain. *IEEE Trans. Circuits Syst. Video Technol.* 2021. [CrossRef]
- 23. Bhardwaj, A.; Verma, V.S.; Jha, R.K. Robust video watermarking using significant frame selection based on coefficient difference of lifting wavelet transform. *Multimed. Tools. Appl.* **2018**, 77, 19659–19678. [CrossRef]
- 24. Asikuzzaman, M.; Pickering, M.R. An overview of digital video watermarking. *IEEE Trans. Circuits Syst. Video Technol.* 2017, 28, 2131–2153. [CrossRef]
- 25. Doërr, G.; Dugelay, J.L. A guide tour of video watermarking. Signal Process. Image Commun. 2003, 18, 263–282. [CrossRef]
- Barni, M.; Bartolini, F.; Cappellini, V.; Piva, A. A DCT-domain system for robust image watermarking. *Signal Process.* 1998, 66, 357–372. [CrossRef]
- 27. Jain, A.K. Fundamentals of Digital Image Processing; Prentice Hall: Englewood Cliffs, NJ, USA, 1988.
- Bhaskaran, V.; Konstantinides, K. Image and Video Compression Standards: Algorithms and Architectures; Springer: Berlin, Germany, 1997.
- 29. Matsuda, I.; Ikegawa, H. ITE/ARIB hi-vision test sequence 2nd edition: Production notes and tips for use (part II). *Kyokai Joho Imeji Zasshi/J. Inst. Image Inf. Telev. Eng.* 2010, 64, 201–204.
- Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* 2004, 12, 600–612. [CrossRef] [PubMed]
- 31. Chaumont, M. *Ensuring Security of H.264 Videos by Using Watermarking;* International Society for Optics and Photonics: Orlando, FL, USA, 2011; Volume 8063, p. 806303. [CrossRef]