Robust Video Watermarking Based on Temporal Modulation with Error Correcting Code

Sang-Woo Lee, Byung-Gil Lee and Dae-Hee Seo

Information Security Research Division,
Electronics and Telecommunications Research Institute
{ttomlee, bglee, dhseo}@etri.re.kr

Abstract

This article presented a novel robust video watermarking algorithm that satisfies robustness and real-time performance requirements in client-side embedding environments such as IP TV set-top boxes. A watermarked video is generated by temporally modulating the mean chrominance value of each chrominance channel in relation to the watermark pattern according to the watermark message. The watermark pattern is generated based on a histogram analysis of the luminance channel. To improve the robustness, we employed the BCH codes followed by repetition codes during the encoding of the watermark message. The experimental results showed that the proposed watermarking algorithm is resistant to geometrical distortions such as rotation, cropping, scaling, and projective transforms. The proposed watermarking algorithm is also robust against signal processing attacks including compression format conversion and other temporal attacks.

Keywords: *Video watermarking, robust watermarking, multimedia security*

1. Introduction

Due to the rapid development of the Internet, a lot of multimedia content has become available through several services such as IP TV. Since digital media is easily reproduced and manipulated, the demand for protecting copyright ownership has increased in IP TV services and video streaming services [1-3]. In particular, the unauthorized copying and illegal distribution of copyrighted content have greatly increased in the area of digital video. Therefore, copyright protection technology is a hot research topic [4, 5]. A cryptographic technique is used to control access to multimedia content. However, once such content is decrypted, there are no countermeasures to block illegal reproduction and distribution. Based on this situation, digital watermarking technology could be a complementary technique of cryptography-based protection methods [6, 7].

Digital watermarking refers to the process of embedding secret information, called a watermark, into digital content. The main requirements of digital watermarking include imperceptibility, capacity, and robustness. Imperceptibility means that embedding a watermark should not have noticeable artifacts in the watermarked video. Capacity refers to the amount of information that is embedded by a watermarking algorithm in a host video. The robustness of a watermarking scheme refers to the detectability of watermarked information from watermarked content that has been modified by various attacks such as geometrical distortions and signal processing attacks. Since video processing requires significant computing resources, simplicity is another important requirement of video watermarking technology. In particular, the real-time performance is a critical requirement in the case of

client-side watermark embedding in IP TV set-top boxes. It is important to note that these requirements are usually in conflict with one another. For example, if we increase watermark energy to increase the robustness, the visual quality of the watermarked video may be decrease. Therefore, we should take this trade-off into consideration in designing a watermarking system on targeted applications.

For a long time, digital watermarking has focused mainly on still images [8-13]. Currently, however, the study of watermarking algorithms for video has become a hot research topic. But while watermarking technologies for still images and video are similar, they are not identical.

Several studies [14-16] have suggested robust video watermarking algorithms that do consider the temporal dimension of video stream. Leest, *et al.*, [14] proposed a watermark embedding method by increasing or decreasing the mean luminance value of each pixel of a frame. However, their algorithm requires the original video during the detection procedure. Do, *et al.*, [15] suggested a blind watermarking algorithm based on temporal modulation, but their algorithm is too complex to support real-time performance. Lee, *et al.*, [16] proposed watermarking algorithms utilizing frame skipping or temporal feature modulation. They showed that their algorithm is robust against compression and temporal attacks. However, their scheme is not robust to other types of attack such as geometric distortions.

This paper presents a robust video watermarking algorithm based on temporal modulation with an error correction code. Basically, we employ the mean value of each frame to support high robustness of the watermarking algorithm. The mean value of each frame is known to be resistant to geometrical distortions, signal processing attacks, and temporal frame dropping [17, 18]. Temporal modulation refers to a watermark embedding technique that embeds a watermark bit by differentiating the pixel values between frames, depending on the watermark message. Furthermore, we employ an error correcting code to improve robustness. One study [15] used a temporal modulation scheme, but their algorithm cannot support real-time embedding performance for video as they employed a complex masking method based on a human visual system (HVS) to improve robustness. On the contrary, our watermarking algorithm does not execute a complex masking method, and thus it can embed a watermark message into a host video in real time. To compensate a decrease in terms of robustness, we used an error correcting code (ECC), which is composed of the BCH code followed by a repetition code. Experimental results showed that our watermarking algorithm satisfies the requirements of both robustness and real-time performance during the embedding and detection procedure.

This paper is organized as follows. The proposed watermarking algorithm is described in Section II, where we explain the embedding and detection procedures of the proposed algorithm in detail. In Section III, the experimental results and an analysis are discussed. Finally, we end with a brief conclusion in Section IV.

2. Proposed Watermarking Algorithm

In general, a watermarking algorithm consists of two procedures: watermark embedding and watermark detection. In this section, we describe the proposed embedding procedure first. The blind detection procedure is then explained.

The proposed watermark embedding algorithm is composed of three steps: watermarking pattern generation, temporal modulation, and watermark message encoding.

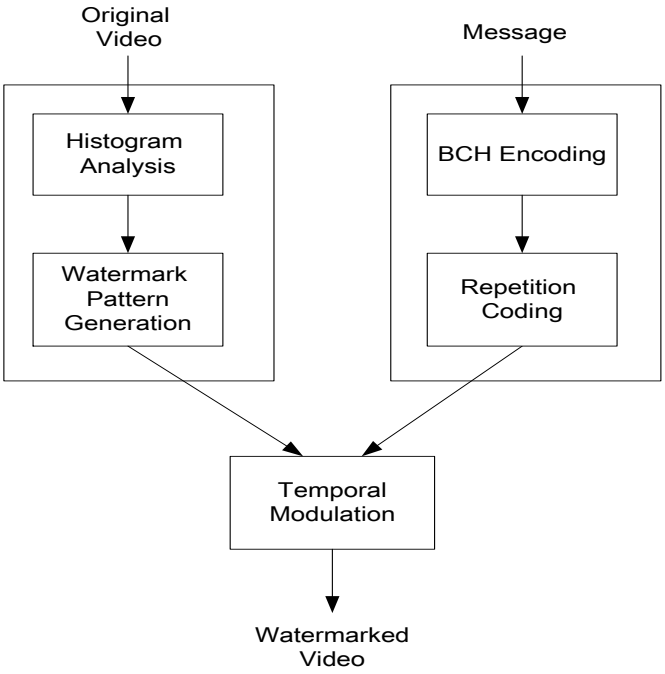


Figure 1. Watermark embedding procedure

Figure 1 shows the watermark embedding procedure of the proposed algorithm. To generate the watermark patterns, we measure the mean luminance value of each frame because the histogram distribution of an image is approximately invariant under a geometric attack. In our watermarking algorithm, each frame of size $M \times N$ is considered. $R(x, y)$, $G(x, y)$, and $B(x, y)$ indicate three primary color values of the RGB color model. Here, (x, y) denotes the pixel position, where $0 \le x < M$, $0 \le y < N$. The watermark pattern generation is performed as follows:

1. All $R(x, y)$, $G(x, y)$, and $B(x, y)$ are transformed into $Y(x, y)$, $Cb(x, y)$, and $Cr(x, y)$ in the YCbCr color model.
2. Calculate the average values of $Y(x, y)$ such as

$$Y_{mean}(x, y) = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} Y(x, y) \tag{1}$$

3. The generated watermark pattern $W(x, y)$ is given by

$$W(x, y) = \begin{cases} +\alpha & \text{if } Y(x, y) \ge Y_{mean}(x, y) + (Y_{mean}(x, y) \times \beta) \\ -\alpha & \text{if } Y(x, y) < Y_{mean}(x, y) - (Y_{mean}(x, y) \times \beta) \\ 0 & \text{else.} \end{cases} \tag{2}$$

Here, α is the gain constant. We divide the watermark pattern into two parts based on the mean luminance value. One part of the watermark pattern is marked with positive value, $+\alpha$, and the other part is marked with the negative value, $-\alpha$. β is the parameter used to choose the

number of pixels that will be modified in one entire frame. The remaining part of the watermark pattern is set to zero. The proposed watermark patterns that are generated by analyzing the mean luminance value based on the histogram do not depend on pixel location.

The watermark information is embedded in each frame of the original video by temporally modulating the mean pixel values. This is done by adding or subtracting the watermark pattern to or from the frame. The proposed temporal modulation can be presented as follows:

For i = 1 to the number of frames in the video

$$Cb_i'(x,y) = Cb_i(x,y) - m \times W_i(x,y) \text{ if } i \le \frac{K}{2},$$
$$Cb_i'(x,y) = Cb_i(x,y) + m \times W_i(x,y) \text{ if } i > \frac{K}{2}, \quad (3)$$

where K refers to the number of consecutive frames; m denotes the watermark information bit and can be 1 or -1; $Cb_i(x,y)$ denotes the chrominance values in the i-th original frame; and $Cb_i'(x,y)$ denotes the watermarked chrominance values in the i-th watermarked frame. One bit of watermark information can be embedded in K consecutive frames. We assume that the average values of K consecutive frames do not vary too greatly.

To improve robustness, we encode a watermark message prior to embedding it into the original video through temporal modulation. The simplest method to prevent errors is to embed the watermark message into a video repeatedly. ECC codes can also be used to recover any bit errors. Here, we consider the BCH and repetition codes. Binary BCH codes can be constructed using parameters (n,k,t), where n represents the length of the code word, k is the length of the message size, and t represents the number of bit errors the code can correct.

In our proposed watermarking algorithm, we use hybrid coding, which refers to a concatenation of repetition and BCH codes. There are two combinations: repetition codes followed by BCH codes and BCH codes followed by repetition codes. We employ the latter case because the bit error rate of the received code is decreased by repetition, and the BCH coding can then correct up to t errors. In the first method, the BCH decoder cannot correct a received code that has more than t errors. In addition, a received code that has fewer than t errors can be corrected using the BCH decoder, and thus the repetition code is useless. This hybrid coding of watermark messages improves the robustness during the watermark detection procedure for modified videos by minimizing the bit error rates from various kinds of attacks.

The proposed watermark detection procedure consists of three steps: watermark pattern generation, computation of the correlation between the watermark pattern and the watermarked frame, and message decoding. The proposed watermark detection does not require the original video. That is, the proposed watermarking algorithm supports blind detection. First, the watermark pattern generation step is performed, which is the same step as in the watermark embedding procedure. Then, the correlation between the generated watermark pattern and watermarked frame is computed. To detect the watermark information correctly, it is important to find the location of the first frame of K consecutive frames. This is called temporal synchronization in this paper. For temporal synchronization, we embed the

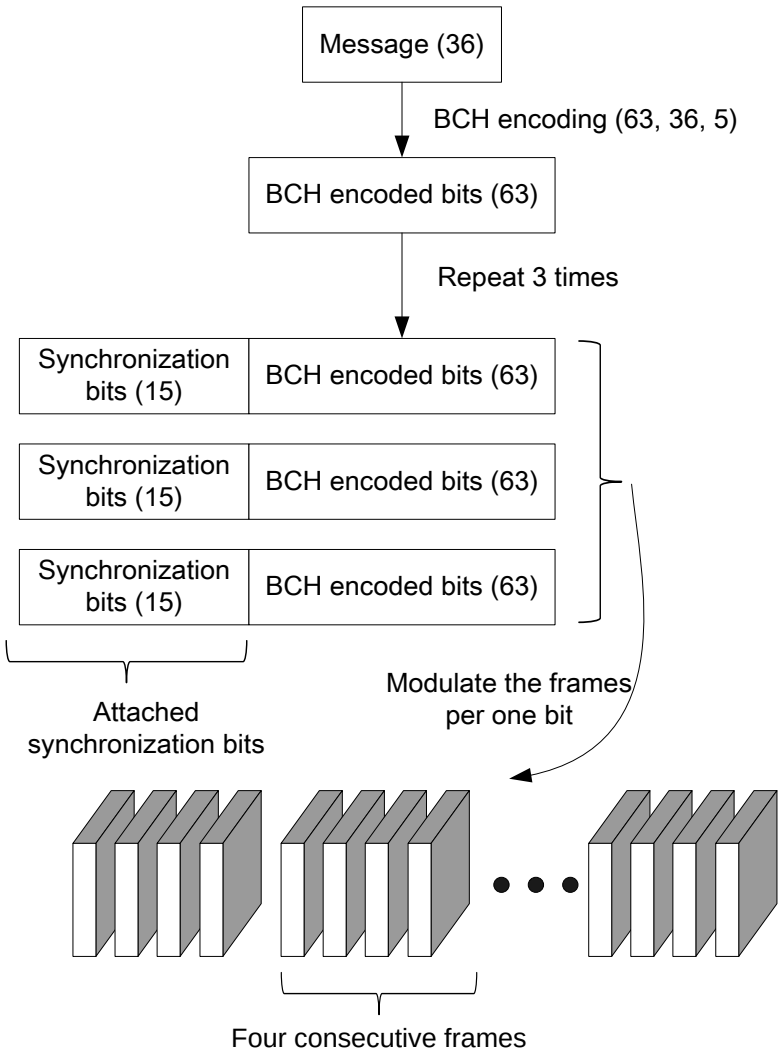


Figure 2. The Watermark Message Format

synchronization bits, which make up a pseudo-random sequence, before the watermark information bits. In the watermark detection procedure, we compute the Hamming distance between the extracted watermark bits and synchronization bits. After the synchronization sequence is detected, the watermark information bits are extracted by calculating the correlation between the generated watermark pattern and the watermarked frame. After extracting a pre-defined number of watermark information bits in every K consecutive frames, we begin finding the synchronization sequence again.

3. Experimental Results and Analysis

In this section, we explain the performance of the proposed watermarking algorithm. Figure 2 shows the watermark message format in the watermark embedding procedure. The length of the message was set as 36 bits. This is a reasonable length of watermark information as we consider that the watermarking data payload is required to be a

minimum of 35 bits in digital cinema specifications [19]. We adopt BCH (63, 36, 5) codes and set the repetition number as 3.

Table 1. Test videos

No.	Test video	Size	FPS	Frame
1	Highway	352x288	29.97	2000
2	Paris	352x288	29.97	1065
3	Drama	640x352	29.97	1000
4	Music show 1	640x352	29.97	1000
5	Music show 2	640x352	29.97	1000

For embedding the algorithm parameters, we set α = 1, β = 0.25, and K = 4. The proposed watermarking algorithm was tested on five test videos, as shown in Table 1. Videos 1 and 2 are frequently used as video watermarking tests. Video 3 is a drama. Videos 4 and 5 are music shows. We choose different video genres because video characteristics, which include frequent scene changes, are different according to their genre. The mean luminance values and the mean chrominance values of the test videos are illustrated in Figures 3. Two famous video clips, Highway and Paris, have an approximately constant mean luminance and chrominance. Thus, we chose other videos that have more scene changes than Videos 1 and 2 in order to evaluate the robustness of our algorithm. In particular, Videos 4 and 5 have much higher variations of luminance and chrominance than the other videos because the stage lighting of music shows change quite frequently.

To evaluate the robustness to various attacks, we measured average bit error rates of the attacked videos in the watermark detection procedure. The robustness of each test video is summarized in Table 3. First, we evaluated the robustness to geometric distortions. A rotation attack was performed by rotating the watermarked video clockwise and anti-clockwise by 15 degrees. For a cropping attack, we kept the central part of the watermarked video. We cut 10 percent of the upper, lower, left, and right sections of the watermarked video. We also scaled the watermarked video to half of its original size. Recorded copies from a camcorder suffer geometric transformations including rotation, scaling changes, and non-linear transformation caused by an angle difference between the camcorder and screen. A projective transform is the modeling of video distortion by a camcorder recording. The experimental results show that the proposed watermarking algorithm is robust to geometric distortions such as rotation, scaling, cropping, and projective transform.

In addition, we evaluated the robustness to temporal modification and signal processing attack. We changed the original frame rates to 20 fps and 40 fps. After restoring the frame rate of the attacked video to its original rate, we performed the watermarking detection algorithm. The proposed watermarking algorithm was robust to frame rate conversion and low-pass and high-pass filtering attack. Finally, the watermarked videos were compressed into H.264/AVC format. Since end-users are likely to display downloaded content on portable devices that have smaller storage, they will commonly convert their video into MPEG-4 format [20, 21]. The experimental results show that the proposed watermarking algorithm is robust to compressed format conversion.

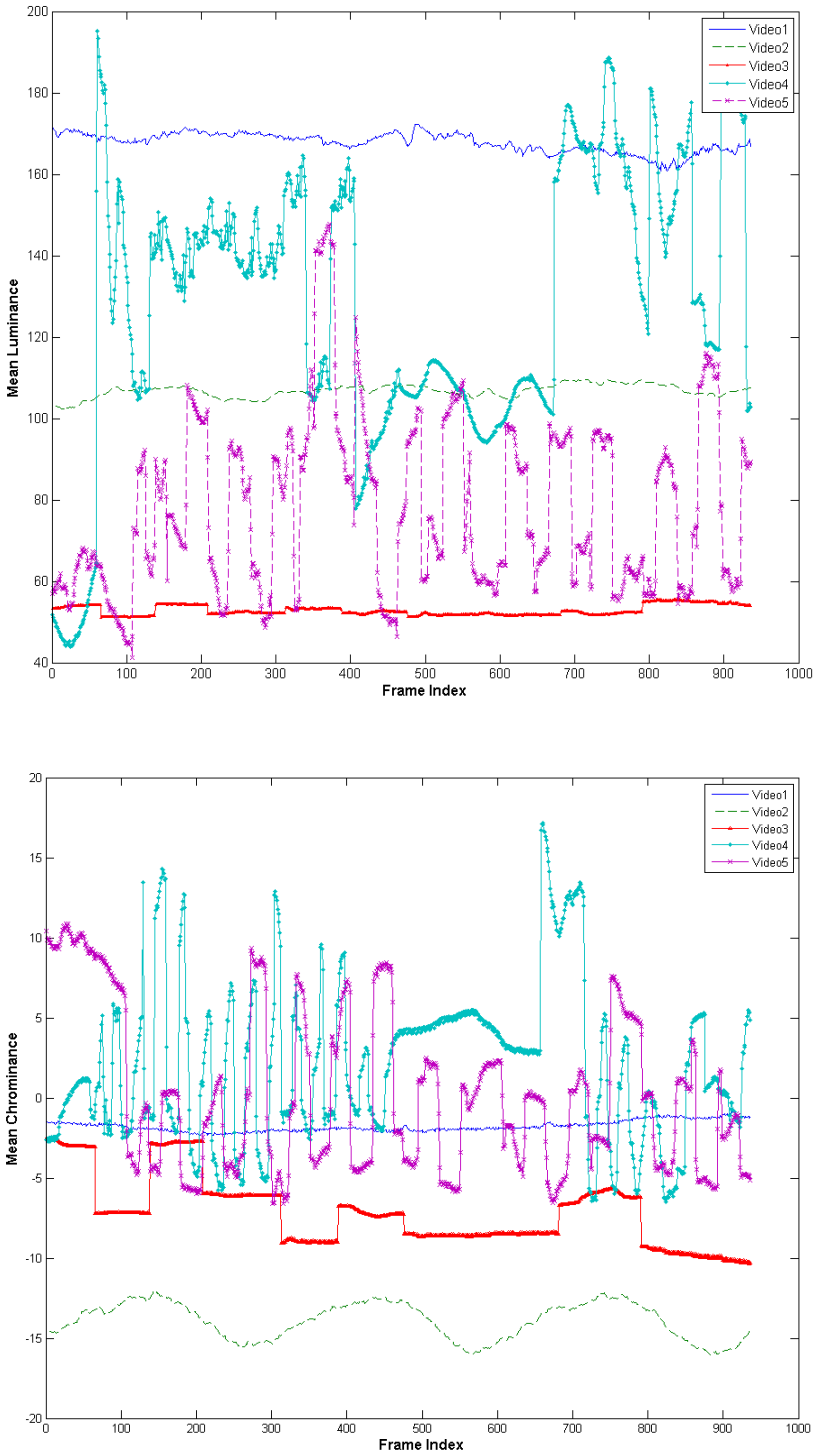


Figure 3. Mean luminance values and mean chrominance values for the test videos

Table 2. PSNRs

β	Min. PSNR	Average PSNR
0.25	48.74	50.84
1	48.13	48.18

Table 3. Bit error rates of test videos

	Video 1 2 3	Video 4		Video 5	
	Without ECC	Without ECC	With ECC	Without ECC	With ECC
Rotation 15°	0	11.42	0	8.56	0
Cropping 10%	0	11.42	0	11.42	0
Scaling x 0.5	0	11.72	0	8.57	0
Projective	0	17.13	0	11.42	0
FRC (20fps)	0	2.85	0	5.71	0
FRC (40fps)	0	11.42	0	5.71	0
Smoother	0	14.27	0	8.57	0
Sharpening	0	11.42	0	5.71	0
H.264/AVC	0	17.13	0	8.57	0

In the watermark embedding procedure, we assume that the mean value of each frame in K consecutive frames is approximately constant. However, some videos such as Videos 4 and 5 have high variations in luminance and chrominance channel, as shown in Figures 3. This causes high bit error rates in the watermark detection procedure of the attacked video. To compensate for this drawback, we employed the BCH and repetition codes. The experimental results of Videos 4 and 5 with and without ECC are shown in Table 3. For the tests on Videos 4 and 5 without ECC, we investigated high bit error rates in the watermark detection procedure against many kinds of attacks. However, we were unable to find bit errors in the tests with ECC. At this point, it is important to note that error-correcting ability is dependent on the characteristics of the BCH and repetition codes. In our experimental case, the number of bit errors was less than the error correcting ability of BCH (63, 36, 5) codes repeated three times. We should adopt BCH codes that have a higher error correcting ability in order to recover more error bits. However, this decreases the capacity of the watermarking algorithm. Therefore, the error correcting codes should be chosen considering a trade-off between capacity and robustness.

Table 2 shows the minimum and average PSNRs of five videos in relation with β. These PSNR values in Table 2 are above 48 dB. In general, it is well known that human eyes have difficulty in recognizing visual artifacts of modified videos that have PSNR values of higher than approximately 37 dB. Therefore, we can say that the watermarked videos have sufficient visual quality in terms of PSNR.

A real-time video watermarking algorithm should embed a watermark message into an original video at 30 fps in 33.3 milliseconds per frame. The computation time per frame for the embedding and detection procedures is less than 13 milliseconds. Therefore, we can conclude that our watermarking algorithm can embed a watermark message in real time.

4. Conclusion

This article presented a novel robust video watermarking algorithm that satisfies robustness and real-time performance requirements in client-side embedding environments such as IP TV set-top boxes. To improve the robustness, we employed the error correcting codes. The experimental results showed that the proposed watermarking algorithm is resistant to geometrical distortions and temporal attacks. It is well known that there are trade-offs between the several requirements in video watermarking technology. From this viewpoint, it is important to note that the proposed algorithm satisfies the requirements of invisibility and real-time performance, and it supports high robustness against not only a specific attack but also many kinds of attacks.

Acknowledgements

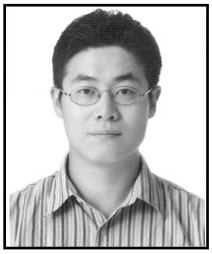
This work was supported by the IT R&D program of MKE/KEIT. [10041686, Cooperative Control Communication/Security Technology and SoC Development for Autonomous and Safe Driving System].

References

[1] S. Lian and Z. Liu, "Secure Media Content Distribution Based on the Improved Set-Top Box in IPTV, IEEE Transactions on Consumer Electronics, vol. 54, no. 2, **(2008)**
[2] S. Park, J. Jeong and T. Kwon, "Contents Distribution System Based on MPEG-4 ISMACryp in IP Set-Top Box Environments", IEEE Transactions on Consumer Electronics, vol. 52, no. 2, **(2006)**.
[3] Y. -H. Kim, J. Shin and J. Park, "Design and Implementation of a Network-Adaptive Mechanism for HTTP Video Streaming", ETRI Journal, vol. 35, no. 1, **(2013)**.
[4] H. K. Lee and J. Kim, "Extended Temporal Ordinal Measurement Using Spatially Normalized Mean for Video Copy Detection", ETRI Journal, vol. 32, no. 3, **(2010)**.
[5] P. K. Dhar and J. -M. Kim, "Digital Watermarking Scheme Based on Fast Fourier Transformation for Audio Copyright Protection", IJSIA, vol. 5, no. 2, **(2011)**, pp. 33-48.
[6] G. Doërr and J. Dugelay, "A Guide Tour of Video Watermarking. Signal Processing", Image Communication, vol. 18, no. 4, **(2003)**, pp. 263-282.
[7] S. Choi, J. -W. Han and H. Cho, "Privacy-Preserving H.264 Video Encryption Scheme", ETRI Journal, vol. 33, no. 6, **(2011)**.
[8] S. -u. Kang, H. J. Hwang and H. J. Kim, "Reversible Watermark Using an Accurate Predictor and Sorter Based on Payload Balancing", ETRI Journal, vol. 34, no. 3, **(2012)**, pp. 429-438.
[9] X. -C. Yuan and C. -M. Pun, "Geometrically Invariant Image Watermarking Based on Feature Extraction and Zernike Transform", IJSIA, vol. 6, no. 2, **(2012)**, pp. 161-166.
[10] C. -Y. Yang and C. -H. Lin "High-Quality and Robust Reversible Data Hiding by Coefficient Shifting Algorithm", ETRI Journal, vol. 34, no. 3, **(2012)**, pp. 429-438.
[11] G. RoslineNesaKumari, B. VijayaKumar, L. Sumalatha and V. V. Krishna, "Secure and Robust Digital Watermarking on Grey Level Images", IJAST, vol. 11, **(2009)**, pp. 1-8.
[12] S. Goyal and R. Gupta, "Optimization of Fidelity with Adaptive Genetic Watermarking Algorithm using Tournament Selection", IJAST, vol. 30, **(2011)**, pp. 55-66.
[13] Y. Lee and J. Kim, "Histogram Rotation-Based Image Watermarking with Reversibility", IJSIA, vol. 6, no. 2, **(2012)**, pp. 197-202.
[14] A. Leest, J. Haitsma and T. Kalker, "On Digital Cinema and Watermarking", Proceedings of the SPIE, vol. 5020, **(2001)**, pp. 526-535.
[15] H. Do, D. Choi, H. Choi and T. Kim, "Digital Video Watermarking Based on Histogram and Temporal Modulation and Robust to Camcorder Recording", Proceedings of 8th IEEE International Symposium on Signal Processing and Information Technology, **(2008)**, pp. 330-335.
[16] Y. Lee, S. Park, C. Kim and S. Lee, "Temporal Feature Modulation for Video Watermarking", IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 4, **(2009)** April, pp. 603-608.
[17] C. Chen, J. Ni and J. Huang, "Temporal Statistic Based Video Watermarking Scheme Robust against Geometric Attacks and Frame Dropping", Proceedings of IWDW 2009, LNCS 5703, **(2009)**, pp. 81-95.

[18] S. Xiang, H. Kim and J. Huang, "Invariant Image Watermarking Based on Statistical Features in the Low-Frequency Domain", IEEE Transactions on Circuit and Systems for Video Technology, vol. 18, no. 6, **(2008)**, pp. 777-790.
[19] Digital Cinema Initiatives, LLC, Digital Cinema System Specification Version 1.2, **(2008)**.
[20] S. Jeong, S. -C. Lim, H. Lee, J. Kim, J. S. Choi and H. Choi, "Highly Efficient Video Codec for Entertainment-Quality", ETRI Journal, vol. 33, no. 2, **(2011)**, pp. 145-154.
[21] X. -F. Li, N. Zhou and H. -S. Liu, "Joint Source/Channel Coding Based on Two-Dimensional Optimization for Scalable H.264/AVC Video", ETRI Journal, vol. 33, no. 2, **(2011)**, pp. 155-162.

Authors



Sang-Woo Lee received his BS, MS, and Ph. D degrees in electronics from Kyungpook National University, Daegu, Rep. of Korea, in 1999, 2001, and 2009 respectively. Since 2001, He has been a senior member of engineering staff in Electronics and Telecommunications Research Institute (ETRI). His research interests include information security based on cryptography and its applications.

Dae-Hee Seo received his Ph.D degree in the Graduate School of Computer Science from Soonchunhyang University, Korea. He is now a Senior Member of Engineering Staff of Electronics and Telecommunications Research Institute, Korea. Dr. Seo has published many research papers in international journals and conferences. Dr. Seo's research interests include Key Management, Network Management, Wireless Security, and Ubiquitous Computing. He is a member of the IARIA, SIG, AICIT, SERSC**.**

Byung-Gil Lee received a Ph. D degree in electrical engineering from the Kyungpook National University in 2003. In 2001, he joined the research member of ETRI in Korea and is currently a team leader of Convergence Security Research Team. His current research interests include wired/wireless network security and convergence security.