Robust Digital Watermarking for High-definition Video using Steerable Pyramid Transform, Two Dimensional Fast Fourier Transform and Ensemble Position-based Error Correcting

Xun Jin[1] and JongWeon Kim[2]
[1] Dept. of Copyright Protection, Sangmyung University, Seoul, Korea
[e-mail: jinxun@cclabs.kr]
[2] Dept. of Electronics Engineering, Sangmyung University, Seoul, Korea
[e-mail:jwkim@smu.ac.kr]
*Corresponding author: JongWeon Kim

Received August 1, 2017; revised October 26, 2017; accepted January 13, 2018; published July 31, 2018

Abstract

In this paper, we propose a robust blind watermarking scheme for high-definition video. In the embedding process, luminance component of each frame is transformed by 2-dimensional fast Fourier transform (2D FFT). A secret key is used to generate a matrix of random numbers for the security of watermark information. The matrix is transformed by inverse steerable pyramid transform (SPT). We embed the watermark into the low and mid-frequency of 2D FFT coefficients with the transformed matrix. In the extraction process, the 2D FFT coefficients of each frame and the transformed matrix are transformed by SPT respectively, to produce two oriented sub-bands. We extract the watermark from each frame by cross-correlating two oriented sub-bands. If a video is degraded by some attacks, the watermarks of frames contain some errors. Thus, we use an ensemble position-based error correcting algorithm to estimate the errors and correct them. The experimental results show that the proposed watermarking algorithm is imperceptible and moreover is robust against various attacks. After embedding 64 bits of watermark into each frame, the average peak signal-to-noise ratio between original frames and embedded frames is 45.7 dB.

Keywords: Video watermarking, steerable pyramid transform, fast fourier transform, error correcting

http://doi.org/10.3837/tiis.2018.07.024 **ISSN : 1976-7277**

1. Introduction

In recent years, the development of the Internet and multimedia transmission technology has led to the explosive growth of the digital video distribution services. The distribution of digital video contents is getting easier as increasing the network speed and the computing power. However, the increase of video distribution has led to an issue of illegal distribution. Copyright infringement has become a serious problem [1-7].

With the rapid increase of the storage capacity, high-definition videos, such as movies, dramas and entertainment programs, have attracted much attention for illegal distribution. In [8], authors proposed an image hashing method using block truncation coding. It extracts perceptual image features from local binary pattern. However, the method can’t extract the same feature if some parts of the image are cropped. Then the original copyrighted version can’t be found. In [9], a fragile image watermarking was proposed using pixel-wise recovery based on overlapping embedding strategy. Fragile watermarking methods are usually used for content authentication because the watermarks are destructible. However, in this work, we mainly focus on the robust watermarking which is generally applied to copyright protection. In [10], a reversible data hiding method was proposed to protect privacy for image content. However, it is not robust enough to survive various attacks such as noise, compression and cropping. Recently, discrete wavelet transform (DWT) based watermarking techniques attract much attention [11-13]. Generally, DWT based watermarking algorithms decompose the image and embed the watermark into the low level components.

In this paper, we propose a robust blind watermarking algorithm for high-definition video contents. The main contribution of this work is the robustness against geometrical and compression attacks. For the security of watermark information, a matrix of random numbers is generated using a secret key for later embedding. Two transform methods are used to increase robustness against common image processing distortions such as noise and compression. One is 2-dimensional fast Fourier transform (2D FFT) [14-17], which is used to transform the video frames from spatial domain to the frequency domain. The other one is steerable pyramid transform (SPT) [18-22]. It is aliasing-free and self-inverting. The matrix of random numbers is transformed by inverse SPT before embedding it into the 2D FFT coefficients. In the extraction process, SPT is applied to the transformed matrix and the 2D FFT coefficients to produce oriented sub-bands, respectively. The SPT transforms the image into several sub-bands of different orientations and scales. Thus, the transformed image is robust against geometrical attacks. The watermarks are embedded into low and mid-frequency of 2D FFT coefficients. Thus, they are robust against compression attacks. The watermark of each frame is extracted based on the cross-correlation between the two oriented sub-bands. If a video is degraded by some attacks, the watermarks of frames contain some errors. To decrease the bit error rate (BER) of the extracted watermark, we propose an ensemble position-based error correcting (EPbEC) algorithm to estimate the errors and automatically correct them. The EPbEC is motivated by error correcting output codes (ECOC) [23], [24] which is a method of recovering the classification errors and provides self-correcting abilities. The EPbEC transfers a multi-label classification problem to a binary classification problem, so that it corrects the bit errors in the extracted watermark. The experimental results show that the proposed watermarking algorithm is imperceptible and moreover is robust against various attacks such as noise, compression and cropping.

The remainder of this paper is organized as follows. In Section 2, we review the methodology of the SPT. In Section 3, we give a brief overview of 2D FFT. In Section 4, the EPbEC algorithm is presented. In Section 5, we present the watermark embedding and extraction process. Finally, we present the experimental results and discuss the evaluation in Section 6.

2. Steerable Pyramid Transform

The SPT is a discrete extension of the curvelet transform which captures curves instead of points. It is a method of multi-resolution image transform like contourlet transform. The construction of the contourlet transform is based on the directional filter banks. The SPT retains and improves some advantages and disadvantages of the orthonormal wavelet transform with steerable orientation decomposition. Its basis functions are localized in spatial-frequency and space. The aliasing is eliminated.

The SPT divides an image into a set of sub-bands which is localized at different orientations and scales [18-22]. It is computed using decimation and convolution operations. A high-pass and a low-pass filters are applied to the target image to produce high-pass and low-pass sub-bands. The low-pass sub-band is further divided into a low-pass and oriented sub-bands. The inverse transform reconstructs the original image with the high-pass, low-pass, and the oriented sub-bands. **Fig. 1** shows the structure of the SPT.

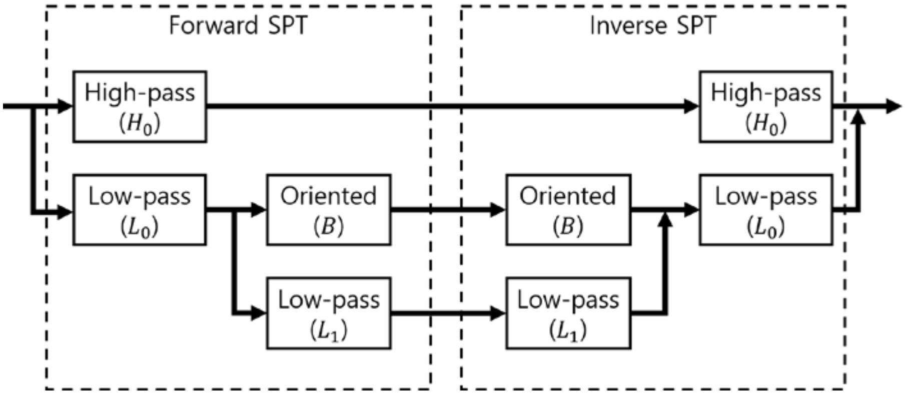


Fig. 1. Structure of SPT

To obtain the perfect reconstruction, some constraints for the filters are defined as follows. To eliminate the aliasing terms in the sub sampling operation, the narrowband low-pass filter $L_1(\omega)$ should be bandlimited.

$$|L_1(\omega)| = 0, |\omega| > \frac{\pi}{2} \tag{1}$$

where ω indicates the frequency vector of the Fourier domain.
In order to prevent amplitude distortion, the system transfer function should be one.

$$|H_0(\omega)|^2 + |L_0(\omega)|^2[|L_1(\omega)|^2 + |B(\omega)|^2] = 1 \tag{2}$$

where $H_0(\omega)$ indicates a non-oriented high-pass filter, $L_0(\omega)$ indicates a low-pass filter and $B(\omega)$ indicates oriented band-pass filter.
The low-pass branch should not be affected by the insertion of recursive portion.

$$\left|L_1\left(\frac{\omega}{2}\right)\right|^2 = \left|L_1\left(\frac{\omega}{2}\right)\right|^2 [|L_1(\omega)|^2 + |B(\omega)|^2] \tag{3}$$

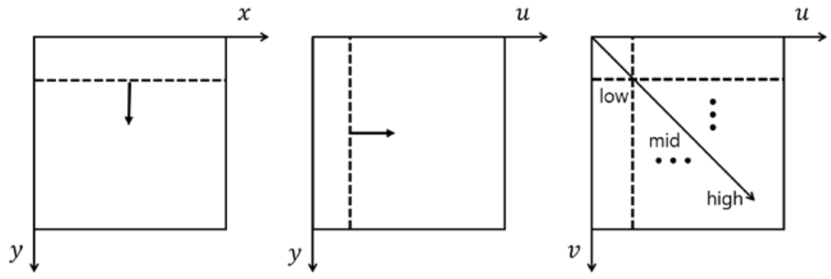
3. 2-dimensional Fast Fourier Transform

2-dimensional Fourier transform (2D FT) involves two steps of 1-dimensional Fourier transform (1D FT). First, a 2D FT of a digital image is achieved by transforming each row of the image, which means replacing each row with its 1D FT as shown in **Fig. 2 (a)**. This first step obtains an intermediary image in which the vertical axis indicates space and the horizontal axis indicates frequency. The second step is to transform each column of the intermediate image by 1D FT as shown in **Fig. 2 (b)**. **Fig. 2 (c)** shows the result of the 2D FT. The forward 2D FT is defined as (4) and the inverse 2D FT is defined as (5).

$$F(u,v) = \frac{1}{WD \cdot HT} \sum_{x=0}^{WD} \sum_{y=0}^{HT} f(x,y)\, e^{-j2\pi\left(\frac{ux}{WD}+\frac{vy}{HT}\right)} \tag{4}$$

$$f(x,y) = \sum_{u=0}^{WD} \sum_{v=0}^{HT} F(u,v)\, e^{j2\pi\left(\frac{ux}{WD}+\frac{vy}{HT}\right)} \tag{5}$$

where x and y denote the pixel coordinates of the image, u and v denote the coordinates of the transformed image. The WD and HT denote the width and height of the image. To calculate the transform more efficiently, we use the fast Fourier transform (FFT), which reduces the complexity of computing the transform from $O(\mathcal{N}^2)$ to $O(\mathcal{N}\log_2\mathcal{N})$, where $\mathcal{N}$ is the data size. To perform the 2D FFT instead of the 2D FT, the data size should be power of 2. Thus, zero-padding scheme is used to achieve the size.



(a) (b) (c)

Fig. 2. Schematic representation of 2D FT

4. Ensemble Position-based Error Correcting

The EPbEC algorithm uses binary positions of data to correct errors. It is inspired by ECOC, which is a powerful framework to deal with multi-label classification problems [23], [24]. It uses the positions of class labels as the objects for optimization. It can further correct classification errors after feature vector based classification.

First the EPbEC classifies a dataset with the size of N using a classifier algorithm. To obtain a classification ensemble with the size of M, it iteratively classifies the data and forms an

$M \times N$ ensemble matrix, EM, of class labels. We apply this concept to watermark extraction procedure. Therefore, we combine the watermarks extracted from M frames to obtain an ensemble matrix EM with size of $M \times N$, where N denotes the length of the watermark. The elements in EM are labelled with $c_l \in \{+1, -1\}, l = 1,2$. The positions of elements labelled with c_l in the mth row is defined as $LP^{m,l}$ where $m = 1,2, \dots, M$. We extract the labels at position $LP^{m,l}$ in all rows of EM to form a partial label set $PL^{m,l}$. The number of elements of c_l in the mth row is defined as $NL^{m,l}$, and the NLs of both labels compose a set NS^m with size of $M \times 2$. The NLs in NS are reordered in descending order to obtain RN. Then we assign temporary new labels nc to form a new partial label set NPL.

$$NPL_{j,k}^{m,l} = \{nc_i | RN_i^m = NS_a^m, c_a = PL_{j,k}^{m,l}\}, \quad i, a = 1,2, \quad j = 1,2, \dots, M, \quad k = 1,2, \dots, NL^{m,l} \tag{6}$$

The number of elements corresponding to a nc at a position k in NPL is defined as NP. The NPs of ncs at position k compose a set $SN^{m,l,k}$. Then we extract the index of the maximum value from the SN. Theoretically, because of the previous descending ordered RN, the index value should be equal to 1, if there are no or very few errors. Otherwise, the k is defined as ke and we correct the errors. We find the labels of other elements in EM at the position where the error occurred as follows:

$$OC_h^{m,l,ke} = \{r | c_r = EM_{h,p}, p = LP_{ke}^{m,l}\}, \quad h = 1,2, \dots, M, r = 1,2 \tag{7}$$

We define a 3-dimensional matrix of binary position, BP, for each label at each position of EM in advance as (8). The similarities of the entire elements to the OC are calculated using the BP as follows:

$$BP_{m,b,l} = \begin{cases} 1, & \text{if } EM_{m,n} = c_l \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

$$S_b^{m,l,ke} = \sum_{h=1}^{M} BP_{h,b,OC_h^{m,l,ke}}, \quad b = 1,2, \dots, N \tag{9}$$

We ignore the elements which have low similarities to select candidates CD. The binary positions corresponding to CD are extracted and accumulated to produce PC which can imply the probability of correct labels. The original label of EM at position p in the mth row will be replaced by the c_{lc}, where the lc denotes the index of the maximum value in PC. In the meantime, the binary position of lth index at position p in the mth row is set to be 0 and that of lcth index is set to be 1. After processing the whole rows and labels, we can obtain a corrected ensemble. The final watermark is the average of the corrected ensemble.

5. Proposed Watermarking Scheme

The proposed watermarking scheme is mainly composed of SPT, 2D FFT, correlation, and EPbEC algorithms. **Fig. 3** shows the embedding process. First, each frame of a video is converted to the NTSC color space to obtain the luminance component. The luminance component is then transformed to the frequency domain using 2D FFT. A secret key is used to generate a $n/2 \times n/2$ matrix of random numbers. It is expanded to two $n \times n$ matrices for high-pass and oriented sub-bands. Each element of the $n/2 \times n/2$ matrix is expanded to 2×2 elements of the $n \times n$ matrix. The inverse SPT transforms the three matrices using one orientation and produces a $n \times n$ transformed matrix TM. The transformed matrix is embedded into the low and mid-frequency of 2D FFT coefficients based on the watermark w as (10). The embedding region is shown as **Fig. 4**. The region is divided into N blocks. The size of each block BL is $n \times n$.

$$BL' = \begin{cases} BL_{s,t} + \alpha \times TM, & if\ w_{s,t} = +1 \\ BL_{s,t} - \alpha \times TM, & otherwise \end{cases} \tag{10}$$

where α denotes the embedding strength and BL' denotes the marked coefficient block.

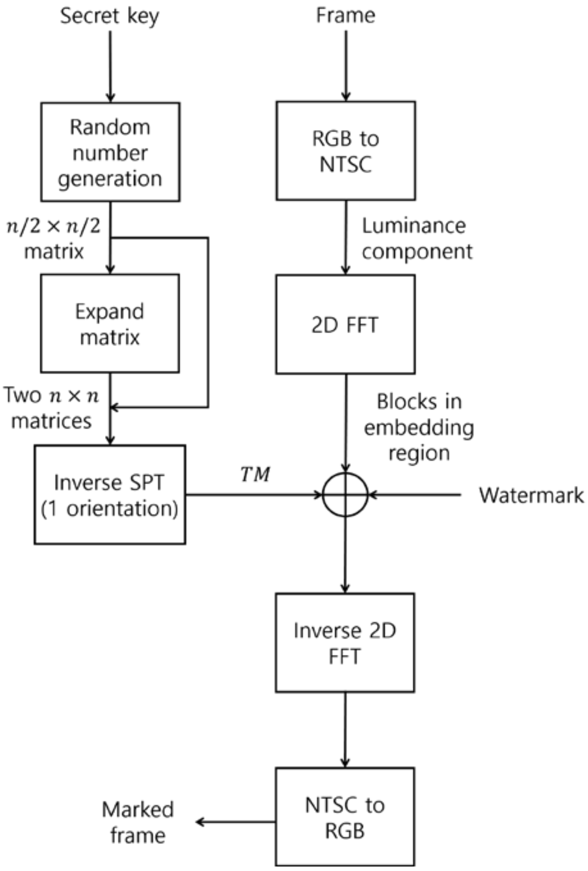


Fig. 3. Watermark embedding process

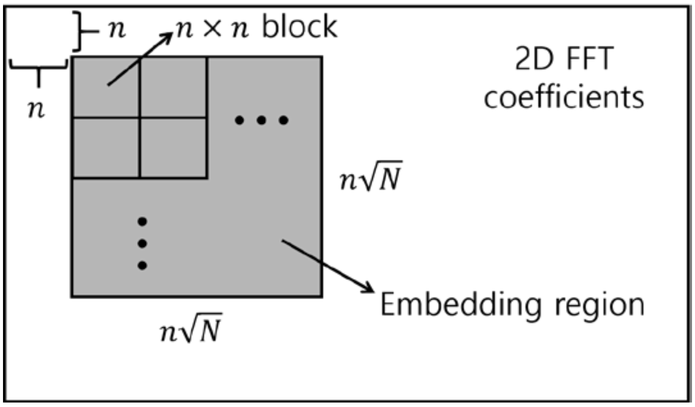


Fig. 4. Embedding region

In the procedure of watermark extraction, the TM is generated in the same way as in the embedding procedure. Each block of the embedding region of the 2D FFT coefficients and the TM are transformed by forward SPT with one orientation. **Fig. 5** shows the extraction process. The cross-correlation coefficients CC between the two oriented sub-bands of 2D FFT coefficient block and the TM are computed. The cross-correlation results of the components obtained by orientation decomposition, are more effective than those of the images in the spatial domain. Therefore, the forward SPT is used in the extracting process, whereas the inverse SPT is used in the embedding process. The watermark of each frame is extracted based on the maximum and the minimum correlation peaks as follows:

$$w_{s,t}' = \begin{cases} +1, & |\max\{CC_{s,t}\}| > |\min\{CC_{s,t}\}| \\ -1, & \text{otherwise} \end{cases} \tag{11}$$

where $w_{s,t}'$ is the extracted watermark of a frame.

The EPbEC combines the watermarks of M frames and corrects the errors to produce final watermark.

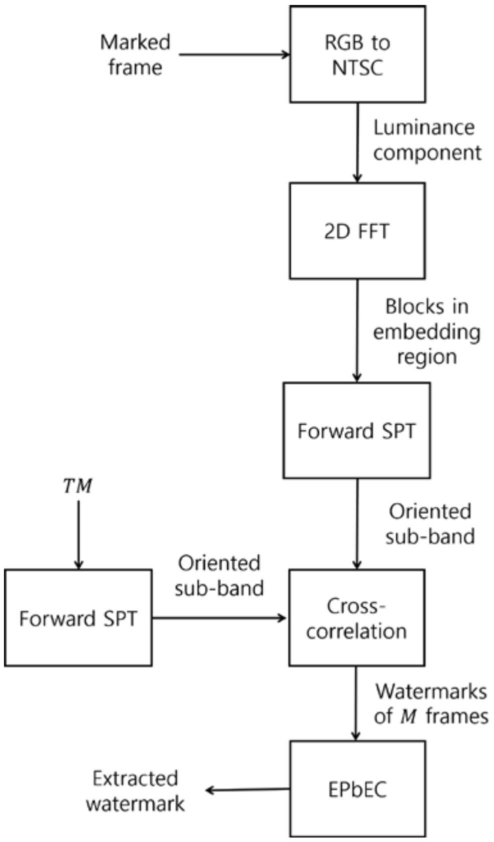


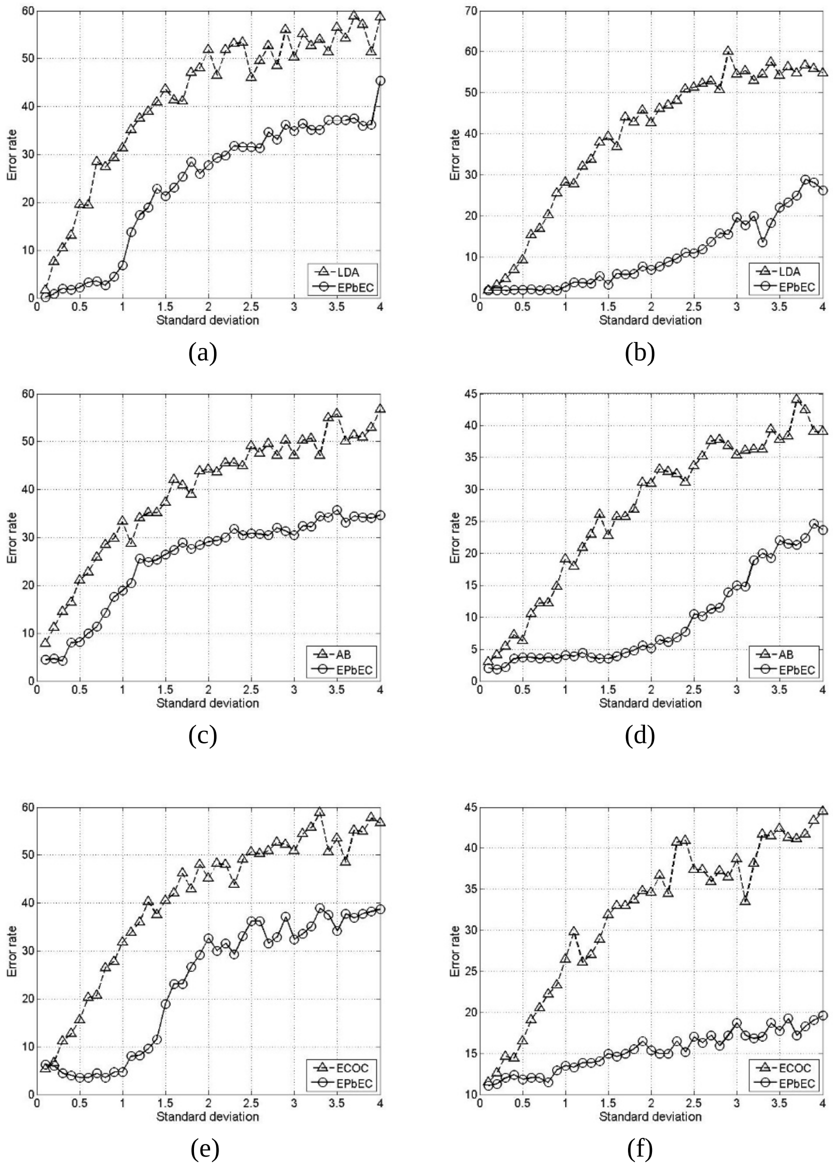
Fig. 5. Watermark extracting process

6. Performance Evaluation

In this section, first we present the performances of the EPbEC algorithm and second we discuss the evaluation of the proposed watermarking scheme. To carry out the comparisons, three classification algorithms were used to evaluate the performance of EPbEC, and they were linear discriminant analysis (LDA) [25], [26], adaboost (AB) of decision tree [27], [28] and ECOC. We selected Fisher's iris dataset and wine dataset from the UC Irvine machine learning repository as experimental data. The evaluation was investigated in terms of minimum classification error rates (MCER) defined as follows:

$$MCER = \min\left\{\frac{E_m}{N} \times 100\%\right\} \tag{12}$$

where E_m indicates the number of sample cases incorrectly classified in mth classification. The number of samples N for iris dataset is 150 and for wine dataset is 178. We set the ensemble size $M = 100$ and selected 70% of each dataset for training. The test data were distorted by adding random noises with zero mean value and σ standard deviation. **Fig. 6** shows MCERs of the three classification algorithms and those of further corrected results using EPbEC. We increased the σ from 0.1 to 4 to strongly distort the test data. **Fig. 6 (a), (c)** and **(e)** show the experimental results of iris dataset. They show the error rates were corrected about 10%. **Fig. 6 (b), (d)** and **(f)** show those of wine dataset. About 20% of error rates were corrected by EPbEC.



(a) (b)

(c) (d)

(e) (f)

Fig. 6. MCERs of three classification algorithms and EPbEC

We evaluated the proposed watermarking scheme using 720×1280 high-definition videos ‘mobcal’ (sample 1), ‘shields’ (sample 2), ‘stockholm’ (sample 3) and ‘pedestrian_area’ (sample 4). The framerates of the four samples are 50, 50, 59 and 25 fps. Their bitrates are 17529, 15795, 15698 and 3953 kbps. **Table 1** shows metadata of the samples.

Table 1. Metadata of video samples

Properties	**Sample1**	**Sample2**	**Sample3**	**Sample4**
Resolution (pixel)	720 *1280	720 *1280	720 *1280	720 *1280
Duration (second)	10	10	10	15
Framerate (fps)	50	50	59	25
Bitrate (kbps)	17529	15795	15698	3953
Number of frames	500	500	590	375

A random number sequence was defined as a watermark. The length of the watermark N was 64 and the size of the transformed matrix n was 32. We performed experiments on an IBM compatible computer with an i7 3.6 GHz CPU, a 16 GB random-access memory (RAM) and Windows 8 64 bit OS. The average watermark embedding and extraction time were 0.12 and 0.16s for each frame. The quality of marked frame was measured with peak signal-to-noise ratio (PSNR). The PSNR is defined as:

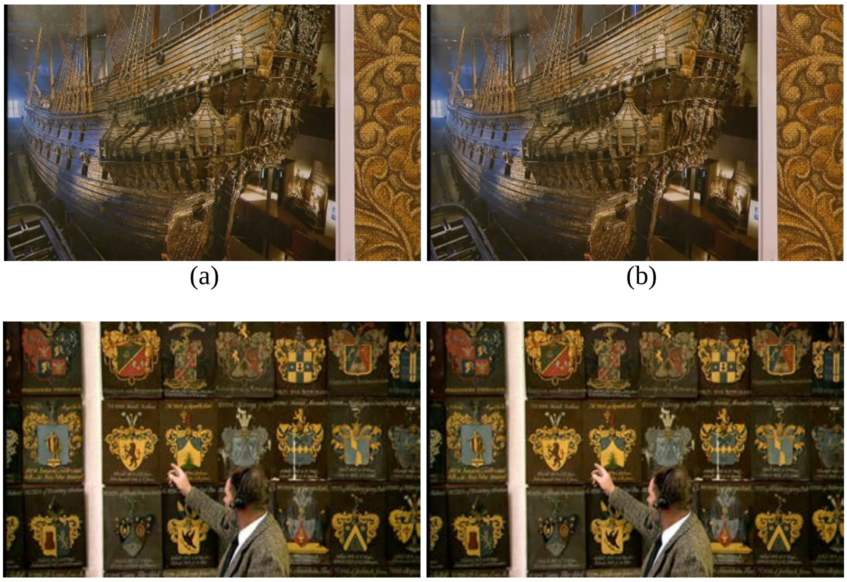
$$PSNR = 10 log_{10} \frac{255^2}{MSE} \tag{13}$$

where MSE (mean square error) is defined as:

$$MSE = \frac{1}{NPX} \sum_{px=1}^{NPX} \left(FR_i - \widehat{FR}_i\right)^2 \tag{14}$$

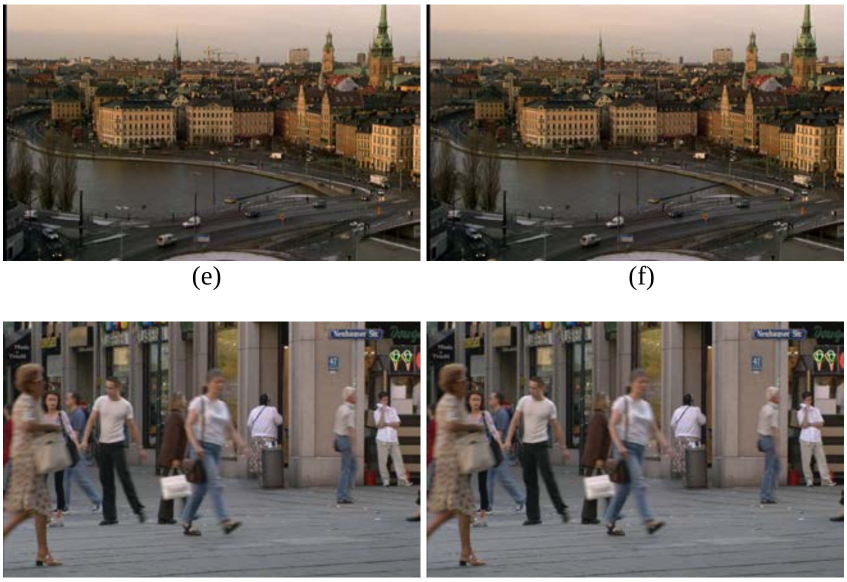
where FR and $\widehat{FR}$ denote the original and marked frames, and NPX denotes the number of pixels in a frame.

The embedding strength α was set to be 20. **Fig. 7 (a), (c), (e)** and **(g)** show the original frames and **Fig. 7 (b), (d), (f)** and **(h)** show the marked frames of the samples. The average PSNR and structure similarity index between the original and marked frames are 45.7 dB and 0.98. They show the proposed watermarking algorithm is imperceptible.



(a) (b)

(c) (d)



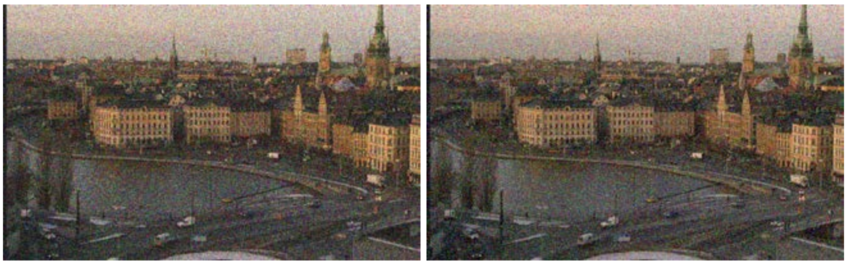
(e) (f)

(g) (h)

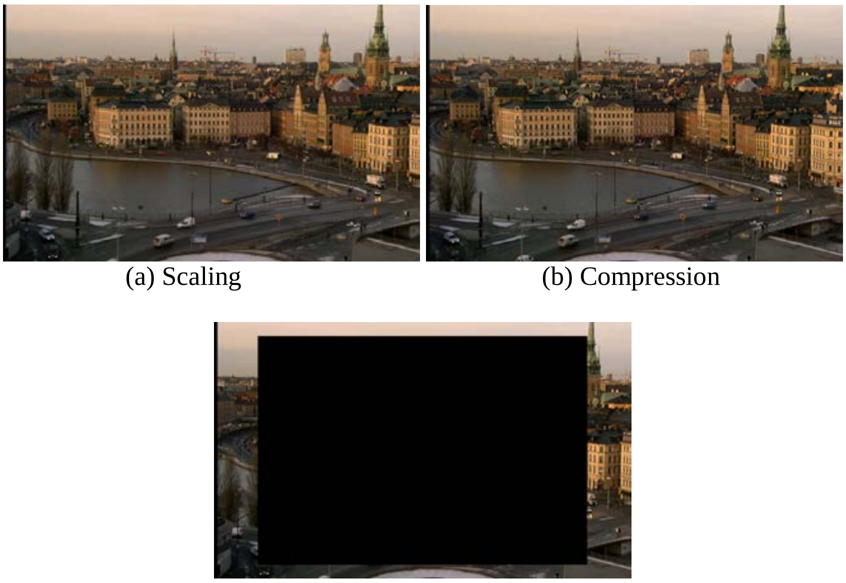
Fig. 7. Original and marked frames

Fig. 8 (a) shows a frame attacked by adding Gaussian noises with mean of 0 and variance of 0.1. **Fig. 8 (b)** shows a frame distorted by adding salt & pepper noises with noise density of 0.2. **Fig. 8(c)** shows a frame scaled by 50%. Two parameters were used to manipulate the compression rate. One was framerate and the other was bitrate. **Fig. 8(d)** shows a frame compressed by framerate of 15 fps and bitrate of 3000 kbps. **Fig. 8 (e)** shows a cropped frame with only 30% of the frame is remained. The BER between the extracted watermark of the cropped frame and the original watermark was 0%, even though they were seriously damaged. The BER is defined as follows:

$$BER = \left(1 - \frac{number\ of\ correct\ bits}{number\ of\ total\ bits}\right) \times 100\% \tag{15}$$



(a) Gaussian Noise (b) Salt & Pepper Noise



(a) Scaling

(b) Compression

(e) Cropping

Fig. 8. Attacked Frames

To evaluate the robustness of the proposed watermarking algorithm, various attacks were applied to the marked videos. A blind and a non-blind watermarking algorithm were proposed using SPT in [18] and [19]. We compared the proposed method (SF) with two recent methods. First one used the SPT and the discrete cosine transform (DCT) [18] (SD). Second one used the DWT and an alpha blending technique [13] (DA). **Table 2** shows the BERs of SF and SD according to the attacks. The mean of the Gaussian noise was set to be 0 and the variance ranged from 0.02 to 0.1. The noise density of salt & pepper noise ranged from 0.12 to 0.2. After adding the Gaussian noises and salt & pepper noises, the BERs of the SF for the four samples were less than 8%, whereas most BERs of the SD were greater than 10%, which means the proposed algorithm is robust against noise attack. If the frame size of videos for watermark extraction is not equal to 720×1280, the video will be resized to 720×1280 before extraction. After scaling half size of the frame, the BERs of the SF were less than 15%, whereas half of the BERs of the SD were greater than 20%. The BERs of the SF for sample 2 were greater than 10%, when it was compressed with bitrate of 3000 kbps, whereas those of the other three samples were less than 2%. The cropping attack removed the areas located at the center of the frame. Even if we removed 90% of the frame area, the BERs of the SF were less than 10% showing good performance. **Table 3** shows the BERs of DA according to the attacks. All of the BERs are greater than 30% under the noise attacks. Most of the BERs of the DA are greater than 10% showing lack of robustness.

Table 2. BERs of SF and SD According to Attacks

Attacks	Sample1		Sample2		Sample3		Sample4	
	SF	SD	SF	SD	SF	SD	SF	SD
Gaussian noise (0.02)	0	25.43	0	20.34	0	14.86	0	3.34
Gaussian noise (0.04)	0	27.36	0	21.88	0	19.7	0	7.17

Gaussian noise (0.06)	0	30.82	1.56	26.1	0	24.39	0	10.14
Gaussian noise (0.08)	1.56	31.71	4.69	26.48	1.56	25.2	3.13	14.37
Gaussian noise (0.1)	3.13	33.03	6.25	28.09	1.56	28.22	4.69	15.75
Salt & pepper noise (0.12)	0	29.01	0	23.74	0	21.35	0	7.78
Salt & pepper noise (0.14)	0	29.96	0	23.87	0	21.66	0	8.36
Salt & pepper noise (0.16)	0	29.99	1.56	24.91	0	22.82	0	9.07
Salt & pepper noise (0.18)	1.56	31.5	3.13	25.46	0	24.14	1.56	12.41
Salt & pepper noise (0.2)	3.13	31.71	7.81	25.8	1.56	26.04	3.13	13.36
Scaling (0.5)	12.5	24.94	14.06	20.47	10.94	9.96	4.69	5.2
Scaling (0.75)	0	22.12	0	16.05	0	6.07	0	0.18
Compression (30, 5000)	0	37.25	6.25	15.47	0	7.94	0	2.11
Compression (15, 5000)	0	31.86	4.96	18.11	0	11.86	0	0
Compression (30, 3000)	0	41.85	28.13	20.22	0	15.35	0	8.98
Compression (15, 3000)	1.56	39.28	12.5	22.76	0	16.73	0	1.29
Cropping (50%)	0	41.48	0	39.98	0	39.09	0	38.88
Cropping (70%)	1.56	41.82	0	40.44	0	39.74	0	39.06
Cropping (90%)	9.38	40.63	9.38	40.63	6.25	40.63	7.81	40.63
Rotation (10°)	1.56	17.62	1.56	15.81	3.13	7.63	4.69	1.69
Rotation (20°)	0	20.31	1.56	19.36	1.56	10.51	6.25	7.29
Rotation (30°)	1.56	23.87	1.56	21.69	1.56	14.19	3.13	9.96

Rotation (40°)	1.56	32.02	1.56	25.8	1.56	21.26	3.13	10.88

Table 3. BERs of DA According to Attacks

Attacks	Sample1	Sample2	Sample3	Sample4
Gaussian noise (0.02)	34.95	38.89	34.34	35.6
Gaussian noise (0.04)	41.48	44.14	40.04	40.49
Gaussian noise (0.06)	44.55	46.2	43.19	42.83
Gaussian noise (0.08)	46.22	47.24	45.15	44.27
Gaussian noise (0.1)	47.2	47.83	46.39	45.22
Salt & pepper noise (0.12)	45.98	46.54	44.5	43.11
Salt & pepper noise (0.14)	46.77	47.16	45.61	44.02
Salt & pepper noise (0.16)	47.35	47.61	46.43	44.7
Salt & pepper noise (0.18)	47.75	47.94	47.04	45.28
Salt & pepper noise (0.2)	48.07	48.19	47.52	45.77
Scaling (0.5)	19.41	22.58	25.06	15.08
Scaling (0.75)	12.65	14.36	19.18	9.39
Cropping (50%)	28.66	29.73	31.7	28.77
Cropping (70%)	37.58	38.02	39.5	37.67
Cropping (90%)	46.55	46.63	47.23	46.49
Rotation (10°)	13.58	15.4	18.17	12.08
Rotation (20°)	16.81	18.25	20.85	15.28

Rotation (30°)	19.02	20.38	22.8	17.58
Rotation (40°)	20.73	22.14	24.44	19.44

7. Conclusion

In this paper, we proposed a robust blind video watermarking scheme with SPT, 2D FFT, and EPbEC. The 2D FFT was used to transform the luminance component of each frame to the frequency domain. Inverse SPT was performed to a matrix of random numbers to produce a transformed matrix. Watermark was embedded into the low and mid-frequency of 2D FFT coefficients using the transformed matrix. In the watermark extraction process, the transformed matrix and the 2D FFT coefficients were transformed by SPT to produce oriented sub-bands, respectively. Cross-correlation was used to extract the watermark from each frame with the oriented sub-bands. The EPbEC algorithm was used to estimate the errors of watermarks and automatically correct them. The experimental results show the PSNR is greater than 45. After adding strong noises, the BER of watermark is still less than 8%. Even if 90% of the frame area is removed, the BER is still less than 10%.

Acknowledgments

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2017.

References

[1] J. Kim, N. Kim, D. Lee, S. Park and S. Lee, “Watermarking two dimensional data object identifier for authenticated distribution of digital multimedia contents,” *Signal Processing: Image Communication*, vol. 25, no. 8, pp. 559-576, September, 2010. Article (CrossRef Link)

[2] G. RoslineNesaKumari, B. VijayaKumar, L. Sumalatha and V. V. Krishna, “Secure and Robust Digital Watermarking on Grey Level Images,” *International Journal of Advanced Science and Technology*, vol. 11, pp. 1-8, October, 2009. Article (CrossRef Link)

[3] D. Li and J. Kim, “Secure Image Forensic Marking Algorithm using 2D Barcode and Off-axis Hologram in DWT-DFRNT Domain,” *Applied Mathematics & Information Sciences*, vol. 6, no. 2, pp. 513-520, January, 2012. Article (CrossRef Link)

[4] J. Nah, J. Kim and J. Kim, “Video Forensic Marking Algorithm Using Peak Position Modulation,” *Applied Mathematics & Information Sciences*, vol. 7, no. 6, pp. 2391-2396, November, 2013. Article (CrossRef Link)

[5] S. Kim, S. H. Lee, and Y. M. Ro, “Rotation and flipping robust region binary patterns for video copy detection,” *Journal of Visual Communication and Image Representation*, vol. 25, no.2, pp. 373-383, February, 2014. Article (CrossRef Link)

[6] M. M. Sathik and S. S. Sujatha, “Wavelet Based Blind Technique by Espousing Hankel Matrix for Robust Watermarking,” *International Journal of Advanced Science and Technology*, vol. 26, pp. 57-72, January, 2011. Article (CrossRef Link)

[7] X. Jin and J. Kim, “Robust Digital Watermarking for High-Definition Video using Steerable Pyramid Transform and Fast Fourier transformation,” in *Proc. of Proceedings of the 4th*

International Conference on Digital Contents and Applications, Jeju, Korea, vol. 120, December 16-19, 2015. Article (CrossRef Link)

[8] C. Qin, X. Chen, D. Ye, J. Wang and X. Sun, "A Novel Image Hashing Scheme with Perceptual Robustness Using Block Truncation Coding," *Information Sciences*, vol. 361-362, no. 20, pp. 84-99, September, 2016. Article (CrossRef Link)

[9] C. Qin, P. Ji, X. Zhang, J. Dong and J. Wang, "Fragile image watermarking with pixel-wise recovery based on overlapping embedding strategy," *Signal Processing*, vol. 138, pp. 280-293, September, 2017. Article (CrossRef Link)

[10] C. Qin and X. Zhang, "Effective Reversible Data Hiding in Encrypted Image with Privacy Protection for Image Content," *Journal of Visual Communication and Image Representation*, vol. 31, pp. 154-164, August, 2015. Article (CrossRef Link)

[11] R. B. Narute and S. R. Patil, "Invisible Video Watermarking for Secure Transmission Based on DWT and PCA Mechanism," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 6, no. 6, pp. 12128-12135, June, 2017. Article (CrossRef Link)

[12] R. Srivastava, "Dwt based Invisible Watermarking on Images," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 3, no. 1, pp. 70-74, January, 2017. Article (CrossRef Link)

[13] N. Asha and P. Bhagya, "Video Watermarking using DWT and Alpha Blending Technique," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 6, no. 5, pp. 348-352, May, 2017. Article (CrossRef Link)

[14] G. Gupta and H. Aggarwal, "Digital image Watermarking using Two Dimensional Discrete Wavelet Transform, Discrete Cosine Transform and Fast Fourier Transform," *International Journal of Recent Trends in Engineering*, vol. 1, no. 1, pp. 616-618, May, 2009. Article (CrossRef Link)

[15] C. V. Loan, "Computational Frameworks for the Fast Fourier Transform," *Society for Industrial and Applied Mathematics*, USA, 1992. Article (CrossRef Link)

[16] D. N. Rockmore, "The FFT: an algorithm the whole family can use," *Computing in Science & Engineering*, vol. 2, no. 1, pp. 60-64, January, 2000. Article (CrossRef Link)

[17] I. H. Sarker, M. I. Khan, K. Deb and F. Faruque, "FFT-Based Audio Watermarking Method with a Gray Image for Copyright Protection," *International Journal of Advanced Science and Technology*, vol. 47, pp. 65-76, October, 2012. Article (CrossRef Link)

[18] A. Hossaini, M. Aroussi, K. Jamali, S. Mbarki, and M. Wahbi, "A New Robust Blind Watermarking Scheme Based on Steerable pyramid and DCT using Pearson product moment correlation," *Journal of Computers*, vol. 9, no. 10, pp. 2315-2327, October, 2014. Article (CrossRef Link)

[19] F. Drira, F. Denis, and A. Baskurt, "Image watermarking technique based on the steerable pyramid transform," in *Proc. of 2004 SPIE*, pp. 165-176, November, 2004. Article (CrossRef Link)

[20] A. Karasaridis and E. P. Simoncelli, "A Filter Design Technique for Steerable Pyramid Image Transforms," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, GA, May 9-9, 1996. Article (CrossRef Link)

[21] W. T. Freeman and E. H. Adelson, "The Design and Use of Steerable Filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891-906, September, 1991. Article (CrossRef Link)

[22] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: a flexible architecture for multi-scale derivative computation," in *Proc. of International Conference on Image Processing*, Washington, DC, October 23-26, 1995. Article (CrossRef Link)

[23] T. Kajdanowicz and P. Kazienko, "Multi-label classification using error correcting output codes," *Int. J. Appl. Math. Comput. Sci.*, vol. 22, no. 4, pp. 829–840, December, 2012. Article (CrossRef Link)

[24] S. Amornsamankul, J. Promrak, and P. Kraipeerapun, "Solving Multiclass Classification Problems using Combining Complementary Neural Networks and Error-Correcting Output Codes," *International Journal of Mathematics and Computers in Simulation*, vol. 5, no. 3, pp. 266-273, January, 2011. Article (CrossRef Link)

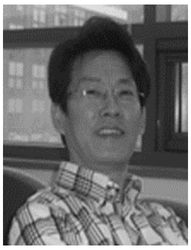
[25] T. Li, S. Zhu and M. Ogihara, "Using discriminant analysis for multi-class classification: an experimental investigation," *Knowledge and Information Systems*, vol. 10, no. 4, pp. 453-472, November, 2006. Article (CrossRef Link)

[26] C. Moulin, C. Largeron, C. Ducottet, M. Gery and C. Barat, "Fisher Linear Discriminant Analysis for text-image combination in multimedia information retrieval," *Pattern Recognition*, vol. 47, no. 1, pp. 260-269, January, 2014. Article (CrossRef Link)

[27] H. Fleyeh and E. Davami, "Multiclass Adaboost Based on an Ensemble of Binary AdaBoosts," *American Journal of Intelligent Systems*, vol. 3, no, 2, pp. 57-70, September, 2013. Article (CrossRef Link)

[28] R. Appel, T. Fuchs, P. Dollar and P. Perona, "Quickly Boosting Decision Trees - Pruning Underachieving Features Early," in *Proc. of Proceedings of the 30th International Conference on Machine Learning*, Atlanta, USA, vol. 28, pp. 594-602, June 16-21, 2013. Article (CrossRef Link)

Xun Jin, received his Ph.D. degree in Copyright Protection from Sangmyung University in Korea (2018). He is currently a product manager at Hangzhou Leaper Technology Co., Ltd. His research interests are digital image/video watermarking, multimedia forensics, pattern recognition, digital signal processing, and information security



JongWeon Kim, received the Ph.D. degree from University of Seoul, major in signal processing in 1995. He is currently a professor of Dept. of Electronics Engineering and Director of Creative Content Labs at Sangmyung University in Korea. He has a lot of practical experiences in the digital signal processing and copyright protection technology in the institutional, the industrial, and academic environments. His research interests are in the areas of copyright protection technology, digital rights management, digital watermarking, and digital forensic marking.