DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069

Improvement of trade-off between compression tolerance and perceptibility of video watermark for broadcasting

Keizo Kato, Kensuke Kuraki, and Ryuta Tanaka
FUJITSU LABORATOLIES LTD.
Kawasaki, Japan

Abstract
Fujitsu is working on the use of video watermarking for digital marketing. In the case of advertising services in Japan, TV viewers can automatically access an E-commerce site synchronized with a home shopping network program and easily order a commodity of their choice by detecting the watermark embedded in the video they are watching with their smart device application. Since watermark signals are generally deteriorated by video compression, the watermark strength needs to be adjusted. In the TV broadcasting, the degree of deterioration is different depending on the form of broadcasting (e.g. digital terrestrial or satellite) because of the associated difference of bitrate and compression format. The complexity of adjusting the watermark strength to each broadcasting form for each video becomes a problem in the real operation while if the strength determined to overcome the largest deterioration is used for all other cases as well, the influence on the image quality of video compressed with a low compression rate may become greater than necessary. To reduce the above-mentioned inconvenience in practical use, we have developed a method that standardizes the strength for more applications than before by improving the trade-off between video compression tolerance and influence on the video quality.

Context
The digital watermark is widely used in content protection; however, its application has recently spread to digital marketing. For instance, Digimarc Discover uses audio/image watermarking [1] and KANTAR MEDIA's Sync now uses audio watermarking [2]. Authors have been developing technology distributing information or directing customers to a URL related to the content by embedding watermark information under the video, and allowing viewers to detect it with smart device cameras (Figure 1). The technology is in practical use in the home shopping network program in Japan [3]. Viewers can access a purchasing site simply by pointing their smart device camera at the screen when they see a commodity they like. It reduces the time required for viewers to telephone or search for the commodity, and allows them to use the program more comfortably. There are several ways to easily access related information from the video using video recognition [4], audio watermarks, etc., as shown in Table1. Video recognition does not require any modifications to the contents, having no impact on video quality. On the other hand, however, it requires the preliminary construction of a database to be used for recognition. Audio watermarking allows viewers to retrieve information only by activating an application; however, the usage is restricted, for example it cannot be used for some applications, such as digital signage, which involves no audio speakers. The best method differs, depending on the scene and usage. To widely develop services solely relying on video watermarking, technological improvement to use it under various environments is necessary.

Table 1 Technologies to connect video and ICT and their features

	video recognition	Audio watermark	Video watermark
Influence on content	None	Imperceptible	Imperceptible
Identification of the same contents	Not possible	Possible by embedding unique watermark	Possible by embedding unique watermark
Construct DB for recognition	Necessary	Unnecessary	Unnecessary
Necessary equipment	Monitor	Monitor and speaker	Monitor
User's task	Holding smart device	Launch app	Holding smart device
Influence of ambient noise	None	Possible	None

OBJECTIVE
When applied to broadcasting content, it is necessary to prevent the watermark from impacting on image quality so as not to hinder viewing. Because of compression, slight variations of pixel values like watermarking tend to be deleted which leads to a deteriorated signal. As a result, the watermark detection speed may become

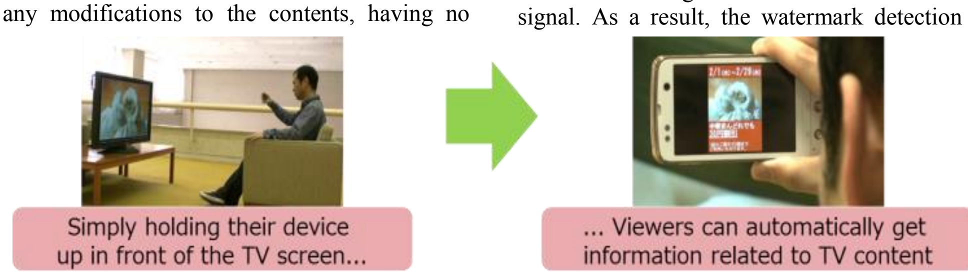


FIGURE 1 FUJITSU'S TECHNOLOGY TO CONNECT VIDEO AND ICT

DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069

slower. Therefore, it is necessary to adjust the strength of the watermark to detect it at high speed even after the video is delivered. However, when one content is delivered in two or more broadcasting forms such as the digital terrestrial or satellite ones, the deterioration degree is different depending on the broadcasting forms because of the difference of bitrate and compression format. The problem in typical workflows is that adjusting the watermark strength to make it appropriate for each broadcasting form requires an additional complexity for the overall process. On the other hand, if the strength is standardized according to the greatest deterioration degrees, the impact on the image quality of video compressed with a low compression rate may become greater than necessary. The study is aimed to reduce the complexity in the strength adjustment in practical use by improving the trade-off between the signal compression tolerance and the impact on perceptibility to achieve more flexible signal standardization.

Video Watermark

Related Work

We assume capture is done from the usual TV viewing distance using smart device camera. In practical use, it may be required to distribute content with digital watermarks through multiple channels or to transcode such content. For this reason, the method must satisfy the following requirements:

(i) Watermarks must be tolerant to changes in the dynamic range of images resulting from a retake of the screen with a camera and to spatial step-out.

(ii)Embedding schemes must be independent of any compression format.

The methods of embedding digital watermarks into video are roughly divided into two types according to whether watermarks are embedded into an uncompressed or compressed domain.

The approaches of embedding watermarks into bit streams in a compressed domain ([5][6][7]) do not involve the problem of watermarks becoming deteriorated by lossy compression, the problem specific to watermarks embedded in an uncompressed domain, if the video is not recompressed. These methods, however, embed watermarks in a manner dependent on a specific compression format, and therefore cannot be used for applications that require the video to be distributed in different compression formats. They also have a problem of difficulty in using temporal characteristics because it is difficult to manipulate images without restraint except in I frame.

The methods of embedding watermarks into an uncompressed domain include approaches that manipulate the values of pixels in the video or coefficients in its orthogonally transformed (such as DFT, DCT, or DWT) space. These methods, if based on the quantization of the pixel values or the coefficients of transformed domain [9] [10] are not adaptable to changes in dynamic ranges resulting from shooting of the monitor with a camera. The methods that use the spatial-direction characteristics of the entire screen [10] are vulnerable to spatial alignment errors caused, for example, by hand-induced camera shake. The method developed by Haitsma et al. [11] embeds digital watermarks by changing the average pixel value of the video. Assume that a 1-bit watermark signal is a pseudorandom number with a length of N, where $w(n) \in \{-1,1\}$. This method is based on the fundamental principle that it changes the average pixel value by adding the value 1 to all pixel values if $w(n) = 1$ in a given time frame t, or adding the value -1 if $w(n) = -1$. It assigns a weight to each embedding position to control the impact on perceptibility at the time of actual embedding. It detects bits by correlating the change in average pixel value with $w(n)$. The method considers only the time-series changes in the average pixel value of the video, meaning that it is tolerant to changes in the dynamic range and spatial step-out resulting from the shooting of the monitor from a distance with a hand-held camera.

Chrominance Video watermarking in time domain

This section describes our method of digital video watermarking. Our method may be considered to be an improved version of the method by Haitsma et al.—a version that makes watermarks less perceptible by making use of the fact that the human eye is insensitive to changes that occur slowly. While Haitsma et al. switched the value to be added between 1 and -1 at regular frame intervals, our watermarks are embedded by slowly changing the average pixel value of the video in the time domain. This is changed by increasing/decreasing the area in which the pixel value is varied every time (Figure 2). Data is transmitted by expressing binary information combining two types of wave with different patterns (Figure 3).

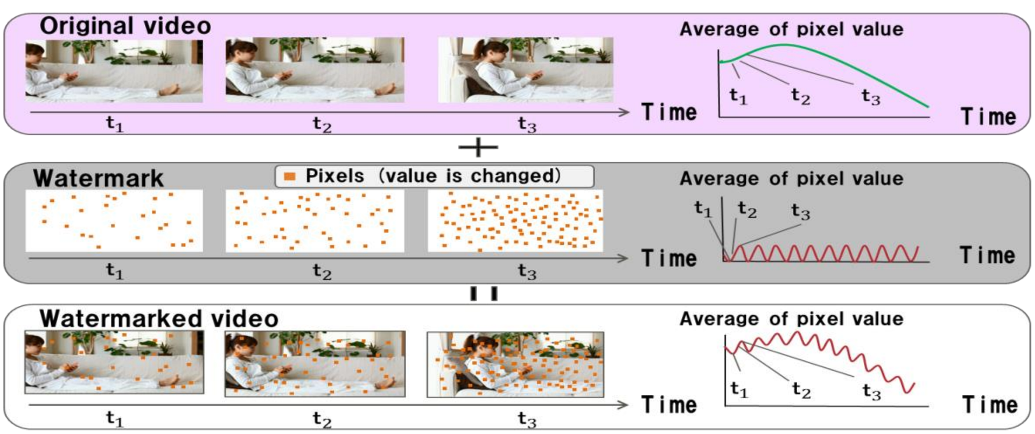


FIGURE 2 WATERMARK USING TEMPORAL VARIATION OF AVERAGE OF PIXEL VALUE

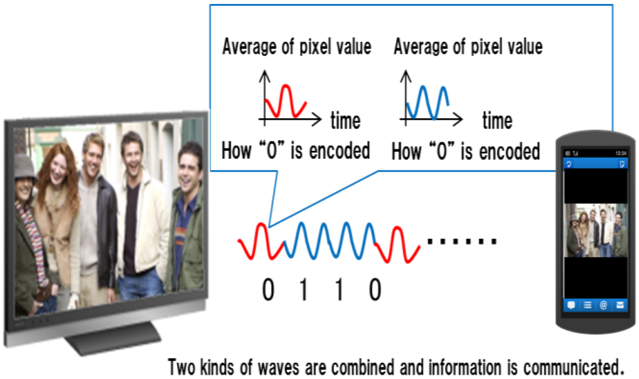


FIGURE 3 COMMUNICATION USING TWO KINDS OF WAVEFORM

Watermarks are embedded so as to be scattered as uniformly as possible on the entire screen with a 1 pixel precision, like random noise. Doing so avoids the steep local variation easily perceived by the human eye. In addition, watermarks are embedded in the chrominance component instead of the luminance component because the former is more difficult to perceive.

This method limits the bit rate due to the temporal resolution of motion video, meaning that it is difficult to directly embed information such as an URL in a short time as a watermark. As a solution to this, we embedded an ID for each video as a watermark signal to distribute information that corresponds to each detected ID, information preliminarily associated with each ID using a database. We thus implemented the service. Figure 4 outlines the system we developed.

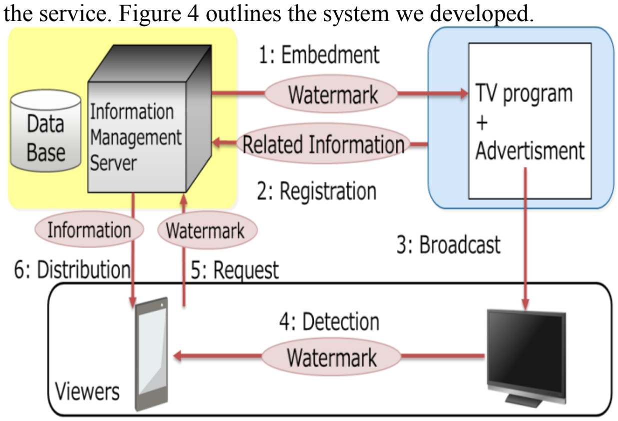


FIGURE 4 OVER VIEW OF TV-WEB CONNECTING SERVICE

Enhancement of compression tolerance

Deteriorated signals may be considered as signals that have noise. As a method to control the impact of noise on signals, spread spectrum method [11] is widely used. While spreading time-series signals may improve our method in noise tolerance, doing so is not desirable because the use of a wider frequency band requires signals to be temporally spread, requiring a longer detection time.

To prevent watermark signals from being deteriorated on each frame it may be effective to take into account that videos are compressed on a per macro block basis. Approaches by Bros et al. [13] and other researchers [14][15] that manipulate the DCT coefficient within each macro block control deterioration caused by compression by embedding watermarks in a low- to medium-frequency component within the macro block considering the fact that in video compression, high-frequency components, which are less perceptible to the human eye, are preferentially removed. However, our approach uses the average pixel value to generate watermark signals, meaning that it is not expected to provide an improved compression tolerance even if the frequencies are distributed in the low-frequency area within each macro block. Neither are approaches by Wu et al. [16] and other researchers that adaptively change the frequency range of the coefficient to be manipulated.

Proposed method

We propose to expand the pixels of the watermark to NxN. The amplitude of the signal does not change when decreasing the number of points embedded instead of increasing the pixel units. For instance, 160 points are varied when embedding in each 2x2 pixel while 640 points are varied in each 1x1 pixel.

Increasing the pixel units is expected to concentrate the temporal variations of the average pixel values in few macro blocks instead of spreading it between all macroblocks. Doing so will reduce the impact of the quantization of the video compression on the information transmitted through the watermarking.

Even if the spatial resolution decreases, the amount of information remains constant since our video watermark only uses temporal variation of average pixel value.

Our method controls each pixel unit so that it will not straddle over more than one macro block. If a unit extends over more than one block, it is, as a result, the same as a watermark embedded based on the pixel unit 1 x 1 as shown in Figure 6, which is not expected to improve the compression tolerance.

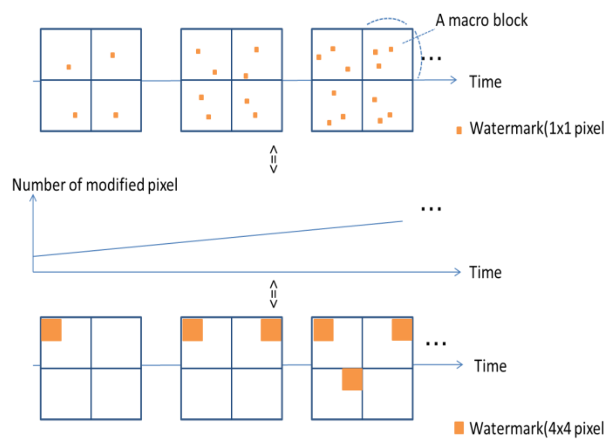


FIGURE 5 EVOLUTION OF THE NUMBER OF MODIFIED PIXELS

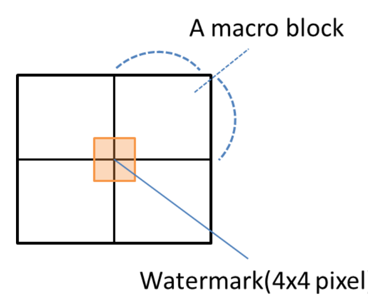


FIGURE 6 WATERMARK STRADDLING OVER MORE THAN ONE MACRO BLOCK

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069

Extremely concentrated local pixel changes, however, may make watermarks more perceptible to the human eye. In the past, studies were also made to control perceptibility based on the human visual characteristics (so called the human visual system). Among them are approaches by Sawazawa et al. [14], Simitopoulos et al. [15], and Chun et al. [17], which only focus on the AC component within each macro block without taking into consideration the unbalanced changes in DC components between blocks. Even approaches by Peter et al. [18] and Mansoor et al. [19] that directly manipulate the DC component have not evaluated the impact on image quality from that aspect.

This study verifies that our approach yields a greater improvement in the trade-off between compression tolerance and video quality than an approach that simply increases the signal strength. To this end, we evaluated:

(i) How much the compression tolerance can be improved by embedding watermark signals in a unit of N x N pixels to localize pixel changes; and

(ii) How the localization of pixel changes affect perceptibility under expected practical use conditions

RESULTS

We embedded 1x1 (previously used) and 4x4 and 8x8 (proposed) pixel unit watermarks in a video, and evaluated from following aspects.

(i) Compression tolerance

We encoded watermarked videos with MPEG-2 and H.264 video codecs assuming use for broadcasting. Encoding conditions, video sample information and watermark configurations are shown in Table2.

Table 2: (a) Encoding conditions

Codec	Resolution(pixel)	Frame rate	Bit rate (Mbps)
MPEG-2	1440x1080	30p	11 ※1
H.264	1920x1080	30p	2 ※2

※1 close to lower bound of digital terrestrial TV broadcasting in Japan based on ISDB-T standard

※2 close to lower bound of actual measurement value when home shopping network program of communication satellite broadcasting is received

(b) Video sample information

Video	Length of time
ITE/ARIB Hi-Vision Test Sequence 2nd Edition sequence C108	15 seconds

(c)Watermark configuration

Configuration of information	Cycle of information
ID code (16bit), check sum code, and error correcting code	1 seconds (embed repeatedly)

Watermark signal power

A watermark signal is expressed as a one-dimensional signal of the discrete time t. If a watermark signal in a compressed video is represented as X(t), the power of the signal is expressed as Equation (1). Figure 7 shows the results. They reveal that as the watermark pixel unit increases, the signal power becomes less likely to decrease.

$$\text{power} = \frac{1}{T}\sum_{t=0}^{T} X(t)^2 \quad (1)$$

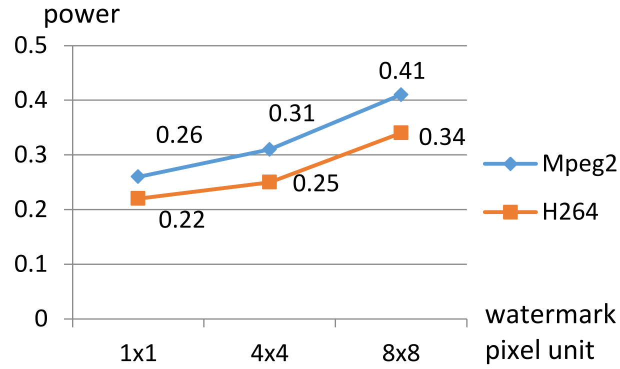


FIGURE 7 SIGNAL POWER AFTER COMPRESSION

Average detection time

We measured the average watermark detection time using a smart device app when the video was repeated three times. The experimental environment is shown in Table 3. Previous research by Shneidermans et al. [20] showed that the average user is willing to endure up to 4 seconds of response time. To meet this requirement, we aimed at a 3 seconds average. The results are shown in Figure 8. The average detection time decreased as watermark pixel units got larger, and the power improved. Notably, when compressing with H.264, using 1x1 pixel units leads to a detection time greater than 3 seconds while using 4x4 or bigger pixel units reduces it below 3 seconds.

Table3: Experimental environment

Display	SHARP: PN-465
Smart device	Fujitsu: Arrows NX F-01F
Distance from monitor	3.5 x monitor's height (Recommended watching distance by manufacturer)

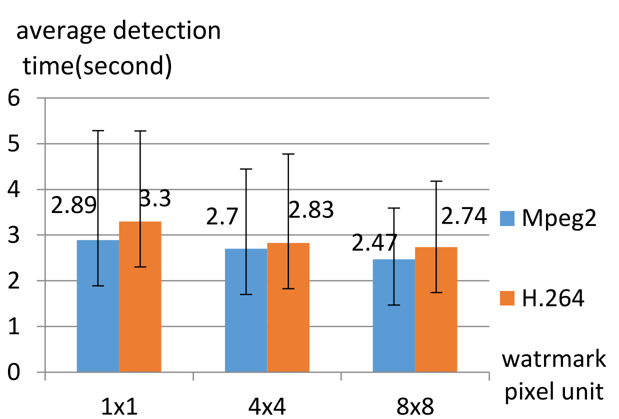


FIGURE 8 AVERAGE DETECTION TIME

DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069

(ii)Influence on image quality

Subjective evaluation of image quality

We made a subjective evaluation on how the watermark signal power and the size of embedding pixel unit affect the perceptibility of watermarks.

(a) Impact of watermark size on video quality (uncompressed)

We played two uncompressed watermarked videos and 7 evaluators judged which, if any, had poorer image quality. The watermark pixel units were 1x1 and 4x4. They evaluated each set of motion video twice. Evaluators were not informed of the pixel units. An 8x8 unit video was also compared with the 1x1 in same way. The monitor and distance were the same as in evaluation (i). The evaluation result was scored as shown in equation (2), and the evaluators' average score of both evaluations is shown in Figure 9. A positive score means the larger watermark pixel unit had greater influence on image quality. The evaluation results indicate that while the 4 x 4 pixel unit has only a slight impact on perceptibility, the 8 x 8 unit has a larger impact, making watermarks more perceptible than the 1 x 1 unit.

$$\text{score} = \begin{cases} -1 \text{ (if 1x1 pixel unit watermark appeared poorer quality} \\ 0 \text{ (if image quality appeared the same)} \\ 1 \text{ (if 4x4 or 8x8 pixel unit watermark appeared poorer quality)} \end{cases} \quad (2)$$

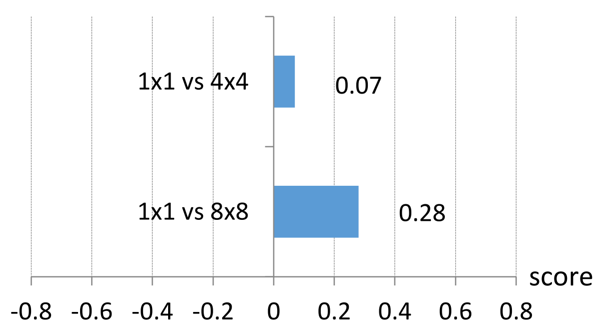


FIGURE 9 SUBJECTIVE EVALUATION OF IMAGE QUALITY

(b) Impact of signal power and watermark size on video quality (uncompressed)

Experiment (i) verified that an increase in watermark size improved the signal power after video compression. The signal power before the video compression was 0.46. We embedded watermarks with a different signal power into the 1 x 1 pixel unit and compressed video with the same conditions as Experiment (i). Figure 10 shows the signal power before and after the compression. This denotes that when the power before compression is increased to 0.48, the power after compression is the same as that for the 4 x 4 pixel unit regardless of whether the compression format is MPEG-2 or H.264. When the power is increased to 0.53 for MPEG-2 format, or to 0.55 for H.264 format, the power after compression is almost the same as that for the 8 x 8 pixel unit.

We evaluated the perceptibility of watermarks before video compression in the following two cases: one case where the signal power was increased with the pixel unit maintained at 1 x 1, and another where the pixel unit was increased to 4 x 4 and 8 x8 with the signal power maintained at 0.46. The evaluation method is the same as for Experiment (a) and the scores are calculated through Equation (3). Figure 11 shows the results. It indicates that when the pixel unit is 4 x 4, the watermarks are less perceptible than where the pixel unit is 1 x 1 with the signal power increased. It also indicates that when the pixel unit is 8 x 8, the watermarks are more perceptible, than where the pixel unit is 1 x 1 with the signal power increased.

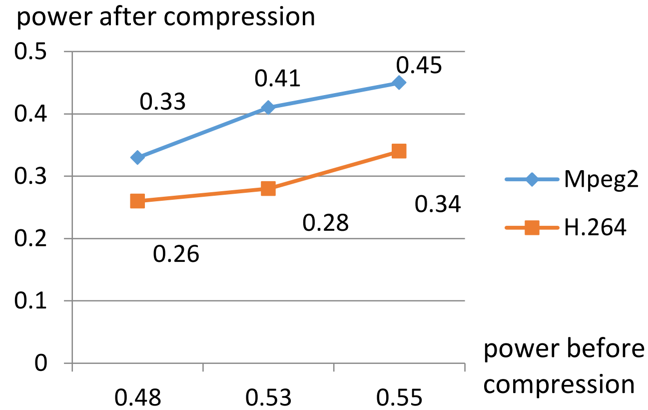


FIGURE 10 WATERMARK POWER BEFORE AND AFTER COMPRESSION

$$\text{score} = \begin{cases} -1 \text{ (if 1x1 pixel unit watermark with higher power appeared poorer quality)} \\ 0 \text{ (if image quality appeared the same)} \\ 1 \text{ (if 4x4 or 8x8 pixel unit watermark with power 0.46 appeared poorer quality)} \end{cases} \quad (3)$$

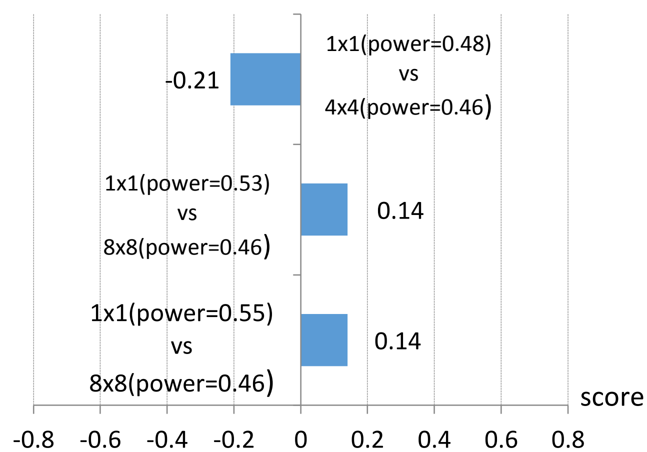


FIGURE 11 SUBJECTIVE EVALUATION OF IMAGE QUALITY

(c) Impact of signal power and watermark size on video quality (compressed)

We compared the perceptibility of watermarks after video compression between the following two cases: one case where the pixel unit is 1 x 1 with the signal power increased to 0.48, 0.53, and then 0.55, and another where the signal power is 0.46 with the pixel unit increased to 4 x 4 and then 8 x 8. The compression conditions are as shown in Table 2(a) and the scores for these evaluations were also determined through Equation (3).Figure 12

DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069

and Figure 13 show the results for MPEG-2 and H.264 compression, respectively.

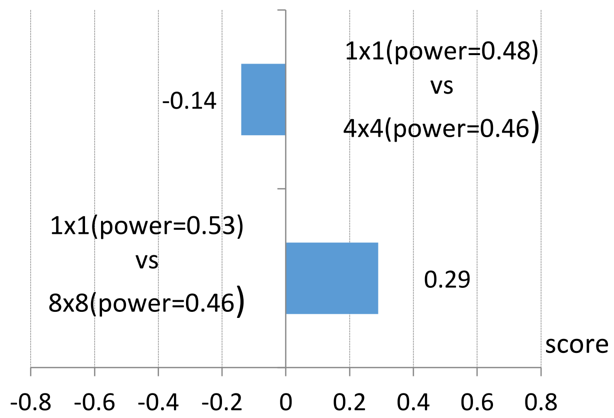


FIGURE 12 SUBJECTIVE EVALUATION OF IMAGE QUALITY (MPEG-2 COMPRESSION)

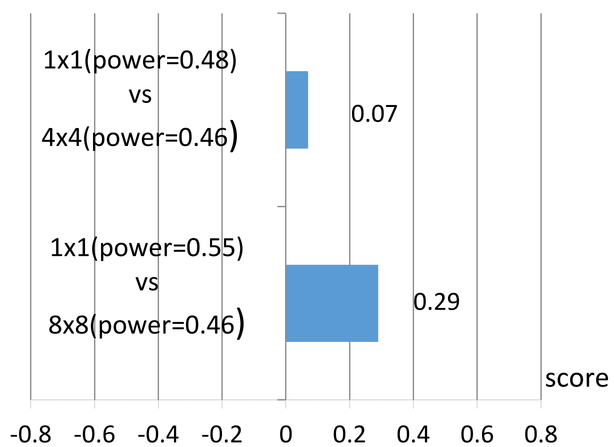


FIGURE 13 SUBJECTIVE EVALUATION OF IMAGE QUALITY (H.264 COMPRESSION)

The results indicate that for H.264 compression, the impact on the compressed video are almost the same between the case where the pixel unit is 1 x 1 with the signal power increased and where the pixel unit is 4 x 4. For MPEG-2 compression, a stronger power has a larger impact on the quality of compressed video as well as uncompressed video. The 8 x 8 pixel unit even has a larger impact on compressed video than where the pixel unit is 1 x 1 with the power increased.

These results reveal that the pixel unit, if increased to 8 x 8, has a larger impact on uncompressed and compressed video than in the case where the pixel unit is 1 x 1 with the signal power increased. They also indicate that for uncompressed videos and MPEG-2-compressed videos, the pixel unit, if increased to only 4 x 4, has a smaller impact on the perceptibility of watermarks than in the case where the pixel unit is 1 x 1 with the power increased. For H.264-compressed video, the impact on the perceptibility is almost the same. In other words, if both cases are required where the video is not compressed or compressed with a high bit rate, and where the video is compressed with a low bit rate, the pixel unit increased to 4 x 4 more effectively improves the trade-off between the compression tolerance and impact on video quality than the 1 x 1 pixel unit with the signal power increased.

Objective evaluation of image quality

An objective evaluation of image quality was made to consider the relation of it with the subjective evaluation. In video quality evaluation, two different objective metrics are widely used: PSNR and SSIM [21]. PSNR and SSIM are expressed as Equations (4) and (5), respectively. SSIM is calculated on a given local window cut out from an image. MSSIM, which is normally used as the metric for the entire image, is expressed as Equation (6).

PSNR: Assuming that Images I and K have M*N pixels;

$$\mathrm{PSNR} = 20\log_{10}\frac{\mathrm{MAX_I}}{\sqrt{\mathrm{MSE}}} \qquad (5)$$

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}}\sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)-K(i,j)]^2$$

where

$\mathrm{MAX_I}$: The possible maximum pixels of the image

SSIM: Assuming that window x and y cut out from Images I and K with M*N pixels have m*n pixels;

$$\mathrm{SSIM} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \qquad (6)$$

where

μ_x, μ_y: The averge pixel values of windows x and y

σ_x, σ_y : The variances of the pixel values of windows x and y

σ_{xy} : Covariance of the pixel values of windows x and y

c_1, c_2: Constant number

(in this paper $c_{1,} = 6.5025$ and $c_2 = 58.5225$).

MSSIM: Assuming L fragments of window with a respective SSIM value equal to $\mathrm{SSIM_l}$;

$$\mathrm{MSSIM} = \frac{1}{L}\sum_{l=1}^{L}\mathrm{SSIM_l} \qquad (7)$$

Irrespective of whether the pixel unit is 1 x 1 or N x N, the PSNR value stays constant so long as the total pixels of the embedded watermark remain unchanged. This fact is obviously inconsistent with the tendency revealed by the subjective evaluation that the pixel unit, if increased to 8 x 8, has a larger impact on the perceptibility of the watermarks.

It is known that compared with PSNR, SSIM is highly correlated with human perception. The SSIM value reflects the nature that steep local variations are more perceptible to the human eye; therefore, it can probably be more effectively used to evaluate the video quality offered by our method. However, considering that the SSIM value is calculated based on the 8 x 8 pixel unit against the image of 16 x 16 pixels as shown in Figure 14, it is true that the SSIM value may not be consistent with the tendency revealed by the subjective evaluation. Assume that all original images have a pixel value of 128 and that the change in pixel value caused by watermark superimposition is always +1. Then, assume that a watermark is superimposed on 16 pixels in each of windows a1 to a4 and on 64 pixels in window b1.

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069

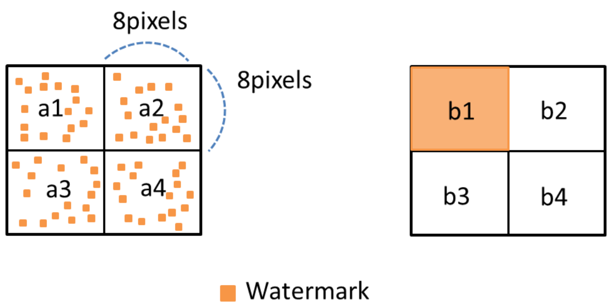


FIGURE 14 MODEL CASE OF SSIM CALCURATION

For each of windows a1 to a4, the SSIM value determined through Equation (6) is 0.999379. For Domain b1, the SSIM value is 0.99997 and for each of windows b2 to b4, it is 1.0; the average value is 0.999992. In other words, the SSIM value increases in accordance with an increase in pixel unit, which is inconsistent with the tendency revealed by the subjective evaluation. Now, assume that all original images have a pixel value of 0. The SSIM values are 9.89 for windows a1 to a4, 8.67 for window b1, and 1.0 for windows b1 to b4, with an average value of 9.67 for windows b1 to b4, which is consistent with the tendency revealed by the subjective evaluation. The reason the results are different is because, when the average pixel value is high, the difference between the average pixel values of original and compared pictures hardly acts as a factor that decreases the SSIM value. For this reason, SSIM is adjusted as Equation (8). In this paper, we call the metric given by Equator (8) RSSIM and the average of the RSSIM values for all windows MRSSIM.

$$\mathrm{RSSIM} = \frac{c_1(2\sigma_{xy} + c_2)}{((\mu_x - \mu_y)^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \qquad (8)$$

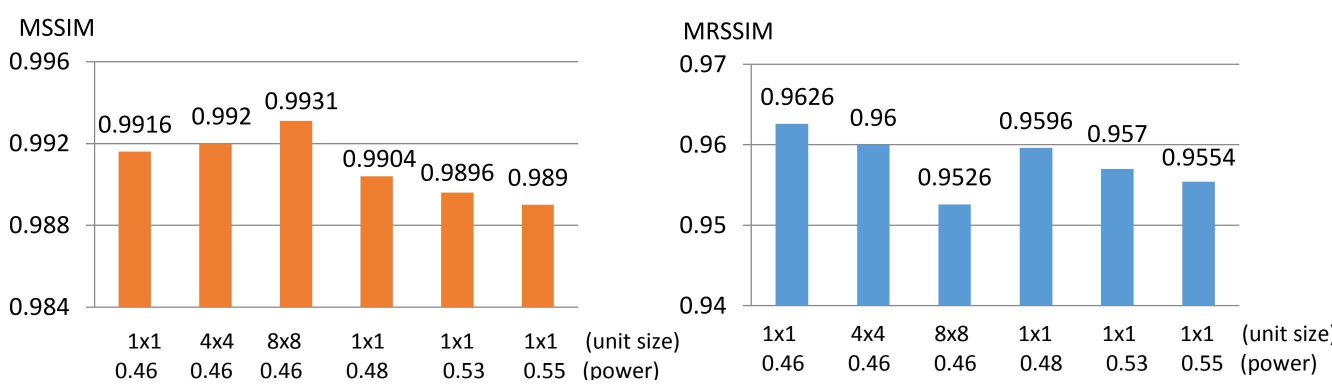
This modification is based on the assumption that the average pixel value for the original images is equal to 0. To get such a property, it is proposed to μ_x with $(\mu_x - \mu_x)$ and replacing μ_y with $(\mu_y - \mu_x)$. By doing so, SSIM value will be constant whether average pixel value is high or low.

Using the two metrics, MSSIM and MRSSIM, we evaluated the impact of the pixel unit and signal power on video quality.

•**Before compression:** Figure 15 shows the results. For the case watermarks are embedded in the actual motion video for evaluation, the MSSIM value improves if the watermark pixel unit is increased with the signal power maintained at a certain level. This result is also inconsistent with the results of the subjective evaluation—the results show that when a comparison is made between the case where pixel unit is 1 x 1 and the signal power is between 0.53 and 0.55 and where the pixel unit is 8 x 8 and the power is 0.46, people feel that the image quality is more deteriorated in the latter case. On the other hand, the MRSSIM value decreases as the pixel unit increases when the signal power remains unchanged. In the other cases where the pixel unit and signal power were changed, the increase and decrease in MRSSIM value are consistent with the results of the subjective evaluation. These facts denote that as the metric for measuring the impact of watermarks on video quality, MRSSIM is more suitable than MSSIM.

• **After compression:** Figures 16 and 17 show the results for MPEG-2 and H.264 compression, respectively. After compression, the MSSIM values indicate that the case where the pixel unit is 1 x 1 and the signal power is between 0.53 and 0.55 has a smaller impact on video quality than where the pixel unit is 8 x 8 and the power is 0.46. However, if a comparison is made between the case where the pixel unit is 1 x 1 and the power is 0.48 and where the pixel unit is 4 x 4 and the power is 0.46, the results are inconsistent with the results of the subjective evaluation. When MRSSIM is used as the metric, the impacts of the pixel unit and power on perceptibility are consistent with the results of the subjective evaluation. For H.264 compression, the MSSIM value is as low as 0.92. However, even if the video is compressed with no watermark embedded, the MSSIM is 0.9293; it should be hence noted that the deterioration is mainly caused by compression.

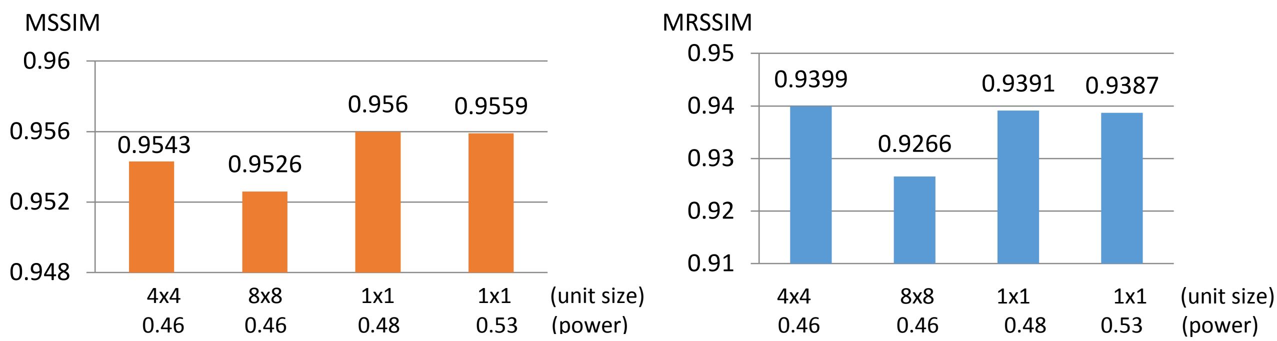
The findings above indicate that before or after compression, the MRSSIM values accurately reflect the results of the subjective evaluation.



(a)MSSIM (b)MRSSIM

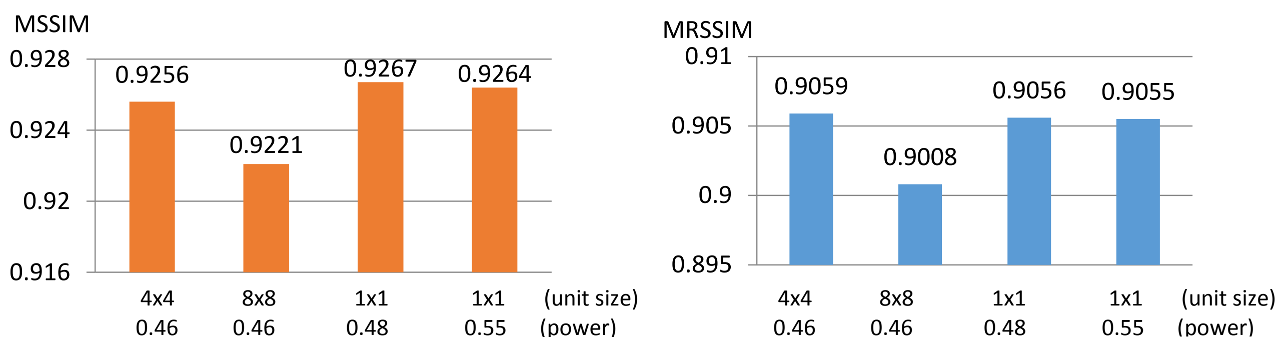
FIGURE 15 MSSIM AND MRSSIM (BEFORE COMPRESSION)

DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069



(a)MSSIM (b)MRSSIM

FIGURE 16 MSSIM AND MRSSIM AFTER MPEG-2 COMPRESSION



(a)MSSIM (b)MRSSIM

FIGURE 17 MSSIM AND MRSSIM AFTER H.264 COMPRESSION

Conclusion

This study researched a method to improve the trade-off between the compression tolerance and perceptibility of camera-detectable digital watermarks for videos based on temporal changes in average pixel value. We have improved the tolerance to compression by expanding the pixel unit in which watermarks are embedded to concentrate the temporal variations of average value in few macro blocks. Specifically, we made a subjective evaluation of the impact on the perceptibility of watermarks to determine a pixel unit suitable for TV broadcasting and viewing. This makes it possible to standardize the signal parameters for more applications that distribute content in different broadcasting forms and at different bit rates, providing a simplified service.

As an objective metric of image quality, we used RSSIM, which excludes the impact of the average pixel value from SSIM, a metric widely used to measure video quality. This has revealed that the RSSIM accurately reflects the tendency revealed by the subjective evaluation. The subjective evaluation conducted this time, however, involved only relative assessment. In the future, we need to examine the correlation with the results of absolute assessment to verify that the RSSIM is a reliable metric. Even though a subject to be discussed remains, considering the fact that the metric accurately reflects the tendencies related to increases and decreases, it is expected that RSSIM can be a useful metric for further development improving the trade-off.

References

[1]. Digimarc Corporation, “Digimarc Discover | Mobile Engagement, Faster Checkout, And More”, http://www.digimarc.com/products/discover.

[2]. Kantar Media, “Enrich the TV viewing experience while collecting individual viewing data with SyncNow | Kantar Media”, http://www.kantarmedia.com/global/thinking-and-resources/latest-thinking/enrich-the-tv-viewing-experience-while-collecting-individual-viewing-data-with-syncnow

[3]. Fujitsu Limited, “Fujitsu Technology Supports Japan's TV Shopping Giant JAPANET TAKATA for a New Shopping Experience”, http://www.fujitsu.com/global/about/resources/news/press-releases/2013/1205-01.html.

[4]. Vobile, “mSync - 2nd Screen | Interactive Viewing”, http://www.vobileinc.com/msync/.

[5]. Taehyun Kim, Kyungoh Park and Youpyo Hong, “Video watermarking technique forH.264/AVC”, Optical Engineering, vol. 51(4), 047402, 2012.

[6]. M. Noorkami and R. M. Mersereau, “Compressed-domain video watermarking for H.264,” in Proceedings of International Conference on Image Processing (ICIP), 2005. .

[7]. Shigeyuki Sakazawa, Yasuhiro Takishima, and Yasuyuki Nakajima, “H.264 Native Video Watermarking Method”, 2006 IEEE

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-069

International Symposium on Circuits and Systems, ISCAS 2006. Proceedings, 2006..

[8]. P. Campisi, "Video watermarking in the 3D-DWT domain using quantization-based methods," presented at IEEE Int.Workshop on Multimedia Signal Processing, 2005.

[9]. Osama S. Faragallah, "Discrete wavelet transform–based spatial-temporal approach for quantized video watermarking", Optical Engineering, vol. 50(7), 077402, 2011

[10]. T. Kalker, G. Depovere, J. Haitsma and M. Maes, "A video watermarking system for broadcast monitoring", Proceedings of IS&T/SPIE/EI25, Security and Watermarking of Multimedia, Content, San Jose, vol. 3657, pp. 103-112, 1999.

[11]. Jaap Haitsma and Ton Kalker, "A WATERMARKING SCHEME FOR DIGITAL CINEMA", Proc. ICIP, Vol.2, pp.487-489, 2001.

[12]. I. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Process.6, 1673–1687 , 1997.

[18]. N Peter H. W. Wong, Oscar C. Au and Justy W. C. Wong, "Data Hiding and Watermarking in JPEG Compressed Domain by DC Coefficient Modification", In Security and Watermarking of Multimedia Contents II, Ping Wah Wong, Edward J. Deip, Editors, Proceedings of SPIE Vol. 3971 ,2000

[19]. Nafees Mansoor , A. K. M. Muzahidul Islam and M. Abdur Razzak," Multiple description discrete cosine transform-based image coding using DC coefficient relocation and AC coefficient interpolation", J. Electron.Imaging.22(2), 023030, 2013.

[20]. B. Shneiderman and C. Plaisant, "DESIGNING THE USER INTERFACE 4TH EDITION," Addison-Wesley Publishing Co., 2005.

[21]. Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.

Author Biography

Keizo Kato belongs to Media Processing Loboratory at FUJITSU LABORATORIES LTD. in, Kawasaki, Japan. He is currently engaged in the research of digital watermarking and visible light communications.

Kensuke Kuraki belongs to Media Processing Loboratory at FUJITSU LABORATORIES LTD. in, Kawasaki, Japan. He is currently engaged in the research of digital watermarking and visible light communications.

Ryuta Tanaka belongs to Media Processing Loboratory at FUJITSU LABORATORIES LTD. in, Kawasaki, Japan. He is currently engaged in the research of digital watermarking and visible light communications, video transmission, and delivery technology.

[13]. A. Bors and I. Pitas, "Image watermarking using DCT domain constraints," in Proc. of IEEE. Int. Conf. Image Processing, Lausanne, pp. 231–234, 1996.

[14]. S. Sakazawa, Y. Takashima and M. Wada, "A Watermarking Method Retrievable from MPEG Compressed Stream", IEICE Trans. Vol. E 85-A No.11, pp. 2535-2543, 2002.

[15]. D. Simitopoulos, S. A. Tsaftaris, N. V Boulgouris', and M. G. Strintzis, "COMPRESSED-DOMAIN VIDEO WATERMARKING OF MPEG STREAMS" in Proc. 2002 IEEE International Conference on Multimedia and Expo, 1, pp. 569‑572, 2002.

[16]. Guo-Zua Wu, Yi-Jung Wang, Wen-Hsing Hsu, "Robust watermark embedding/detection algorithm for H.264 video", Journal of Electronic Imaging 14(1), 013013, 2005.

[17]. T.-Y. Chung, M.-S. Hong, Y.-N. Oh, D.-H. Shin, and S.-H. Park, "Digital watermarking for copyright protection of MPEG2 Compressed video," IEEE Trans. Consum. Electron., vol. 44, no. 3, pp.895–901,1998