Impact of Rounding Error on Spread Spectrum Fingerprinting Scheme

Kuribayashi, Minoru

Kato, Hiroshi

Impact of Rounding Error on Spread Spectrum Fingerprinting Scheme

Minoru Kuribayashi *Member, IEEE,* and Hiroshi Kato *nomember*

***Abstract*—In spread spectrum fingerprinting, it has been considered that the strength of the embedded signal is reduced to $1/c$ of its original value when c copies are averaged by colluders. In this study, we analyze the model of the averaging attack by considering quantization that causes nonlinear changes in the fingerprint sequence. Our detailed analysis reveals that the attenuation of the signal energy strongly depends on the quantization performed during the embedding and averaging stages. We estimate the actual attenuation factor from the perspective of a stochastic model in the spatial domain and derive an attenuation factor that differs considerably from the conventional one. Our simulation result indicates that the actual attenuation factor is classified into the best and worst cases from the detector's perspective. Furthermore, we demonstrate that colluders can select the worst case by comparing their fingerprinted copies. A countermeasure for preventing the worst-case scenario is also proposed in this paper.**

***Index Terms*—fingerprinting, collusion attack, quantization error**

I. Introduction

Digital fingerprinting is a technique that is used for tracing illegal users; in this technique, a unique ID called a digital fingerprint [1] is embedded into a content before distribution. When a suspicious copy of a content is found, an owner can identify an illegal user by extracting the fingerprint. One of the serious issues in a fingerprinting system is its susceptibility to a collusion attack; in this type of attack, several users combine their copies of the same content to modify/delete the embedded fingerprints. In a simple form of this attack, multiple copies of the same content are averaged. By sufficiently combining several copies, the embedded fingerprints can be weakened or removed by this attack.

Cox et al. [2] proposed a framework for collusion-resistant fingerprinting. In this method, a spread spectrum sequence is assigned to each user and embedded into the digital content using a watermarking technique. Because these spread spectrum sequences are statistically orthogonal to each other, a detector can identify the sequence contained in an illegal copy. In their study, the correlation scores of colluders' sequences were estimated to be reduced to $1/c$ of their original value if c copies were averaged by them. This estimation is supported by related works [3], [4]. Zhao et al. [5] analyzed the effects of other collusion attacks on spread spectrum watermarking and reported that numerous nonlinear collusion attacks such as the interleaving attack can be well approximated by averaging collusion along with additive noise. Thus far, it has been considered that the signal energy of spread spectrum fingerprinting is linearly attenuated by a factor of c and the collusion resistance is evaluated under the attenuation factor.

In this study, we carry out a detailed analysis of the effects of rounding error on spread spectrum fingerprinting. When a fingerprint signal is embedded into the frequency domain with a floating-point number, a fingerprinted image is derived by transforming it into the spatial domain. Because the pixel value is an integer in the range $[0, 255]$, a rounding operation must be performed. Suppose that when c malicious users collect their copies of the same image, a pirated copy can be produced by averaging these copies. During the computation for averaging, each pixel value must be rounded to an integer if the sum of c pixels cannot be divided by c. Among several rounding operations that can be used for this purpose, the implementation "flooring," "ceiling," and "rounding to nearest integer" are the most. We consider the combination of such rounding operations for the embedding and averaging stages, and an attenuation factor is estimated with respect to the truncated decimal numbers for the averaging stage. For the purpose of detection, an averaged copy is transformed into the frequency domain and compared with the original copy to detect the fingerprint sequence. The extracted sequence comprises multiplexed fingerprint sequences and quantization noise. Under the assumption that the distorted sequence can be restored in the nearest normal distribution, the variance is estimated and the associated attenuation factor c' is obtained. This assumption is validated via a comparison with the experimental value. The result indicates that the detected fingerprint strength of an averaged copy is $1/c'$ times the original strength and this is completely different from the conventional value $1/c$.

First, we analyze the distortions caused by the rounding operation under the condition that either the "flooring" or "rounding to nearest integer" operation is selected during the embedding and averaging stages. If the rounding operation selected during the embedding stage differs from that selected during the averaging stage, the attenuation of the fingerprint signal decreases considerably if the operations are the same, however the attenuation increases considerably. In the case of colluders, it is possible to determine the rounding operation performed at the embedding stage by comparing their copies. In order to avoid the worst-case scenario of colluders selecting the same rounding operation, we propose an approach for making this selection difficult. Because the rounding operation is performed after transforming the fingerprinted frequency

The authors are with the Graduate School of Engineering, Kobe University, Kobe 657-8501, Japan. (e-mail: kminoru@kobe-u.ac.jp)

components back to the spatial domain, we can randomly select the rounding operation at the pixels. With this conversion, the attenuation of the fingerprint signal can be restricted. Next, we increase the number of candidates for the rounding operation and improve the attenuation factor c'. Such a rounding operation is closely related to the dithered operation reported in [6] that is used to perform a randomized quantization in the frequency domain of an image. In this study, we propose methods for estimating the attenuation factor of an embedded fingerprint with respect to the number of colluders and for theoretically deriving the factor considering the rounding operations performed at the embedding and averaging stages on the basis of the statistical behavior of the fingerprint signal in the spatial domain.

The remainder of this paper is organized as follows. In Section II, we review related works and the model of the collusion attack. Section III presents an analysis of the effects caused by performing the rounding operation during the embedding and averaging stages. Section IV describes an attack strategy that works in favor of the colluders. Section V presents a countermeasure for this attack strategy along with experimental results. Finally, Section VI concludes this paper.

II. Preliminaries

In this section, we review spread spectrum fingerprinting schemes and describe the model of collusion attacks.

A. Spread Spectrum Fingerprinting

In Cox's spread spectrum fingerprinting scheme [2], a fingerprint sequence is independently selected from random values that follow the normal distribution $N(0,1)$. Such a random sequence is amplified using the characteristic of the selected frequency components of an image and is embedded into the components. The fingerprint sequence is extracted from the frequency components of a suspicious copy by subtracting them from the frequency components of an original image. The fingerprinted image for each distinct fingerprint differs slightly. Hence, malicious users compare c copies and try to eliminate this difference.

Let $\boldsymbol{W}$ be a watermark signal composed of ℓ elements $w_j \in N(0,1), (1 \leq i \leq \ell)$. These elements are embedded into a selected DCT coefficient $d_i, (1 \leq i \leq \ell)$, based on the following equation:

$$d_i' = d_i(1 + \alpha w_i), \tag{1}$$

where $N(0,1)$ is a normal distribution with mean 0 and variance 1 and α is an embedding strength. At the detector side, we determine which SS sequence is present in a test image by evaluating the similarity of the sequences. A sequence $\hat{\boldsymbol{W}}$ is extracted by calculating the difference between the suspicious copy and the original image; the similarity of $\hat{\boldsymbol{W}}$ with $\boldsymbol{W}$ is obtained as follows:

$$\mathsf{sim}(\boldsymbol{W}, \hat{\boldsymbol{W}}) = \frac{\boldsymbol{W} \cdot \hat{\boldsymbol{W}}}{\sqrt{\hat{\boldsymbol{W}} \cdot \hat{\boldsymbol{W}}}}. \tag{2}$$

If this value exceeds a certain threshold, the embedded sequence is regarded as $\boldsymbol{W}$. When an original image is available, the above similarity measurement is valid because the main interference term, the frequency components of the original image, can be completely eliminated at the detection stage. In a fingerprinting system, it is assumed that an original image is available at the detection stage because the operation is performed by the author or his agent. Hence, at the detection stage, DCT coefficients of a test image are subtracted from those of the original image, after which the correlations with all candidates of the watermark signal are computed. Thus, a non-blind and informed watermarking scheme can be applied. Thus far, several variants of spread spectrum watermarking schemes based on Cox's method have been proposed [7], [8], [9], [10], [11], particularly for applications to sequences whose elements are randomly selected from normally distributed values.

A common disadvantage of Cox's scheme and its variants is that considerable computational resources are required for the detection because the correlation scores of all spread spectrum sequences are required to be calculated. For the reduction of computational costs, hierarchical spread spectrum fingerprinting schemes have been proposed. Wang et al. [9] proposed a scheme in which a set of users is divided into different subsets and each subset is assigned to a specific group whose members are more likely to collude with each other than with members from other groups. With the assumption that the users in the same group are equally likely to collude with each other, the fingerprints within one group have equal correlation. At the detection stage, the independence among the groups limits the number of innocent users falsely placed under suspicion within a group as the probability of accusing another group is very large. Suppose that each group can accommodate up to M users. The fingerprint sequence $\boldsymbol{W_{i,j}}$ assigned to the j-th user within the i-th group consists of two components:

$$\boldsymbol{W_{i,j}} = \sqrt{1-\rho}\boldsymbol{e_{i,j}} + \sqrt{\rho}\boldsymbol{a_i}, \tag{3}$$

where $\{\boldsymbol{e_{i,1}}, \boldsymbol{e_{i,2}}, \ldots, \boldsymbol{e_{i,M}}, \boldsymbol{a_i}\}$ are the orthogonal basis vectors of group i with equal energy and ρ is referred to as the intra-group correlation. Because of the presence of the common vector $\boldsymbol{a_i}$, when colluders from the same group average their copies, the energy of the vector is not attenuated; hence, the detector can accurately identify the group. The detection algorithm consists of two stages. one involves the identification of groups containing colluders, and the other, the identification of colluders within each suspicious group.

This concept of grouping has been applied to variants of spread spectrum fingerprinting [10], [11] and to the construction of a collusion secure code [12]. In [11], two components of the fingerprint sequence given by Eq.(3) were designed using DCT basic vectors modulated by PN sequences such as the M-sequence and Gold sequence [13] in order to further reduce the computational costs. With a fast DCT algorithm, the computation of correlation scores at the detector is reduced to the logarithmic scale. The fingerprint sequence assigned to the j-th user within the i-th group is represented as follows:

$$\boldsymbol{W_{i,j}} = \boldsymbol{pn(i)} \otimes \boldsymbol{dct(j, \beta_u)} + \boldsymbol{pn(s)} \otimes \boldsymbol{dct(i, \beta_g)}, \tag{4}$$

where $\boldsymbol{pn}(\boldsymbol{x})$ is a PN sequence of length ℓ generated using an initial value x; s, a secret key; and $\boldsymbol{dct}(\boldsymbol{i}, \boldsymbol{\beta})$, the i-th DCT basic vector of strength β and length ℓ. $\otimes$ implies element-wise multiplication. The terms $\boldsymbol{pn}(\boldsymbol{i}) \otimes \boldsymbol{dct}(\boldsymbol{j}, \boldsymbol{\beta_u})$ and $\boldsymbol{pn}(\boldsymbol{s}) \otimes \boldsymbol{dct}(\boldsymbol{i}, \boldsymbol{\beta_g})$ in Eq.(4) respectively correspond to $\sqrt{1-\rho}\boldsymbol{e_{i,j}}$ and $\sqrt{\rho}\boldsymbol{a_i}$ in Eq.(3). The energy of the fingerprint sequence is represented by $\beta^2 = \beta_g^2 + \beta_u^2$. The additional correspondence relationships are $\sqrt{1-\rho} = \beta_u$ and $\sqrt{\rho} = \beta_g$.

As in the case of Cox's method, the fingerprint sequence $\boldsymbol{W_{i,j}}$ is embedded into randomly selected frequency components of an image. In the spatial domain, the embedded fingerprint sequence is spread over the entire image; the spread sequence in the spatial domain is denoted by $\boldsymbol{w_i}$. The changes introduced in different pixels by embedding the fingerprint signal at one frequency component are not independent. However, the eventual changes in pixels are summations of the changes caused by the fingerprint signal embedded into ℓ frequency components. At each pixel, the sum of the spread signal is considered to be the sum of random variables from a statistical point of view. If ℓ is sufficiently large, the sum approaches a normal distribution because of the central limit theorem. The mean of the sum should necessarily be zero because the DC component is excluded when embedding the fingerprint signal. Therefore, when the number of pixels is L, the variance of distribution at each pixel is statistically equal to β^2/L.

B. Collusion Attack

When colluders come together with c differently fingerprinted copies D_i containing a fingerprint sequence $\boldsymbol{w_i}$ in the same content D, they combine these copies to produce a colluded version $\hat{D}_c$. Because no colluder is willing to face more risk than other colluders, attackers usually agree to share the risk evenly among themselves. In [4], [14], collusion attacks were modeled by averaging and adding noise. Several types of collusion attacks were studied in [3], and combinations of several nonlinear collusion attacks were analyzed in [5]. In [15], the fairness of the collusion process in the case of collection of different resolutions of copies was studied. Hence, averaging with an equal weight is a reasonable assumption. Based on this attack model, the pirated copy $\hat{D}_c$ is expressed as

$$\hat{D}_c = \frac{1}{c}\sum_{i=1}^{c} D_i + \epsilon = \frac{1}{c}\sum_{i=1}^{c} \boldsymbol{w_i} + D + \epsilon. \tag{5}$$

From this model, the energy of the embedded fingerprint is attenuated by a factor of c. It is found that the colluded content retains better perceptual quality than the individual fingerprinted content. When the strength of the fingerprint sequence is β, the expected value that a detector outputs is β/c. From the viewpoint of energy, the original fingerprint sequence retains an energy β^2; this value reduces to β^2/c^2 after averaging. Because c fingerprint sequences with energy β^2/c^2 are retained in an averaged copy, the total energy corresponding to the embedded fingerprints is β^2/c. In this analysis, the attenuation factor with respect to energy is estimated as c. Because a fingerprint is spread throughout the entire original image, the cut-and-paste collusion attack, occasionally referred to as the interleaving attack, has similar effects on the averaging attack; all colluders contribute equally to the attack [5]. In both attacks, the attenuation of the fingerprint corresponds to the number of copies involved in the collusion. In this paper, we focus on averaging without noise ϵ.

A previous study [16] investigated the risks caused by selfish colluders who break the fairness agreement of sharing the risk evenly among themselves during the collusion. The presence of selfish colluders increases the complexity involved in the analysis of the attenuation of the fingerprint. Hence, in this paper, we assume that all colluders adhere to the fairness agreement, provide each other with their respective fingerprinted copies, and average these copies to produce a pirated copy.

III. Analysis of Rounding Operations

In this section, we study the effect of rounding errors caused during the embedding and averaging stages and measure the attenuation factor, denoted by c'. Most studies typically consider that $c' = c$.

A. Rounding Operation

There are several ways of rounding a number x to an integer y. The most common ones are "round to nearest (RN)," "round towards zero (TRUNCATE)," "round down (FLOOR)," and "round up (CEIL)". Most programming languages provide functions or a special syntax to round fractional numbers in various ways. Some programming languages such as FORTRAN and C provide only one method, usually the TRUNCATE operation. Other types of rounding methods have to be programmed explicitly; for example, RN can be implemented by adding 0.5 to the number to be rounded and then employing the TRUNCATE operation. In this paper, we assume that our programing language provides only the TRUNCATE operation.

Because a fingerprint sequence is embedded into the frequency domain of an image in spread spectrum fingerprinting, a rounding operation must be performed when the fingerprinted frequency components are transformed into the spatial domain because the pixel value is an integer in the range $[0, 255]$. In addition, a fractional part must be rounded to an integer when an averaging collusion is performed using c copies of the same image. With the exception of the underflow, only positive numbers are rounded in these operations, and hence, the TRUNCATE operation is equivalent to the FLOOR operation. In such a case, for any integer x and a number $\delta \in \mathbb{R}, 0 \leq \delta < 1$, a number within the range $[x-\delta, x+1-\delta)$ is rounded to x. Hereafter, we employ the RN ($\delta = 0.5$) and FLOOR ($\delta = 0$) operations for the analysis of the rounding error because these operations are easy to implement on a computer.

The quantization effects on an additive watermark such as a spread spectrum watermark were analyzed in [17]. In this study, the uniform scalar quantization of watermarked contents was theoretically investigated, and the results were extended to dithered quantization. These studies mainly focused on the effect of a single watermark signal embedded in a content,

whereas in the present study, we aim to estimate the quantization effects on an averaged copy by considering the rounding operations performed during the embedding and averaging stages.

B. RN During Embedding

The energy β^2 inserted in the frequency domain of an image is spread over the spatial domain by employing an orthogonal operation with floating-point numbers. If the number of pixels is L, the change at each pixel is expected to follow the normal distribution $N(0, \beta^2/L)$. The change at each pixel is equal to the fingerprint signal in the spatial domain, and the sum of variances is equal to the energy of the fingerprint signal β^2. For convenience, we denote the variance β^2/L as σ^2. Suppose that a floating-point number is rounded to the nearest integer (RN) by employing the TRUNCATE operation with $\delta = 0.5$. Because the distribution $N(0, \sigma^2)$ is considered to be the probability density function (PDF) of the change at each pixel, the probability $P(x)$ that the change is rounded to a value $x \in \mathbb{Z}$ is calculated as follows:

$$P(x) = \int_{x-\delta}^{x+1-\delta} f(t, \mu, \sigma^2)dt, \tag{6}$$

$$= \int_{x-0.5}^{x+0.5} f(t, \mu, \sigma^2)dt, \tag{7}$$

where

$$f(t, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right). \tag{8}$$

It should be noted that the mean of the distribution is $\mu = 0$, as mentioned in Sect.II-A.

After quantization by RN, the fingerprint sequence $\boldsymbol{w_i}$ in the spatial domain is rounded to an integer, denoted by $\hat{\boldsymbol{w}}_{\boldsymbol{i}}$. The mean $\hat{\mu}$ and variance $\hat{\sigma}^2$ of the distribution are represented by

$$\hat{\mu} = \sum (x - \mu)P(x), \tag{9}$$

$$\hat{\sigma}^2 = \sum (x - \mu)^2 P(x). \tag{10}$$

Because the PDF $f(t, 0, \sigma^2)$ is an even function centered on zero, the mean becomes $\hat{\mu} = 0$ after quantization. It is noticed that after the rounding operation during the embedding stage, the fingerprint $\hat{w}_i$ for each user is no longer Gaussian but a Gaussian random variable quantized to an integer. The variance $\hat{\sigma}^2$ of the discrete values is, however, almost equal to the original variance σ^2. Due to the limitation of the range $[0, 255]$ of the pixel value, only the tail of the original distribution is deleted by the rounding operation, resulting in a very small difference. Hence, we regard the distribution of the discrete values as $N(0, \hat{\sigma}^2)$.

1) FLOOR During Averaging: In the averaging attack involving c colluders, the value of each pixel is summed and divided by c. The operation performed using floating-point numbers is equivalent to summing the pixel values divided by c. In this case, the distribution of the averaged values is Gaussian with mean 0 and variance $\hat{\sigma}^2/c$. The averaged values are discrete because of the rounding operation performed during the embedding stage. For the sake of convenience, we assume that they are represented by analog values and consider the distribution $N(0, \hat{\sigma}^2/c)$ to be the PDF in order to maintain a correspondence relationship with the original fingerprint signal spread over the spatial domain.

It should be noted that the probability $P(x)$ is derived by the integral of $f(t, 0, \sigma^2)$ for the range $[x - 0.5, x + 0.5)$ during the embedding stage. After c copies are averaged, the range is reduced to $[x - 0.5/c, x + 0.5/c)$, which is the minimum scale to indicate the averaged values if they are represented by analog values. Without loss of generality, for the distribution $N(0, \hat{\sigma}^2/c)$ of averaged values, a value within the range $[x - 0.5/c, x + 0.5/c)$ is rounded to x when RN is performed during the embedding stage. Similarly, a value within $[x + 1 - 0.5/c, x + 1 + 0.5/c)$ is rounded to $x + 1$. Therefore, when FLOOR is performed on the sum of the pixel values of c copies, a value within the range $[x - 0.5/c, x + 1 - 0.5/c)$ is rounded to x for the distribution $N(0, \hat{\sigma}^2/c)$. If this value is positive, then the corresponding fingerprint strength decreases; otherwise, it increases. In this case, the probability $P_c(x)$ that the fingerprint value is rounded to x after averaging c copies is derived from the integral computation of the PDF in the range $[x-0.5/c, x+1-0.5/c)$:

$$P_c(x) = \int_{x-0.5/c}^{x+1-0.5/c} f\left(t, 0, \frac{\hat{\sigma}^2}{c}\right) dt. \tag{11}$$

By performing FLOOR, the averaged fingerprint signal $w_{i,j}/c, (1 \leq j \leq L)$, is rounded to $\lfloor w_{i,j}/c \rfloor$, resulting in the degradation of the fingerprint sequence $\boldsymbol{w_i}$.

The variance of the PDF is reduced to $\hat{\sigma}^2/c$ after averaging c copies, and $P_c(0)$ is increased accordingly. The fingerprint signals that are generated from spread spectrum sequences are not orthogonal but quasi-orthogonal to each other. Thus, the fingerprint signals embedded in fingerprinted copies interfere with each other slightly, and the variance of the averaged fingerprint signals at each pixel must be decreased; this variance is expected to be reduced by a factor of c. Hence, the amplitude of the remaining signals at each pixel is narrowed down. It is noteworthy that when the summed value is negative, the value is never rounded to 0 by FLOOR irrespective of how small a decimal number it is. Therefore, the rounded value asymptotically becomes -1 with an increase in c if the sum is negative; otherwise, it becomes 0.

$$P_c(-1) + P_c(0) \simeq 1 \tag{12}$$

In this case, probabilities $P_c(-1)$ and $P_c(0)$ are the integral values in the ranges $[-1 - 0.5/c, -0.5/c)$ and $[-0.5/c, 1 - 0.5/c)$, respectively. With an increase in c, the ranges approach $[-1, 0)$ and $[0, 1)$, and hence, $P_c(-1) \simeq P_c(0)$. Next, the mean of the multiplexed c fingerprint signals at each pixel statistically approaches $\hat{\mu}_c = -0.5$, implying changes in the DC component in the frequency domain. Even if the DC component is changed, the embedded fingerprint sequence is not affected because the DC component is avoided during the embedding stage. In such a case, the variance $\hat{\sigma}_c^2$ is calculated as follows:

$$\sum_x (x - \mu_c)^2 P_c(x) = 0.5^2 P_c(-1) + 0.5^2 P_c(0) = 0.25. \tag{13}$$

Thus, a portion of the fingerprint signal remains in the averaged copy even if the number of colluders increases. It is observed that the fingerprint sequence is distorted by a rounding operation. Because the quantization of the spatial domain affects the entire frequency domain, the detectable energy corresponding to the fingerprint is expected to be smaller than the total energy remaining in an averaged copy.

2) RN During Averaging: The effects of RN during the averaging stage differ greatly from those of FLOOR. In this case, the value within the range $[x-0.5-(c \bmod 2)/c, x+0.5-(c \bmod 2)/c)$ is rounded to x at the averaging stage for the distribution $N(0, \hat{\sigma}^2/c)$ of the averaged value. Therefore, the range of $P_c(x)$ differs slightly irrespective of whether c is an odd or an even number. If c is odd, the probability $P_c(x)$ is represented by

$$P_c(x) = \int_{x-0.5}^{x+0.5} f\left(t, 0, \frac{\hat{\sigma}^2}{c}\right) dt; \tag{14}$$

otherwise, it is represented by

$$P_c(x) = \int_{x-0.5-0.5/c}^{x+0.5-0.5/c} f\left(t, 0, \frac{\hat{\sigma}^2}{c}\right) dt. \tag{15}$$

Because the range $[x-0.5-(c \bmod 2)/c, x+0.5-(c \bmod 2)/c)$ is approximated by $[x-0.5, x+0.5)$ with the increase in c, the difference between the above two cases becomes negligible. By performing RN, $w_{i,j}$ shifts to $\lfloor w_{i,j}/c + 0.5 \rfloor$; hence, the fingerprint sequence is distorted. With the increase in c, the distribution of the fingerprint $w_{i,j}$ shrinks and the probability $P_c(0)$ is increased because the summed value within the range $[-0.5, 0.5)$ is rounded to 0. As a result, the attenuation of the fingerprint energy is much greater than that in the conventional estimation.

C. FLOOR During Embedding

Suppose that a floating-point number is quantized by employing the FLOOR operation during the embedding stage. Therefore, the probability that a fingerprint value is rounded to a value $x \in \mathbb{Z}$ in a spatial domain is calculated as follows:

$$P(x) = \int_{x}^{x+1} f(t, 0, \sigma^2) dt. \tag{16}$$

The mean value of the quantized signal at the spatial domain becomes $\hat{\mu} = -0.5$ from Eq.(9) and Eq.(16) because of the interval of integration.

When FLOOR is performed during the averaging stage, the value within the range $[x-0.5/c, x+1-0.5/c)$ is rounded to x for the distribution $N(-0.5, \hat{\sigma}^2/c)$ of the averaged value. In this case, the probability $P_c(x)$ is the integral value of $f(t, -0.5, \hat{\sigma}^2/c)$ in the range $[x-0.5/c, x+1-0.5/c)$. The mean value $\hat{\mu}$ becomes -0.5 and the DC component in the frequency domain is decreased; the effects on the fingerprint sequence in the frequency domain can be ignored. Therefore, $P_c(x)$ is represented by

$$P_c(x) = \int_{x-0.5-0.5/c}^{x+0.5-0.5/c} f\left(t, -0.5, \frac{\hat{\sigma}^2}{c}\right) dt. \tag{17}$$

As in the case in which RN is performed both during the embedding and the averaging stages, the probability $P_c(0)$ is increased because the integral value within the range $[-0.5, 0.5)$ is rounded to 0, and hence, the attenuation of the fingerprint energy is much greater than that in the conventional estimation.

Further, when colluders perform RN at the averaging stage, the distribution of $w_{i,j}$ shifts from 0 to $+0.5$; a part of the fingerprint signal still remains in an averaged copy even if the number of colluders increases. Therefore, we can consider that the combination of rounding operations during both the embedding and the averaging stages is a very important factor for evaluating the collusion resistance.

D. Attenuation Factor

A fingerprint sequence $\boldsymbol{W_i}$ embedded into the frequency domain of an image is spread over the spatial domain by means of an inverse orthogonal transform. Because the pixel value must be an integer within the range $[0, 255]$, the spread fingerprint signal $\boldsymbol{w_i}$ is quantized. At the averaging stage, the effect of the quantization error causes a critical change in the multiplexed fingerprint sequences spread over the entire frequency domain. At the detection stage, the frequency components in which the fingerprint sequences are embedded are examined, whereas the other components that contain some energy are ignored. We evaluate the amount of detectable fingerprint energy and calculate the attenuation factor.

After the averaging collusion, c fingerprint sequences $\boldsymbol{w_i}, (1 \le i \le c)$, are multiplexed in the spatial domain and are distorted by quantization, denoted by $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$.

$$\hat{\boldsymbol{w}}_{\boldsymbol{c}} = \begin{cases} \left\lfloor \sum \frac{\boldsymbol{w_i}}{c} \right\rfloor & \text{FLOOR} \\ \left\lfloor \sum \frac{\boldsymbol{w_i}}{c} + 0.5 \right\rfloor & \text{RN} \end{cases}, \tag{18}$$

The probability $P_c(x)$ is derived from the distribution of $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$. The energy of $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$ is spread over L pixels, and at each pixel, the expected value of the spread energy is represented by $\sum x^2 P_c(x)$ using the probability.

If all operations are performed using floating-point numbers, the distribution of $\sum \boldsymbol{w_i}/c$ becomes approximately $N(0, \sigma^2/c)$ because no quantization error occurs. Thus, the distribution of the distorted sequence $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$ is composed of two elements: the mixed fingerprint sequence and the quantization error. For the sake of convenience, they are denoted by $\boldsymbol{w}_{\boldsymbol{c}}^*$ and $\boldsymbol{\epsilon}$, respectively. The purpose of this study is to estimate the energy of $\boldsymbol{w}_{\boldsymbol{c}}^*$ and, furthermore, to evaluate the individual energy of the mixed sequences. At the detector side, however, the number of sequences involved in $\boldsymbol{w}_{\boldsymbol{c}}^*$ is unknown, and hence, we use the method of undetermined coefficients.

Let c^* be the undetermined coefficient. Then, the energy of $\boldsymbol{w}_{\boldsymbol{c}}^*$ is β^2/c^*. Let $N(\hat{\mu}_c, \hat{\sigma}_c^2)$ be the distribution of $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$ that is obtained from the difference between D and $\hat{D}_c$. In a manner similar to Eq.(9) and Eq.(10), the mean $\hat{\mu}_c$ and variance $\hat{\sigma}_c^2$ are represented as follows:

$$\hat{\mu}_c = \sum (x - \hat{\mu}) P_c(x) \tag{19}$$

and

$$\hat{\sigma}_c^2 = \sum (x - \hat{\mu})^2 P_c(x), \tag{20}$$

It is observed that the mean value $\hat{\mu}_c$ is a DC component of $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$, and hence, this component is included in the quantization error of $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$. Considering the fact that the energy of noise ϵ is very small, we assume that the variance is given as $\hat{\sigma}_c^2 = \sigma^2/c^*$. We aim to determine the parameter c^* that minimizes $(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2$.

Now, we assume that the variance σ^2/c^* is very small and that the probability $P_c(x)$ for $|x| > \tilde{x} (\neq 0)$ is negligible. The mean $\hat{\mu}_c$ can be easily calculated by comparing the suspicious copy with the original one. Under the assumption, the following equation is derived:

$$\sum_{x=-\tilde{x}}^{\tilde{x}} P_c(x) = 1. \tag{21}$$

First, we derive the parameters t_x such that the integral value within the range $[t_x, t_{x+1})$ of $N(\hat{\mu}_c, \sigma^2/c^*)$ corresponds to $P_c(x)$.

$$P_c(x) = \int_{t_x}^{t_{x+1}} f\left(t, \hat{\mu}_c, \frac{\sigma^2}{c^*}\right) dt \tag{22}$$

and

$$P_c(-\tilde{x}) = \int_{-\infty}^{t_{-\tilde{x}}} f\left(t, \hat{\mu}_c, \frac{\sigma^2}{c^*}\right) dt. \tag{23}$$

Hence, the parameters t_x are sequentially calculated from $t_{-\tilde{x}}$ to $t_{\tilde{x}}$. Considering the average energy spread over L pixels, $(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2$ is calculated using parameters t_x as follows:

$$\begin{aligned}(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2 &= \hat{\boldsymbol{w}}_{\boldsymbol{c}}{}^2 - 2\hat{\boldsymbol{w}}_{\boldsymbol{c}}\boldsymbol{w}_{\boldsymbol{c}}^* + \boldsymbol{w}_{\boldsymbol{c}}^{*2} \\ &= L\sum_{x=-\tilde{x}}^{\tilde{x}} (x-\hat{\mu}_c)^2 P_c(x) \\ &\quad -2L\sum_{x=-\tilde{x}}^{\tilde{x}} (x-\hat{\mu}_c) g(x) + \frac{\beta^2}{c^*}, \end{aligned} \tag{24}$$

where

$$g(x) = \int_{t_x}^{t_{x+1}} t \cdot f\left(t, \hat{\mu}_c, \frac{\sigma^2}{c^*}\right) dt. \tag{25}$$

Using the integral formula

$$\int x e^{-x^2} dx = -\frac{1}{2} e^{-x^2}, \tag{26}$$

Eq.(25) is rewritten as follows:

$$\begin{aligned} g(x) &= \sqrt{\frac{c^*}{2\pi\sigma^2}} \int_{t_x}^{t_{x+1}} t \cdot \exp\left\{ -\frac{c^*(t-\hat{\mu}_c)^2}{2\sigma^2} \right\} dt \\ &= \frac{\sigma}{\sqrt{2\pi c^*}} \left\{ \exp\left(-\frac{c^*}{2\sigma^2}(t_x - \hat{\mu}_c)^2 \right) \right. \\ &\quad \left. - \exp\left(-\frac{c^*}{2\sigma^2}(t_{x+1} - \hat{\mu}_c)^2 \right) \right\} \end{aligned} \tag{27}$$

By varying c^*, the above operations are performed until $(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2$ is at its minimum. Although the attenuation factor of the energy is c^*, that of each fingerprint sequence is $c' = \sqrt{c^* \cdot c}$, because it is expected that the total energy β^2/c^* is equally separated into c fingerprint sequences by averaging. When colluders select the same rounding operation for averaging, the attenuation factor c' becomes larger than c; otherwise, it becomes smaller.

As an example, we consider the best-case scenario from the detector's perspective, in which RN and FLOOR are performed during the embedding and averaging stages, respectively. As mentioned in Sect.III-B1, the approximation given by Eq.(11) becomes valid with the increase in c; in an extreme case, the multiplexed c fingerprint signals at the pixels are rounded to only two values, -1 and 0, with an equal probability, i.e., $P_c(-1) = P_c(0) = 1/2$. Therefore, the distribution of $\boldsymbol{w}_{\boldsymbol{c}}^*$ becomes $N(0, \sigma^2/c^*)$, and the ranges of integral for $P_c(-1)$ and $P_c(0)$ can be considered as $[-\infty, 0)$ and $[0, \infty]$, respectively. Using these ranges, we can calculate $g(-1)$ and $g(0)$ as follows:

$$g(-1) = \int_{-\infty}^{0} t \cdot f\left(t, 0, \frac{\sigma^2}{c^*}\right) dt = \frac{-\sigma}{\sqrt{2\pi c^*}} \tag{28}$$

and

$$g(0) = \int_{0}^{\infty} t \cdot f\left(t, 0, \frac{\sigma^2}{c^*}\right) dt = \frac{\sigma}{\sqrt{2\pi c^*}}. \tag{29}$$

By substituting these expressions into Eq.(24), we obtain

$$(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2 = \hat{\boldsymbol{w}}_{\boldsymbol{c}}^{\boldsymbol{2}} - \frac{2\sigma L}{\sqrt{2\pi c^*}} + \frac{\beta^2}{c^*}. \tag{30}$$

When $(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2$ is differentiated with respect to c^*, we obtain

$$\left\{(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2\right\}' = \frac{\sigma L}{\sqrt{2\pi}(c^*)^{\frac{3}{2}}} - \frac{\beta^2}{(c^*)^2}. \tag{31}$$

Let us reiterate that $c^* \neq 0$ and $\sigma^2 = \beta^2/L$. Because $(\hat{\boldsymbol{w}}_{\boldsymbol{c}} - \boldsymbol{w}_{\boldsymbol{c}}^*)^2$ becomes minimum when the differential value is zero, the optimal c^* is calculated by

$$c^* = \frac{2\pi\beta^2}{L} \tag{32}$$

and the corresponding attenuation factor c' is given by

$$c' = \sqrt{\frac{2\pi\beta^2 c}{L}}. \tag{33}$$

It should be noted that the energy of $\boldsymbol{w}_c^*$ approaches a constant value $\beta^2/c^* = L/2\pi \simeq 0.1592L$ with the increase in c. From Eq.(13), the variance of $\hat{\boldsymbol{w}}_{\boldsymbol{c}}$ is $\hat{\sigma}_c^2 = 0.25$ and its energy is $0.25L$. The energy of noise $\boldsymbol{\epsilon}$ is, therefore, $(0.25 - 1/2\pi)L$. As a result, the detectable energy corresponding to the fingerprint signals is approximately 63.7% of the remaining signals in this case.

E. Consideration

From the above analysis, the attenuation factor c' increases when the same rounding operation is performed during both the embedding and the averaging stages. If different rounding operations are performed, c' decreases. During the embedding of the fingerprint sequence, one of the two rounding operations, RN or FLOOR, is performed when the frequency components are transformed into the spatial domain. Similarly, one of these operations is selected by colluders when executing the averaging attack. Because the rounding operation is generally fixed in the entire process, there are only four possible combinations for combination of the rounding operations: "RN-RN," "RN-FLOOR," "FLOOR-RN," and "FLOOR-FLOOR."

In such a case, the best strategy for the colluders is to predict the rounding operation performed during the embedding stage and to perform the same operation during the averaging stage.

From the detector's perspective, the worst-case scenario is that the colluders successfully predict the rounding operation performed during the embedding stage. On the other hand, the best-case scenario is the failure of the prediction. In other words, the best-case scenario is the "RN-FLOOR" or "FLOOR-RN" rounding operation combination, and the worst-case scenario is the "RN-RN" or "FLOOR-FLOOR" rounding operation combination..

Although the above analysis is only carried out for two types of rounding operations, RN and FLOOR, we can choose other operations by selecting δ under the assumption that our programming language provides the TRUNCATE operation, as mentioned in Sect.III-A. Without loss of generality, we select δ_1 and δ_2, where $\delta_1 < \delta_2$ in the range $[0, 1)$, to satisfy the following condition:

$$|\delta_1 - \delta_2| = 0.5. \tag{34}$$

Here, we denote the two rounding operations by R_{δ_1} and R_{δ_2}, respectively. Hence, the best-case scenario from the detector's perspective is derived by selecting "R_{δ_1}-R_{δ_2}" or "R_{δ_2}-R_{δ_1}" and the worst-case scenario, by selecting "R_{δ_1}-R_{δ_1}" or "R_{δ_2}-R_{δ_2}." The reasons for these selections are the same as those for the selections when $\delta_1 = 1$ and $\delta_2 = 0.5$, and therefore, they are omitted.

The difference $|\delta_1 - \delta_2|$ has a value in the range $[0, 1]$ because the step size for rounding a decimal number is 1. Because of the symmetric property, we can consider the difference within the range $[0, 0.5]$ without loss of generality. When the difference is 0 and 0.5, the attenuation factor c' becomes maximum and minimum, respectively. If it is in the range $(0, 0.5)$, c' varies between the maximum and minimum values. We can calculate this value by an analysis similar to that described in the above. Hereafter, we refer to the best-case and worst-case scenario from only the detector's perspective.

F. Numerical Comparison

We use a standard "lena" image with a 256-level grayscale and a size of 512×512 pixels. To simplify the analysis, we directly add the fingerprint sequence $\boldsymbol{W_{i,j}}$ to the DCT coefficients selected randomly from low- and middle-frequency components without any weighting method that would utilize the perceptual feature of an image. The energy of the fingerprint sequence is fixed as $\beta^2 = 520000 (\beta_g = 400, \beta_u = 600)$ and the length is $\ell = 8192$. The attenuation factor c' is calculated from the extracted energy of the fingerprint sequence by subtracting the DCT coefficients of the averaged image from that of original one, and the values are averaged for 100 trials by randomly selecting combination of fingerprint sequences.

We use the peak signal-to-noise ratio (PSNR) for evaluating the image quality. The PSNR of the 256-level grayscale is expressed as follows:

$$\mathrm{PSNR} = 10 \log \frac{255^2}{\frac{1}{L} \sum \left\{ (\Delta D)^2 - (\overline{\Delta D})^2 \right\}}, \tag{35}$$

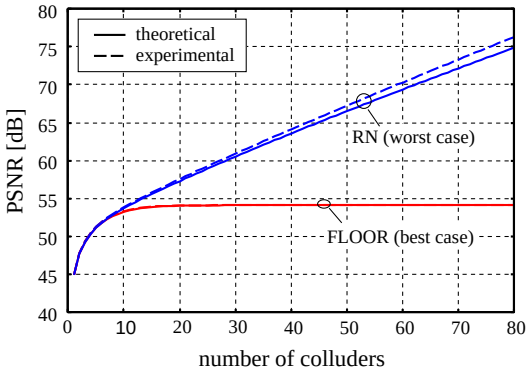


Fig. 1. Comparison of PSNR when RN is performed during embedding.

where ΔD is the difference between an original content D and the averaged copy $\hat{D}_c$. Figure 1 shows the results for the best and worst cases. Although the PSNR value for the best-case scenario approaches a constant value, that for the worst-case scenario increases in proportion to the number of colluders. This observation confirms that the fingerprint signal is maintained in the best-case scenario even if the number of colluders is large and that it is attenuated in the worst-case scenario.

The attenuation factor c' is derived both theoretically and experimentally in the case in which RN is performed during the embedding stage; these results are shown in Fig.2. The conventional attenuation factor is indicated by the line $c' = c$. The theoretical value is very similar to the experimental value; hence, the validity of our analysis is confirmed. It is noteworthy that c' is considerably smaller than c for the best case. On the other hand, we can see that fingerprint sequences are degraded much more for the worst case. Contrary to the result of the conventional analysis, our analysis and experimental result reveal that the attenuation of the embedded fingerprint strength is affected by the quantization method.

When the FLOOR operation is performed during the embedding stage, the results are similar to those derived in the case in which RN is performed. If the same rounding operation is performed during both the embedding and the averaging stages, the attenuation factor c' increases considerably; otherwise, it decreases considerably.

Because the energy of the fingerprint signal in Cox's scheme depends on the characteristics of the image, it is not constant. Therefore, we have omitted the theoretical analysis of the performance in this paper. Only the average values of the experimental results are shown in Fig.3; These values are obtained by changing the embedding strength α using a length $\ell = 1000$. When $\alpha = 0.1$, the PSNR values of the best and worst cases are very similar and continue to increase monotonically. We can observe that the difference between the PSNR values increases with the decrease in α. Because the variance of the fingerprint signal is large when $\alpha = 0.1$, the distortions caused by the rounding operations do not affect the performance.

It is interesting to note that the PSNR stops increasing when its value approaches 54.15 [dB] in both Fig.1 and Fig.3. This

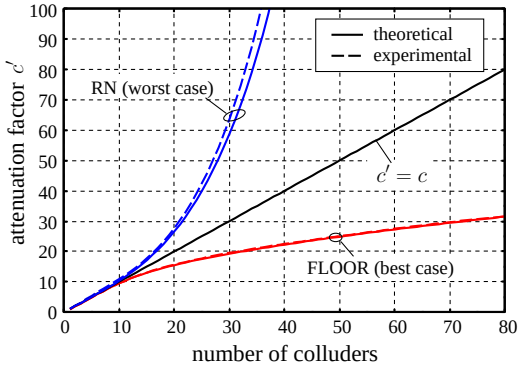


Fig. 2. Comparison of attenuation factor c' when RN is performed during embedding.

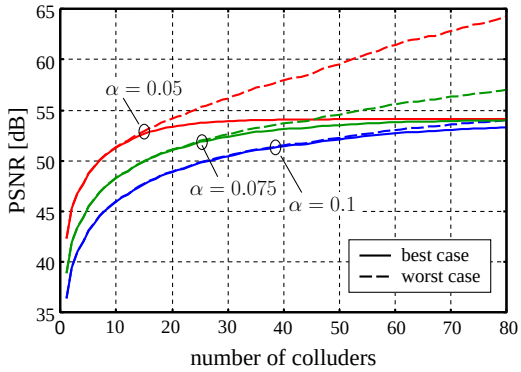


Fig. 3. Comparison of PSNR when RN is performed during embedding using Cox's method.

is because the variance in the fingerprint signal becomes 0.25 by the averaging attack, as shown in Eq.(13). Therefore, the PSNR value is calculated as follows:

$$10 \log \left(\frac{255^2}{0.25} \right) \simeq 54.15, \tag{36}$$

Therefore, the upper limit of the PSNR in the best-case scenario is 54.15 [dB].

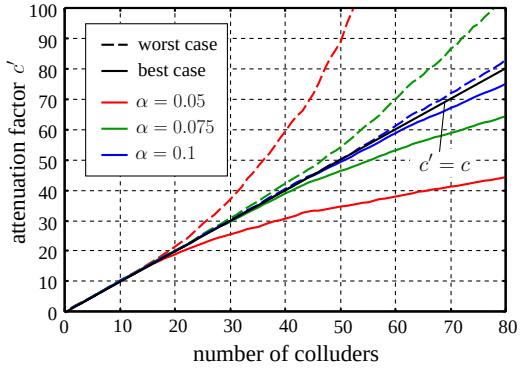


Fig. 4. Comparison of attenuation factor c' when RN is performed during embedding using Cox's method.

IV. Attack Strategy

The worst-case scenario, in which the colluders successfully predict the rounding operation, has a 1/2 probability of occurring. In this section, we show that the colluders can identify the rounding operation employed during the embedding stage.

In a fingerprinting system, each colluder's copy contains its distinctive fingerprint signal and the original content is not revealed. Therefore, the colluders compare their copies and analyze the difference in order to remove/modify the fingerprint signal. Suppose that the number of colluders is c and their copies are D_i, where $i = 1, 2, \ldots, c$. After the averaging attack, the colluders perform a rounding operation using a certain δ. The statistical observation of the fingerprint signals during both the embedding and the averaging stages, as presented in Sect.III, implicitly states the following properties. If δ selected during the averaging stage is the same as that used during the embedding stage, the mean square error (MSE) between the averaged copy and the original one is at its minimum. On the contrary, if the difference between the δ value during the embedding and the averaging stages is equal to ± 0.5, the MSE is at its maximum. Without the original copy, colluders can determine the pair δ_1 and $\delta_2 (= \delta_1 + 0.5)$ in Eq.(34) by calculating $\max_{\delta} \{MSE(c, \delta, \delta + 0.5)\}$, where

$$MSE(c, \delta, \delta + 0.5) = (\hat{D}_{c,\delta} - \hat{D}_{c,\delta+0.5})^2 \tag{37}$$

and $\hat{D}_{c,\delta}$ is produced by a rounding operation R_δ after averaging c copies.

The next step is to identify the rounding operation performed during the embedding stage from the two obtained parameters δ_1 and δ_2. As mentioned in Sect.III, the strength of the fingerprint signals is attenuated by increasing c, and the level of attenuation for the best-case scenario is considerably lower than that for the worst-case scenario. This also indicates that the MSE between the two averaged copies produced from c and $\tilde{c} (< c)$ copies during the best-case scenario is smaller than that in the worst-case scenario. For $\tilde{c}$ randomly selected copies, two averaged copies, denoted by $\hat{D}_{\tilde{c},\delta_1}$ and $\hat{D}_{\tilde{c},\delta_2}$, are produced using two types of rounding operations determined by δ_1 and δ_2. If the difference in MSE, D_{MSE}, is

$$\begin{aligned} D_{MSE} &= \tilde{MSE}(c, \delta_1) - \tilde{MSE}(c, \delta_2) \\ &= (\hat{D}_{c,\delta_1} - \hat{D}_{c',\delta_1})^2 - (\hat{D}_{c,\delta_2} - \hat{D}_{c',\delta_2})^2 \\ &= > 0, \end{aligned} \tag{38}$$

then δ_1 is the rounding parameter for the worst case and δ_2, that for the best case. On the contrary, if $D_{MSE} < 0$, δ_1 and δ_2 are the parameters for the best and worst cases, respectively. When $D_{MSE} = 0$, then classification is not possible. Therefore, for determining an appropriate $\tilde{c}$ value, colluders can uniquely determine the rounding operation performed during the embedding stage; hence, we must evaluate the worst case of the collusion resistance.

V. Countermeasure

One of the drawbacks of the rounding operation performed during the embedding stage is the fixed rounding parameter δ. This information enables colluders to identify δ by observing

the statistical differences between the pixel values of their copies. In this section, we study the effect on the attenuation of fingerprint signals and the performance of detector when δ is randomly selected at each pixel.

A. Random Selection of FLOOR and RN During Embedding

One of the approaches that can be used for increasing the difficulty of the analysis is to randomly select δ at each pixel using a secret key when a fingerprint is embedded. For example, if the FLOOR and RN operations are randomly selected, colluders are expected to predict the rounding operation at $L/2$ pixels. Therefore, the attenuation factor c' fluctuates between the best- and the worst-case scenario. Usually, the factor approaches a stable value with the increase in L; this factor is smaller than that of the worst-case scenario.

If the RN operation ($\delta = 0.5$) is performed during the embedding stage, the mean value of the quantized signal becomes $\hat{\mu} = 0$, as mentioned in Sect.III-B. If the FLOOR operation ($\delta = 0$) is performed, then $\hat{\mu} = -0.5$. This indicates that $\hat{\mu}$ of the quantized signal depends on δ, and this is formulated as follows:

$$\hat{\mu} = \delta - 0.5. \tag{39}$$

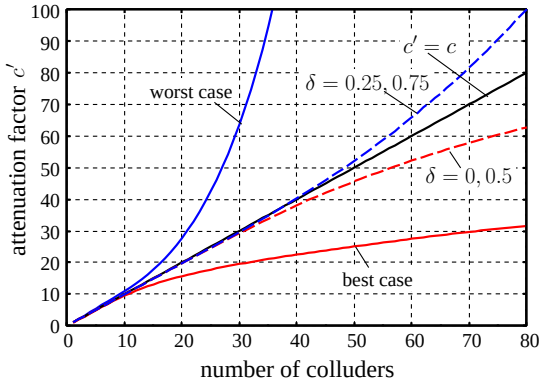
When $\delta = 0$ and 0.5 is selected randomly, the probability $P_c(x)$ that the fingerprint signal is rounded to a value $x \in \mathbb{Z}$ at each pixel is represented by

$$P_c(x) = \frac{1}{2}\int_{x-\delta}^{x+\delta} \left\{ f\left(t, 0, \frac{\hat{\sigma}^2}{c}\right) + f\left(t, -0.5, \frac{\hat{\sigma}^2}{c}\right) \right\} dt, \tag{40}$$

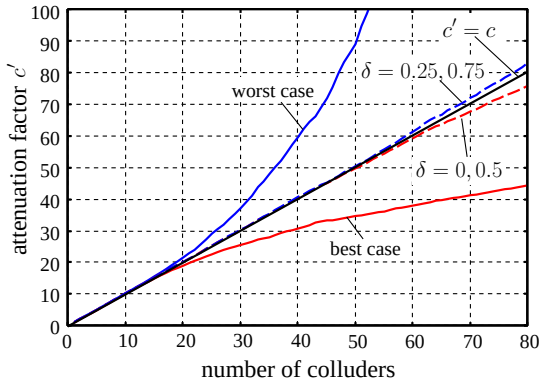
because it is a combination of two types of distributions: $N(0, \hat{\sigma}^2/c)$ and $N(-0.5, \hat{\sigma}^2/c)$. When colluders select $\delta = 0$ during the averaging stage, the first term of the integration in Eq.(40) takes the maximum value as the best case and the second term takes the minimum value as the worst case. On averaging, these values correspond to the best case for half of the pixels $L/2$. The same phenomenon is observed for $\delta = 0.5$. From the analysis presented in Sect.III, the best case is derived when the rounding parameter δ at the averaging stage is $\delta_e + 0.5$, where δ_e is the parameter at the embedding stage. On the other hand, when $\delta = 0.25$, both terms have an intermediate value that decreases with an increase in c. In this case, the attenuation of the fingerprint signals is at its maximum. Considering the symmetric shape of the distribution, the same phenomenon is observed for $\delta = 0.75$. Unlike the situation involving a fixed rounding operation, the best case in this situation is when colluders select $\delta = 0$ or 0.5, and the worst case, when they select $\delta = 0.25$ or 0.75.

The numerical comparisons are shown in Fig.5 using the same parameters as those used in Sect.III-F. Although the attenuation factor c' in the best case is higher than that in the fixed rounding operation, it is still lower than c. On the other hand, the worst case can be improved by randomly selecting the FLOOR and RN operations. The effectiveness of the random selection is confirmed by the numerical results.

It is interesting to note that the condition given in Eq.(34), $|\delta_1 - \delta_2|$, is not 0.5 but 0.25. Accordingly, the attack strategy



(a) CDMA-based scheme



(b) Cox's scheme ($\alpha = 0.05$)

Fig. 5. Comparison of attenuation factor c' with random selection of FLOOR and RN during embedding.

mentioned in Sect.IV also changes. The relationship between the pair δ_1 and δ_2 becomes $\delta_2 = \delta_1 + 0.25$, and colluders try to determine the pair by calculating $\max_{\delta}\{MSE(c, \delta, \delta + 0.25)\}$. Even if the above-mentioned minor modification occurs, colluders can still obtain the rounding parameters δ_1 and δ_2 from statistical observation.

B. Random Rounding Operation

We consider the attenuation factor c' for the case in which the rounding operation performed during the embedding stage is randomly and independently changed for pixels by selecting the rounding parameter δ in the range [0,1).

In the above discussion, the rounding operations available for selection are FLOOR and RN, i.e., $\delta = 0$ and 0.5; these operations are selected randomly with a probability 1/2. First, the number of available rounding operations is expanded to 3, i.e., $\delta = 0, 1/3$, and $2/3$, and the probability of selection is 1/3. Next, the probability $P_c(x)$ is represented by

$$P_c(x) = \frac{1}{3}\int_{x-\delta}^{x+\delta} \left\{ f\left(t, -\frac{1}{2}, \frac{\hat{\sigma}^2}{c}\right) + f\left(t, -\frac{1}{6}, \frac{\hat{\sigma}^2}{c}\right) + f\left(t, \frac{1}{6}, \frac{\hat{\sigma}^2}{c}\right) \right\} dt, \tag{41}$$

where the three terms in the integration are three distributions: $N(-1/2, \hat{\sigma}^2/c)$, $N(-1/6, \hat{\sigma}^2/c)$, and $N(1/6, \hat{\sigma}^2/c)$. As in the

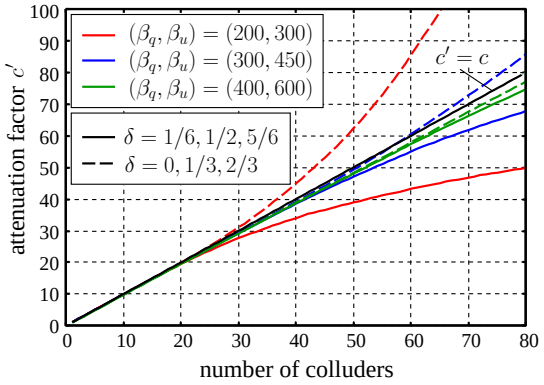


Fig. 6. Comparison of attenuation factor c' using three randomly selected rounding parameters during averaging. The solid lines indicate for the maximum attenuation factors and the dashed lines, the minimum ones.

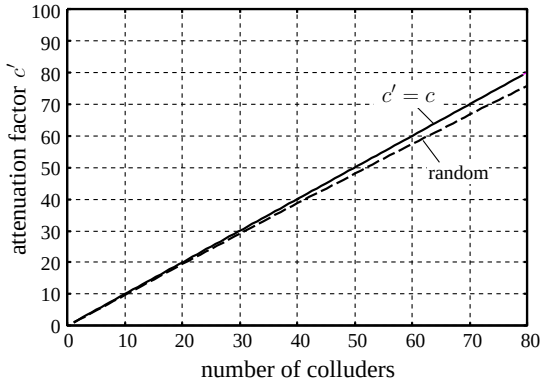


Fig. 7. Comparison of attenuation factor c' using random rounding operations, where "random" uses 100 candidates of δ in the range [0,1).

case involving the FLOOR and RN operations, one of these three terms is at its maximum value when $\delta = 1/2, 5/6$, and $1/6$ $(= 7/6 - 1)$ during the averaging stage because $\delta_e = 0, 1/3, 2/3$; therefore, $P_c(x)$ attains the maximum value. On averaging, the best case is observed at $L/3$ pixels. On the other hand, if the intermediate values $\delta = 0, 1/3$, and $2/3$ are applied, $P_c(x)$ will attain its minimum value. To confirm the analysis, we implemented a method involving the CDMA-based fingerprinting scheme and evaluated the attenuation factor for the maximum and minimum cases. The experimental results are plotted in Fig.6, which shows a change in the amount of energy to be embedded into an image. When the amount of energy $\beta^2 = \beta_g^2 + \beta_u^2$ is large $(\beta_g = 400, \beta_u = 600)$, there is no remarkable difference between the two cases. However, this difference increases with a decrease in the energy β^2.

From the above results, it is expected that the number of candidates for the rounding operation is increased, and the attenuation factor c' approaches c. If the number approaches infinity, the probability $P_c(x)$ is represented by

$$P_c(x) = \lim_{n\to\infty} \left\{ \frac{1}{n} \int_{x-\delta}^{x+\delta} \sum_{i=0}^{n-1} \left\{ f\left(t, \frac{i}{n} - \frac{1}{2}, \frac{\hat{\sigma}^2}{c}\right) \right\} dt \right\}. \quad (42)$$

In such a case, the rounding parameter during the averaging stage coincides with δ_e, and hence, the difference in the attenuation factor becomes negligible. Figure 7 shows the experimental results of the attenuation factor c' using 100 candidates of δ in the range [0,1). The number of rounding parameters depends on the energy of the fingerprint signal to be embedded into an image. When the CDMA-based scheme is applied for embedding a fingerprint using the parameters $\beta_g = 400$ and $\beta_u = 600$, the recommended number is obtained as 3 from the above results. We reiterate that these results are expected values and that the attenuation factor oscillates between the best and worst cases. Without knowledge of the selection of the rounding operation for the embedding stage, the expected attenuation factor is obtained from the above analysis. When the number of candidates for δ is properly selected and randomly applied for the pixels during the embedding stage, it is difficult for colluders to attenuate the embedded fingerprint signal to more than $1/c$.

The above random rounding operation is analogous to the dithered operation described in [6] that performs randomized quantization in the frequency domain to improve the robustness against the collusion attack. The objective of that operation is to estimate the probability of successfully catching one colluder, whereas our objective is to estimate the attenuation factor of the embedded fingerprint with respect to the number of colluders. Furthermore, the important contribution of this paper is to clarify the attenuation factor with respect to the rounding operations performed during both the embedding and the averaging stages by using the statistical distribution of the fingerprint signal in the spatial domain. Our results demonstrate the property of the spread spectrum fingerprinting scheme that the rounding operation performed during embedding must be unpredictable for colluders when they generate a pirated copy.

Although we discuss the attenuation factor under the averaging attack, there are several nonlinear types of collusion attacks such as Median, Minimum, Maximum, and Min-Max attacks [5]. Because of the nonlinearity of these attacks, it is difficult to carry out a theoretical analysis based on the PDF of the fingerprint signal in the spatial domain. However, the experimental results reported in [6] (Fig.6) suggest that the distribution of the colluded fingerprint after the attacks can be regarded as a discrete version of Gaussian distribution. Thus, a similar analysis for such attacks is possible and it is expected to obtain a similar attenuation factor. Further processing methods such as filtering and addition of noise may be applied to a pirated copy; these methods are occasionally modeled as additive Gaussian noise. Considering the independent and identically distributed signals of both the noise and the fingerprint signals, the interference of noise can be discussed separately to estimate the performance of the detector. Moreover, as reported in [6], even if the noise is quantized, there is no significant difference in the result. Therefore, it is also expected that the attenuation factor is dependent on the amount of noise added to a pirated copy.

VI. Conclusions

In this paper, we analyzed the effects of the rounding error introduced during the embedding and averaging stages in the spread spectrum fingerprinting scheme. Considering the characteristics of the RN and FLOOR operations, the truncation of several decimals was conducted and analyzed for each averaged pixel value. We classified the strategy of the rounding functions into four cases and measured the attenuation factor c'. Our simulation revealed that the best case from the detector's perspective can be derived from the selection of different rounding operations during the embedding and averaging stages and the worst case, from the selection of the same operation. Unfortunately, colluders can increase the probability of selecting the worst case by comparing their copies. As a countermeasure, we randomly selected two types of rounding operations during the embedding stage. Although the collusion resistance is degraded relative to that in the best case, it is stable for any type of strategy employed by the colluders. Furthermore, the value of the attenuation factor c' approaches that of c with an increase in the number of candidates for the randomly selected rounding parameter δ.

Our analysis of the attenuation factor is conducted under the assumption that all colluders share the same risk during collusion and that no noise is added to the averaged copy. In future works, we aim to provide a detailed analysis of the attenuation factor in cases in which selfish colluders are involved and where the pirated copy is distorted by the addition of noise.

Acknowledgment

This research was partially supported by the Ministry of Education, Culture, Sports Science and Technology, Grant-in-Aid for Young Scientists (B) (21760291), 2010.

References

[1] M. Wu, W. Trappe, Z. J. Wang, and K. J. R. Liu, "Collusion resistant fingerprinting for multimedia," *IEEE Signal Processing Mag.*, pp. 15–27, 2004.

[2] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamson, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, 1997.

[3] H. Stone, "Analysis of attacks on image watermarks with randomized coefficients," *NEC Res. Inst., Tech. Rep.*, vol. 96–045, 1996.

[4] J. Kilian, T. Leighton, L. Matheson, T. Shamoon, R. Trajan, and F. Zane, "Resistance of digital watermarks to collusive attacks," in *Proc. IEEE Int. Symp. Infomation Theory*, 1998, p. 271.

[5] H. V. Zhao, M. Wu, Z. J. Wang, and K. J. R. Liu, "Forensic analysis of nonlinear collusion attacks for multimedia fingerprinting," *IEEE Trans. Image Process.*, vol. 14, no. 5, pp. 646–661, 2005.

[6] A. L. Varna, S. He, A. Swaminathan, and M. Wu, "Fingerprinting compressed multimedia signals," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 3, pp. 330–345, 2009.

[7] Z. J. Wang, M. Wu, H. V. Zhao, W. Trappe, and K. J. R. Liu, "Anti-collusion forensics of multimedia fingerprinting using orthogonal modulation," *IEEE Trans. Image Process.*, vol. 14, no. 6, pp. 804–821, 2005.

[8] Y. Zhu, D. Feng, and W. Zou, "Collusion secure convolutional spread spectrum fingerprinting," in *IWDW 2005*. 2005, vol. 3710 of *LNCS*, pp. 67–83, Springer, Heidelberg.

[9] Z. J. Wang, M. Wu, W. Trappe, and K. J. R. Liu, "Group-oriented fingerprinting for multimedia forensics," *EURASIP J. Appl. Signal Process.*, vol. 14, pp. 2142–2162, 2004.

[10] S. He and M. Wu, "Joint coding and embedding techniques for multimedia fingerprinting," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 231–247, 2006.

[11] N. Hayashi, M. Kuribayashi, and M. Morii, "Collusion-resistant fingerprinting scheme based on the CDMA-technique," in *IWSEC2007*. 2007, vol. 4752 of *LNCS*, pp. 28–43, Springer, Heidelberg.

[12] Y. T. Lin, J. L. Wu, and C. H. Huang, "Concatenated construction of traceability codes for multimedia fingerprinting," *Optical Engineering*, vol. 46, no. 10, pp. 107202.1–107202.15, 2007.

[13] R. Gold, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Trans. Infom. Theory*, vol. 14, no. 1, pp. 154–156, 1968.

[14] F. Ergun, J. Kilian, and R. Kumar, "A note on the limits of collusion-resistant watermarks," in *Advances in Cryptology — EUROCRYPT'99*. 1999, vol. 1592 of *LNCS*, pp. 140–149, Springer, Heidelberg.

[15] H. V. Zhao and K. J. R. Liu, "Behavior forensics for scalable multiuser collusion: Fairness and effectiveness," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 3, pp. 311–329, 2006.

[16] H. V. Zhao and K. J. R. Liu, "Traitor-within-traitor behavior forensics: Strategy and risk minimization," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 4, pp. 440–456, 2006.

[17] J. J. Eggers and B. Girod, "Quantization effects on digital watermarks," *Signal Process.*, vol. 81, no. 2, pp. 239–263, 2001.

Minoru Kuribayashi received the B.E., M.E., and D.E degrees from Kobe University, Kobe, Japan, in 1999, 2001, and 2004 respectively. From 2002 to 2007, he was a Research Associate in the Department of Electrical and Electronic Engineering, Kobe University. Since 2007, he has been an Assistant Professor at Division of Electrical and Electronic Engineering, Kobe University. His research interests are in digital watermark, information security, cryptography, and coding theory.

Hiroshi Kato received the B.E. degree from Kobe University, Kobe, Japan, in 2008. Since 2008, he has been a student at Division of Electrical and Electronic Engineering, Kobe University. His research interests are in digital watermark, and information security.