Hindawi Publishing Corporation
EURASIP Journal on Information Security
Volume 2009, Article ID 293031, 11 pages
doi:10.1155/2009/293031

Research Article

Recoverable Privacy Protection for Video Content Distribution

Guangzhen Li,[1] Yoshimichi Ito,[1] Xiaoyi Yu,[2] Naoko Nitta,[1] and Noboru Babaguchi[1]

[1] *Division of Electrical, Electronic and Information Engineering, Graduate School of Engineering, Osaka University, 2-1, Yamada-oka Suita, Osaka 565-0871, Japan*
[2] *School of Software and Microelectronics, Peking University, No. 24, Jinyuan Industry Development Zone, Daxing District, Beijing 102600, China*

Correspondence should be addressed to Yoshimichi Ito, ito@comm.eng.osaka-u.ac.jp

Received 16 April 2009; Revised 14 October 2009; Accepted 26 November 2009

Recommended by Andrew Senior

This paper presents a method which attains recoverable privacy protection for video content distribution. The method is based on discrete wavelet transform (DWT), which generates scaling coefficients and wavelet coefficients. In our method, scaling coefficients, which can be regarded as a low-resolution image of an original image, are used for producing privacy-protected image. On the other hand, wavelet coefficients, which can be regarded as privacy information, are embedded into the privacy-protected image via information hiding technique. Therefore, privacy protected image can be recovered by authorized viewers if necessary. The proposed method is fully analyzed through experiments from the viewpoints of the amount of the embedded privacy information, the deterioration due to the embedding, and the computational time.

1. Introduction

Recently, video surveillance has received a lot of attention as a useful technology for crime deterrence and investigations and has been widely deployed in many circumstances such as airports, convenience stores, and banks. Video surveillance allows us to remotely monitor a live or recorded video feed which often includes objects such as people. Although video surveillance contributes to realizing a secure and safe community, it also exposes the privacy of the object in the video.

Over the past few years, a lot of techniques on privacy protection in video surveillance system have been proposed [1–7]. Newton et al. [1] proposed an algorithm to protect the privacy of the individuals in video surveillance data by deidentifying faces. Kitahara et al. [2] proposed a video capturing system called Stealth Vision, which protects the privacy of the objects by blurring or pixelizing their images. In [3], Wickramasuriya et al. protect object's privacy based on the authority of either object or viewers. In [4], Boyle et al. considered face obscuring for privacy protection and discussed the effects of blurring and pixelizing. Crowley et al. [5] proposed a method for privacy protection by replacing an socially inappropriate original image with a socially acceptable image using eigen-space coding technique. Chinomi et al. [6] proposed privacy-protected video surveillance system called PriSurv, which adaptively protects objects' privacy based on their privacy policies which are determined according to closeness between objects and viewers.

Although these techniques fulfill some requirements of privacy protection, it also has a potential security flaw when privacy-protected videos produced by the above techniques are distributed on the Internet, because these techniques do not provide methods for recovering the original videos from privacy-protected videos. For example, suppose that a surveillance video camera is installed around school route, and the camera distributes a privacy-protected video on the Internet in usual case. When a crime has occurred around school route, police wants to observe the original image of a suspect in privacy-protected video. In addition, when parents want to observe the situation of their children, they require the video as they are. Thus, in order to improve the security of privacy-protected surveillance system, the privacy protection which can recover the original image from privacy-protected image is strongly required. We refer to such privacy protection as *recoverable privacy protection*.

Concerning recoverable privacy protection, several techniques have been proposed [8–11]. Dufaux and Ebrahimi [8] and Dufaux et al. [9] proposed a method based on transform domain scrambling of regions of interest in a video sequence. A pioneering work was done by Zhang et al. [10]. They proposed a method for storing original privacy information in video using information hiding technique, and it can recover the original privacy information if necessary. However, the method has the drawback that the large amount of the privacy information must be embedded to recover the original image since the privacy information is obtained from the whole information of the object regions. Even if all the privacy information could be embedded using data compression technique, it requires huge computational loads. In [11], Yu and Babaguchi proposed another method to realize recoverable privacy protection. Their method masks a real face (privacy information) with a virtual face (newly generated face for anonymity). To deal with the huge payload problem of privacy information hiding, the method uses statistical active appearance model (AAM) [12] for privacy information extraction and recovering. It is shown that the method can embed the privacy information into video without affecting its visual quality and keep its practical usefulness. However, the method requires a set of face images for training statistical AAM.

In this paper, we propose a method for recoverable privacy protection based on discrete wavelet transform (DWT). It is well known that DWT is one of the useful tools for multiresolution analysis. DWT generates scaling coefficients and wavelet coefficients. Since an image consisting of scaling coefficients can be regarded as a reduced-size image of its original, we refer to it as a low-resolution image. A low-resolution image is used for producing privacy-protected image by expanding it to the size of the original image. Using wavelet coefficients, together with a low-resolution image, one can recover its original image. Therefore, wavelet coefficients are regarded as privacy information. In order to prevent unauthorized viewers from recovering privacy-protected image, our method embeds wavelet coefficients into the privacy-protected image via information hiding technique. By this, the privacy-protected image can only be recovered by authorized viewers if necessary. Furthermore, it is shown that the amount of the privacy information of the object can significantly be reduced compared to Zhang's method [10]. In addition, in contrast with Yu's method [11], our method requires no training beforehand.

Some results of this paper have already been reported in [13], where a method for bitmap image is developed. In this paper, we extend the method so as to deal with compression technique such as JPEG [14] and JPEG2000 [15] for content distribution on the Internet and provide detailed algorithms for privacy information extracting, hiding, and recovering. Furthermore, we analyze the effectiveness of the proposed method through numerical experiments from the viewpoints of the amount of the embedded privacy information, the deterioration due to the embedding, and the computational time.

This paper is organized as follows. In Section 2, we show the architecture of the proposed system. In Section 3, the discrete wavelet transform is introduced and the new image processing method for privacy information extraction is proposed. The privacy information hiding and recovering are described in Section 4. Experimental results on the proposed method are presented in Section 5. Conclusions are made in Section 6.

2. System Architecture

Figure 1 shows the architecture of the proposed system. In the encoding procedure, the object region is extracted using adaptive Gaussian mixture model [16, 17], where the object region is defined as the least rectangular area containing human body. Then, the privacy-protected image of the object region is produced by expanding the low-resolution image obtained by DWT. Next, the privacy information of the object is extracted. In this case the privacy information of the object region is defined as the information from which the person corresponding to the region is identified, that is, a set of wavelet coefficients obtained by DWT for the region. Finally, the extracted privacy information and region information are embedded into the surveillance video using the amplitude modulo modulation-based information hiding scheme [18]. The locations and the order of the embedded pixels are described by their corresponding secret key.

When there are multiple object regions in a single image, the obscuration for privacy protection would differ according to each object region. However, in this paper, we do not deal with this issue for simplicity. We have considered such an issue in [7].

For the decoding procedure, the privacy information and region information are extracted with the procedure of the information hiding method using the secret key. Then the original image of the objects could be recovered by using the encoding process in the reversed order.

3. Privacy Information Extraction

In our system, the original image of the object region is transformed into two sets of data: a low-resolution image and a set of wavelet coefficients. This process is carried out by using discrete wavelet transform. In what follows, first, the discrete wavelet transform is introduced and then, the proposed method is described.

3.1. Discrete Wavelet Transform. The discrete wavelet transform (DWT) is computed by successive lowpass and highpass filtering of discrete signal. We use a Haar discrete wavelet transform to extract privacy information. The Haar DWT of image $\{I(m,n) \mid 0 \leq m \leq 2^M - 1, 0 \leq n \leq 2^N - 1\}$ is given by the following equations:

$$\begin{aligned}
&S^{(j+1)}(p,q)\\
&\quad = \sum_{m=0}^{(2^M/2^j)-1} \sum_{n=0}^{(2^N/2^j)-1} h(m-2p)h(n-2q)S^{(j)}(m,n),\\
&W_H^{(j+1)}(p,q)\\
&\quad = \sum_{m=0}^{(2^M/2^j)-1} \sum_{n=0}^{(2^N/2^j)-1} g(m-2p)h(n-2q)S^{(j)}(m,n),\\
&W_V^{(j+1)}(p,q)\\
&\quad = \sum_{m=0}^{(2^M/2^j)-1} \sum_{n=0}^{(2^N/2^j)-1} h(m-2p)g(n-2q)S^{(j)}(m,n),\\
&W_D^{(j+1)}(p,q)\\
&\quad = \sum_{m=0}^{(2^M/2^j)-1} \sum_{n=0}^{(2^N/2^j)-1} g(m-2p)g(n-2q)S^{(j)}(m,n),
\end{aligned} \tag{1}$$

where $S^{(0)}(m,n) = I(m,n)$, $0 \leq p \leq (2^M/2^{j+1}) - 1$, $0 \leq q \leq (2^N/2^{j+1}) - 1$. The sequences $h(n)$ and $g(n)$, which correspond to the impulse responses of lowpass and highpass filters, respectively, are defined as follows:

$$h(n) = \begin{cases} \dfrac{1}{\sqrt{2}}, & n = 0,\\ \dfrac{1}{\sqrt{2}}, & n = 1,\\ 0, & \text{otherwise} \end{cases} \qquad g(n) = \begin{cases} -\dfrac{1}{\sqrt{2}}, & n = 0,\\ \dfrac{1}{\sqrt{2}}, & n = 1,\\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

$S^{(j)}(p,q)$ is the scaling coefficients of level j, which is extracted by the lowpass filter $h(n)$. The image composed of the scaling coefficients $S^{(j)}(p,q)$ $(0 \leq p \leq 2^M/2^j - 1, 0 \leq q \leq 2^N/2^j - 1)$ is referred to as the low-resolution image of level j. $W_H^{(j)}(p,q)$, $W_V^{(j)}(p,q)$, and $W_D^{(j)}(p,q)$ are, respectively, the wavelet coefficients of level j in vertical, horizontal, and diagonal direction, which are extracted by the highpass filter $g(n)$. We refer to these wavelet coefficients as a set of wavelet coefficients of level j. Figure 2 shows the result of level 2 DWT, which is composed of a low-resolution image of level 2 and a set of wavelet coefficients from level 1 to level 2. As defined above, low-resolution image given by DWT can be regarded as a down sampling of the original image. Therefore, if we expand the low-resolution image to the size of the original image, a mosaic image could be obtained, in which the original information can be protected. If the level of DWT is large enough, the low-resolution image of the object region can protect the object's privacy, although the amount of the set of wavelet coefficients to be embedded becomes large. Figure 3 shows the results of low-resolution images, which are expanded to the original image size.

3.2. DWT-Based Privacy Information Extraction. We can extract the bounding box of the object in the surveillance video, using the background subtraction method of adaptive Gaussian mixture model [16, 17]. The bounding box is referred to as an object region. We transform the object region from (R,G,B) color space to (Y,Cb,Cr) color space where Y is the luminance component, and Cb and Cr are the blue and red chrominance components, respectively. According to the fact that human eyes are only sensitive to the luminance but not sensitive to the chrominance, the sensitive privacy information is only included in Y image. Therefore, we apply the DWT-based method presented in Section 3.1 for Y images and produce the low-resolution image Y_s and the set of wavelet coefficients Y_w. When Y image is given by $\{I(m,n) | 0 \leq m \leq 2^M - 1,\ 0 \leq n \leq 2^N - 1\}$ and level j DWT is employed, Y_s and Y_w are defined as follows:

$$\begin{aligned}
Y_s &= \begin{bmatrix} S^{(j)}(0,0) & \cdots & S^{(j)}\left(0,\overline{N}_j\right)\\ \vdots & \ddots & \vdots \\ S^{(j)}\left(\overline{M}_j,0\right) & \cdots & S^{(j)}\left(\overline{M}_j,\overline{N}_j\right) \end{bmatrix},\\
Y_w &= \left[Y_w^{(1)} \ \cdots \ Y_w^{(j)}\right],\\
Y_w^{(k)} &= \left[Y_{w,H}^{(k)} \ Y_{w,V}^{(k)} \ Y_{w,D}^{(k)}\right] \ (k = 1,\ldots,j),\\
Y_{w,X}^{(k)} &= \left[W_X^{(k)}(0,0) \ W_X^{(k)}(0,1) \ \cdots \ W_X^{(k)}\left(\overline{M}_k,\overline{N}_k\right)\right]\\
&\qquad (X \in \{H,V,D\}),
\end{aligned} \tag{3}$$

where $\overline{N}_l = 2^{(N-l)} - 1$ and $\overline{M}_l = 2^{(M-l)} - 1$. Since Y_w is a one-dimensional array consisting of $2^{M+N} - 2^{M+N-2j}$ wavelet coefficients, we also express Y_w as $Y_w = \{a_z \mid z = 1,\ldots,2^{M+N} - 2^{M+N-2j}\}$. Let Y_s' be the image obtained by expanding image Y_s to the size of image Y. The privacy-protected image is produced by replacing Y by Y_s' and transform to (R,G,B) color space. Finally, the set of wavelet coefficients Y_w and the coordinates of the most top-left and bottom right pixels of the object region are embedded by applying amplitude modulo modulation. In usual case, images are compressed before Internet transmission by JPEG and so on. For protecting the embedded data from the image compression, we embed the privacy information and region information into the frequency domain of the privacy-protected image after quantization. As compression format, JPEG and JPEG2000 are used in our method. Figure 4 shows the structure of the privacy-protected image compression and information hiding. The details of the embedding method are described in the next section.

4. Privacy Information Hiding and Recovering

Let the size of image Y be $2^M \times 2^N$ and let the level of DWT be j. Then, the encoded data sequence $E = \{e_1, e_2, \ldots\}$ which is embedded in privacy-protected image is generated by the set of wavelet coefficients $Y_w = \{a_z \mid z = 1,\ldots,2^{M+N} - 2^{M+N-2j}\}$. The process is shown in Algorithm 1.

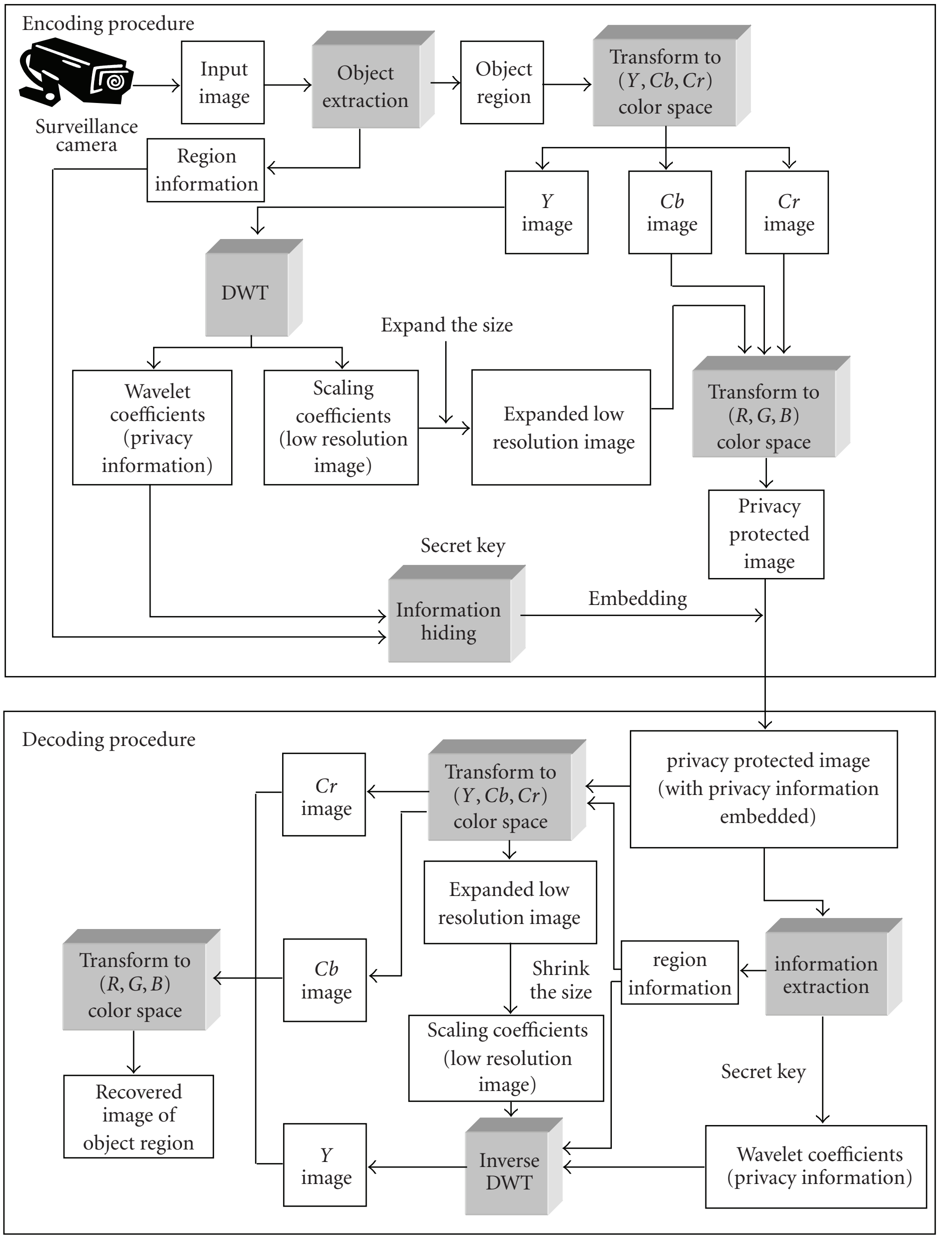

Encoding procedure
Surveillance camera
Input image
Object extraction
Object region
Transform to (Y, Cb, Cr) color space
Region information
Y image
Cb image
Cr image
DWT
Expand the size
Wavelet coefficients (privacy information)
Scaling coefficients (low resolution image)
Expanded low resolution image
Transform to (R, G, B) color space
Privacy protected image
Secret key
Information hiding
Embedding
Decoding procedure
Cr image
Transform to (Y, Cb, Cr) color space
privacy protected image (with privacy information embedded)
Expanded low resolution image
Transform to (R, G, B) color space
Cb image
Shrink the size
region information
information extraction
Scaling coefficients (low resolution image)
Secret key
Recovered image of object region
Y image
Inverse DWT
Wavelet coefficients (privacy information)

Figure 1: Schematic diagram of the system.

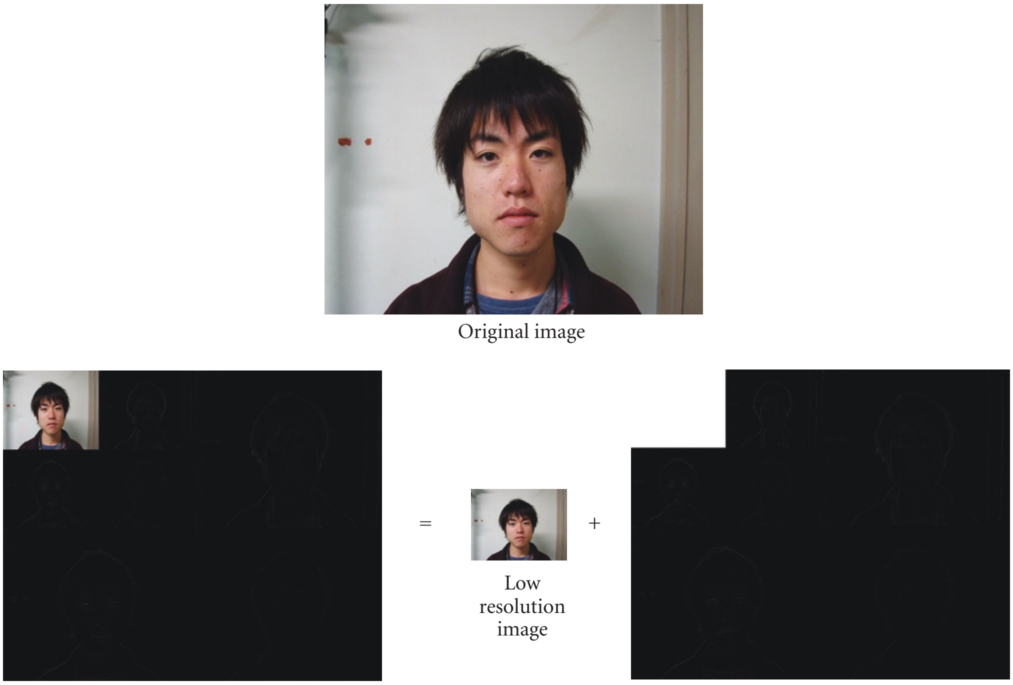
We apply run length coding for the intervals consisting of successive zeros, since, as shown in Figure 5, a histogram of wavelet coefficients of an image is distributed around 0 with small variance in general.

Next, embed the encoded data sequence $E = \{e_1, e_2, \cdots\}$ to the frequency domain of the privacy-protected image after quantization via amplitude modulo modulation (AMM) [18] according to the following equation:

$$\overline{F}_k = \underset{F \in \mathcal{S}(e_k,3)}{\arg\min} |F - F_k|, \tag{4}$$

where $\mathcal{S}(e_k, 3)$ is the set of integers given by $\mathcal{S}(e_k, 3) = \{F \mid F \equiv e_k \pmod 3\}$. And F_k and $\overline{F}_k$ are the frequency components at frequency k before and after embedding, respectively, and the frequencies for embedding are described by a secret key. Therefore, only the viewer who has the secret key can extract the embedded information from the image. The embedded color component is in the order of Cr, Cb, Y. Finally, we can obtain a compressed privacy-protected image after the entropy coding of JPEG/JPEG2000.

For the recovering procedure, the privacy information and region information are extracted by taking the congruence modulo 3 of the corresponding pixel values. Then the original image of the object is obtained by the recovering process of the Figure 1.



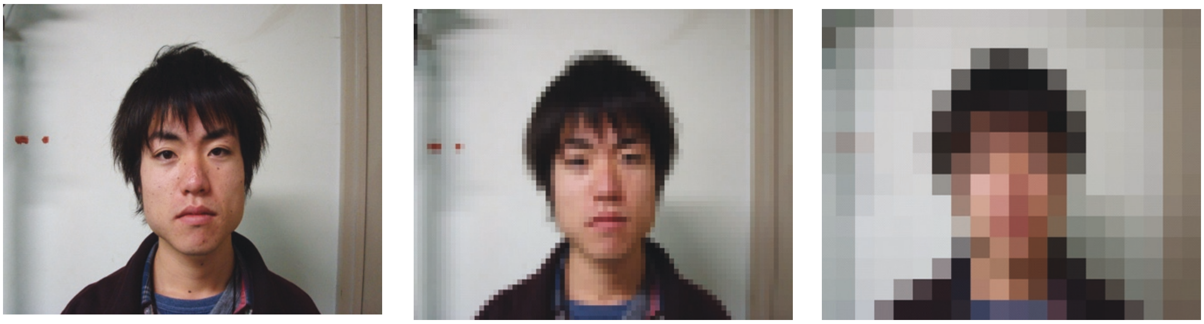
Original image

Low resolution image

Result of level 2 DWT

Set of wavelet coefficients

Figure 2: Result of level 2 DWT.



Original image

(a)

Level 3

(b)

Level 5

(c)

Figure 3: Results of low-resolution images, which are expanded to the original image size. (a) original image; (b) result of level 3 DWT; (c) result of level 5 DWT.

5. Experiments

In this section, we evaluate the performance of proposed method through several experiments. The video sequences used for the experiments are *ice* (352 × 288), *ice* (704 × 576), and *deadline* (352 × 288). In experiments, we apply JPEG for compression. In Figure 6, an original image, a privacy-protected image with privacy information embedded, and a recovered image of ice are shown in Figures 6(a) and 6(b), and 6(c), respectively. In a similar manner, corresponding images of deadline are shown in Figure 7.

For each video sequence, the average number of pixels of object regions per frame and the average number of bits for embedded data sequences per frame at different DWT levels are shown, respectively, in Tables 1 and 2.

From Tables 1 and 2, we can observe that the number of pixels of object regions and the amount of embedded data sequence of ice (704 × 576) are about four times larger than those of ice (352 × 288). This is quite natural because the resolution of ice (704 × 576) is four times larger than that of ice (352 × 288). We can also observe that the amount of embedded data sequences of deadline (352 × 288) is larger than that of ice (704 × 576), whereas the numbers of pixels of object regions of deadline (352 × 288) and ice (352 × 288) are similar.

In the following, we consider the influence of the above values on the performance of the proposed method. We employ the following three measures for performance evaluation: API, PSNR, and processing time.

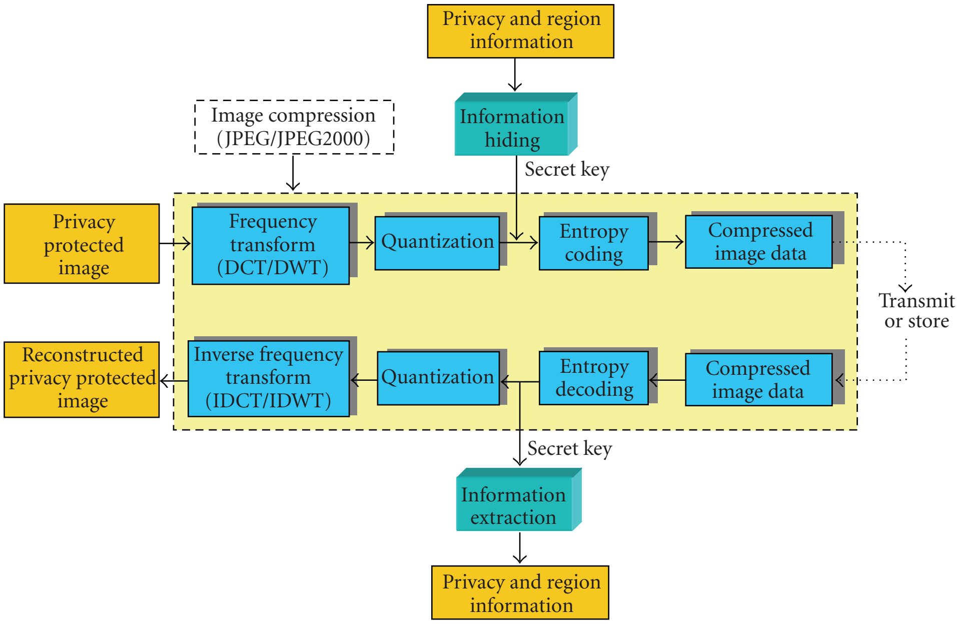

Privacy and region information
Information hiding
Image compression (JPEG/JPEG2000)
Secret key
Privacy protected image
Frequency transform (DCT/DWT)
Quantization
Entropy coding
Compressed image data
Transmit or store
Reconstructed privacy protected image
Inverse frequency transform (IDCT/IDWT)
Quantization
Entropy decoding
Compressed image data
Secret key
Information extraction
Privacy and region information

Figure 4: The structure of the privacy-protected image compression and information hiding.

API is an abbreviation of Average of Privacy Information and is defined as follows:

$$\text{API} = \frac{\text{total number of bits for embedded data sequences}}{\text{total number of pixels of object regions}} \quad (\text{bit/pixel}). \tag{5}$$

Namely, API is the average of the required bits of data sequences for recovering one pixel in the object regions and is regarded as a measure of the amount of privacy information which should be embedded. API is also calculated by using the following equation, which is equivalent to (5):

$$\text{API} = \frac{\text{average number of bits for embedded data sequences per frame}}{\text{average number of pixels of object regions per frame}}. \tag{6}$$

Therefore, API can be calculated by Tables 1 and 2.

PSNR (Peak Signal-to-Noise Ratio) is used as a measure of deterioration of recovered image and is also used for evaluating the influence of embedded data sequence on privacy-protected image. PSNR between $\tilde{K}[m,n,l]$ and$K[m,n,l]$ $(m = 0,\ldots,H-1; n = 0,\ldots,W-1; l = 0,\ldots,C-1)$ is defined as follows:

$$\text{PSNR} = 20\log\left(\frac{255}{\sqrt{\text{MSE}}}\right), \tag{7}$$

where MSE (Mean Square Error) is defined by

$$\text{MSE} = \frac{\sum_{m=0}^{H-1}\sum_{n=0}^{W-1}\sum_{l=0}^{C-1}\left|\tilde{K}[m,n,l] - K[m,n,l]\right|^2}{H \cdot W \cdot C}. \tag{8}$$

5.1. *Evaluation of the Amount of Privacy Information and the Deterioration Due to Embedding*

APIs of each video sequence for different levels of DWT under the condition $\Delta = 1$ are shown in Figure 8. From Figure 8, we can observe that API tends to be large as DWT level increases. However, API hardly increases when the level is larger than 2 and does not exceed 3 bit/pixel. Therefore, we could embed all the privacy information into three color channels Y, Cb, and Cr of the privacy-protected image, even if the level of DWT is large and the object region size is equal to the size of the whole image. On the other hand, using the method of [10], API becomes 24 bit/pixel (= 8 bit/pixel/channel ×3 channel) when the data to be embedded consist of whole information of the object regions.

Next, we consider the deterioration of the privacy-protected image due to the privacy information embedding.

Table 1: Average number of pixels of object regions per frame (pixel/frame).

	Level 1	Level 2	Level 3	Level 4	Level 5
ice (352 × 288)	3.28×10^4	3.65×10^4	3.57×10^4	3.35×10^4	2.92×10^4
ice (704 × 576)	1.60×10^5	1.58×10^5	1.55×10^5	1.49×10^5	1.36×10^5
deadline (352 × 288)	3.52×10^4	3.47×10^4	3.39×10^4	3.17×10^4	2.76×10^4

Table 2: Average number of bits for embedded data sequences per frame (bit/frame).

	Level 1	Level 2	Level 3	Level 4	Level 5
ice (352 × 288)	3.83×10^4	5.55×10^4	5.84×10^4	5.63×10^4	4.84×10^4
ice (704 × 576)	1.63×10^5	2.08×10^5	2.18×10^5	2.14×10^5	1.98×10^5
deadline (352 × 288)	6.10×10^4	7.68×10^4	7.98×10^4	7.63×10^4	6.68×10^4

Table 3: Average CPU time for generating recoverable privacy-protected image per frame (sec/frame).

	level 1	level 2	level 3	level 4	level 5
ice (352 × 288)	0.139	0.142	0.144	0.144	0.140
ice (704 × 576)	0.646	0.662	0.668	0.664	0.661
deadline (352 × 288)	0.137	0.142	0.142	0.143	0.140

Table 4: Average CPU time for recovering image per frame (sec/frame).

	Level 1	Level 2	Level 3	Level 4	Level 5
ice (352 × 288)	0.110	0.112	0.112	0.110	0.110
ice (704 × 576)	0.452	0.472	0.478	0.478	0.467
deadline (352 × 288)	0.109	0.113	0.114	0.113	0.112

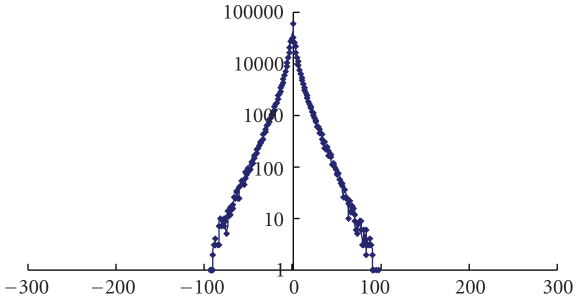


Figure 5: Histogram of wavelet coefficients.

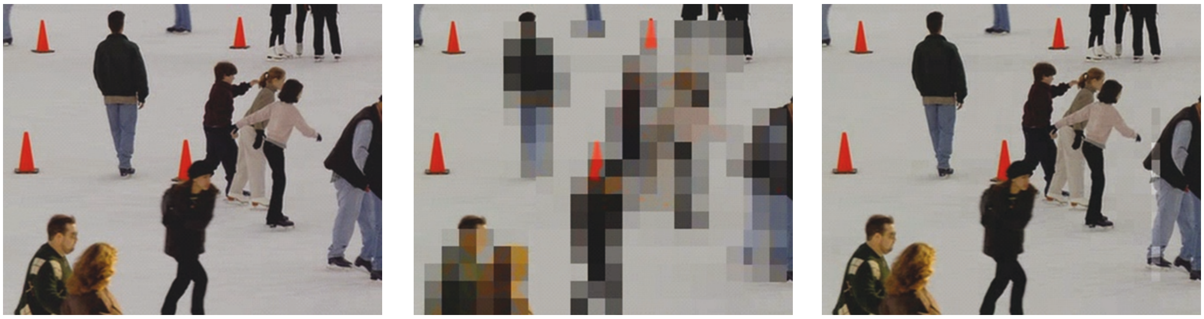
The deterioration due to embedding can be estimated by (7) and (8). Since amplitude modulo modulation in Section 4 uses congruence modulo 3, $|\widetilde{K}[m,n,l] - K[m,n,l]|$ in (8) becomes less than or equal to 2. Therefore, an upper bound of MSE can be calculated as follows:

$$\begin{aligned}\mathrm{MSE} &= \frac{\sum_{m=0}^{H-1}\sum_{n=0}^{W-1}\sum_{l=0}^{C-1}\left|\widetilde{K}[m,n,l]-K[m,n,l]\right|^2}{H\cdot W\cdot C}\\ &\le \frac{\sum_{m=0}^{H-1}\sum_{n=0}^{W-1}\sum_{l=0}^{C-1}2^2}{H\cdot W\cdot C} = 4.\end{aligned} \tag{9}$$

By this inequality, we obtain

$$\mathrm{PSNR} = 20\log\left(\frac{255}{\sqrt{\mathrm{MSE}}}\right) \ge 42.1. \tag{10}$$

This result implies that the deterioration of the privacy-protected image due to the embedding of privacy information is small enough so that we can ignore the influence of the embedding.

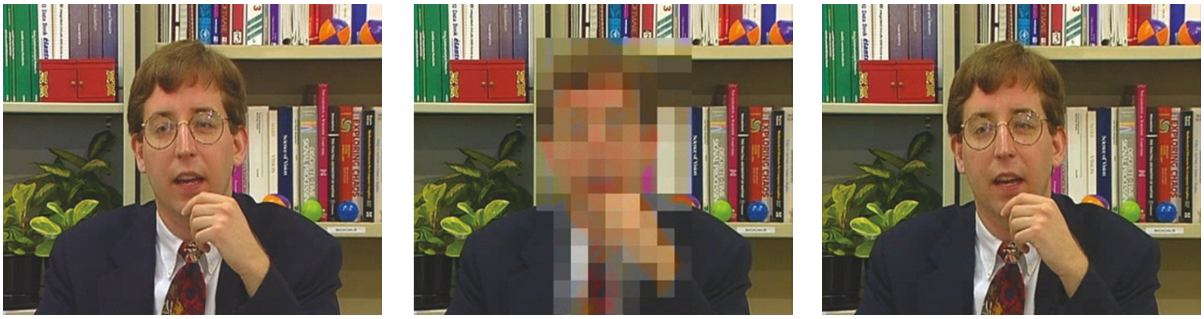


Original image (a)

Privacy protected image (b)

Recovered image (c)

Figure 6: Video sequence: ice (352 × 288), DWT level: $j = 4$, quantization step size: $\Delta = 1$, compression: JPEG.



Original image (a)

Privacy protected image (b)

Recovered image (c)

Figure 7: Video sequence: deadline (352 × 288), DWT level: $j = 4$, quantization step size: $\Delta = 1$, compression: JPEG.

5.2. Evaluation of the Deterioration of the Recovered Image. Here, we evaluate the deterioration of the recovered image by PSNR between the original image and the recovered image. Figures 9 and 10 show the PSNR of the recovered image for ice (352 × 288) and deadline (352 × 288) at the different quantization step size Δ and the different DWT levels. From Figures 9 and 10, we observe that PSNR becomes small as Δ becomes large. Almost PSNRs are larger than 30 (dB). Therefore, the proposed method can recover the image with low deterioration by appropriate choice of DWT level and quantization step size. We can also observe that PSNR of deadline (352 × 288) is worse than that of ice (352 × 288). This is due to the fact that the embedded data sequence of deadline (352 × 288) is larger than that of ice (352 × 288) as shown in Table 2.

5.3. Evaluation of Computational Time. Computational time for generating recoverable privacy-protected image and that for recovering image are shown in Tables 3 and 4, respectively, for each video sequence at different DWT levels under the condition $\Delta = 1$. From Tables 3 and 4, together with Tables 1 and 2, we observe that the influence of the resolution on computational time is dominant, whereas DWT level, the amount of embedded data sequence, and the number of pixels of object regions have an insignificant effect on the computational time.

The generation of recoverable privacy-protected image consists of the following four processes: object extraction, expanding low-resolution image after DWT, JPEG compression, and privacy data embedding. As for the image recovering, we have the following three processes, that is, privacy data extraction, JPEG decompression, and IDWT after shrinking low-resolution image. The rate of each process in generating recoverable privacy-protected image and that for image recovering are shown in Figures 11 and 12, respectively. From Figures 11 and 12, we can observe that computational costs for object extraction, privacy data embedding, privacy data extraction, and image recovering are very small compared to the costs for image compression. Therefore, our proposed method can be applied for real time processing, provided that the computational time for compression can be small.

6. Conclusion

In this paper, we have presented a method which attains recoverable privacy protection for video content distribution. By the proposed method, all the privacy information

```
• Input:
  Set of wavelet coefficients: $Y_w = \{a_z \mid z = 1, \dots, \alpha\}$ $(\alpha = 2^{M+N} - 2^{M+N-2j})$
  Quantization step size of the privacy information: $\Delta$
• Processing:
  Step1 : Generate the quantized data sequence $Y_w(\Delta) = \{a'_z = \lfloor a_z/\Delta \rfloor \mid z = 1, \dots, \alpha\}$ by
          quantizing $Y_w = \{a_z \mid z = 1, \dots, \alpha\}$, where $\lfloor c \rfloor$ is the largest integer
          that does not exceed $c$.
  Step2 : Find the intervals $[x_i, y_i](i = 1, \dots, \beta)$ consist of successive zeros (but the number
          of zeros is more than 2) from the data sequence $Y_w(\Delta)$, where $\beta, x_i, y_i$ are the number of
          such intervals,the $i$th smallest element of the set of starting points of successive zeros
          $\{x \mid 1 \leq x \leq \alpha, a'_x = a'_{x+1} = 0, a'_{x-1} \neq 0\}$, and the $i$th smallest element of the set
          of end points of successive zeros $\{y \mid 1 \leq y \leq \alpha, a'_y = a'_{y+1} = 0, a'_{y-1} \neq 0\}$.
          For this calculation, we suppose $a'_0 = a'_{\alpha+1} = 1$.
  Step3 :
    For $(z = 1, \dots, \alpha)$:
      For $(i = 1, \dots, \beta)$:
        If $(z \in [x_i, y_i])$:
          If $(z = y_i)$: Encode the data sequence of the interval $[x_i, y_i]$ with the run length
            coding, and add it to the data sequence $E$. Then Goto Next $z$
          Else $(x_i \leq z < y_i)$: Goto Next $z$
      If $(a'_z \geq 0)$: Encode $a'_z$ to binary bits, and add the binary bits to the data sequence $E$ with
        delimiter digit 2.
      Else $(a'_z < 0)$: Encode $a'_z$ to binary bits, and add the binary bits sandwiched by
          the sign bit 0 and delimiter digit 2 to the data sequence $E$.
• Output:
  Data sequence to be embedded: $E = \{e_1, e_2, \dots\}$
```

Algorithm 1

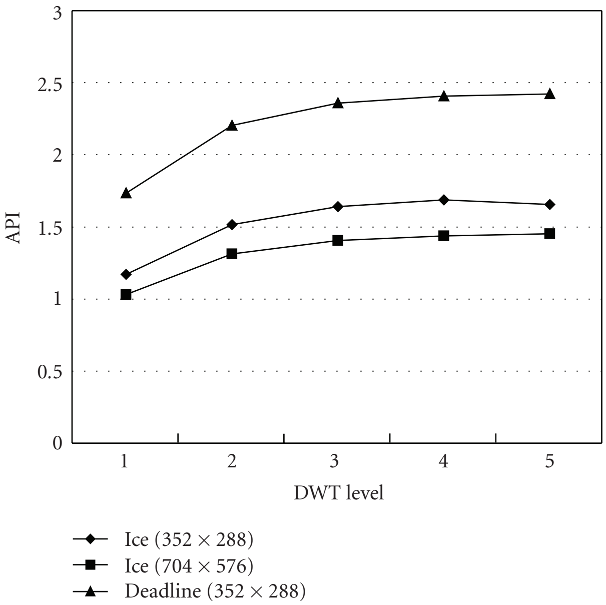


Figure 8: API for the different level of DWT.

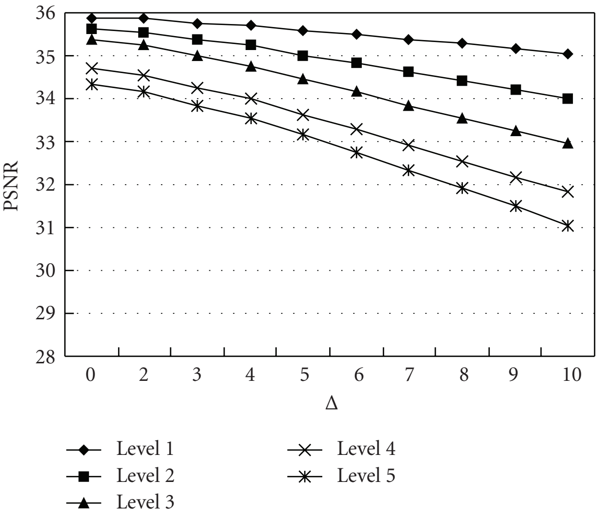


Figure 9: PSNR of recovered image. (ice (352 × 288)).

can be embedded into the privacy-protected image even if the level of DWT is large and the object region size is equal to the size of the whole image. We also show that proposed method recovers the privacy-protected image with low deterioration, and the computational time for privacy protection and image recovering is small.

The proposed method is based on the idea that the privacy information needed for recovering video sequence is embedded in the video sequence itself (which is referred to as self-recoverable), and only authorized viewers can extract the privacy information. An alternate approach is that the privacy information for recovering video sequence is stored

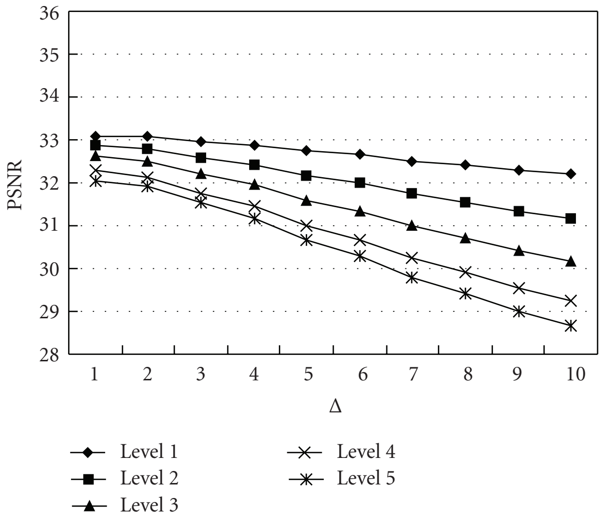


FIGURE 10: PSNR of recovered image. (deadline (352 × 288)).

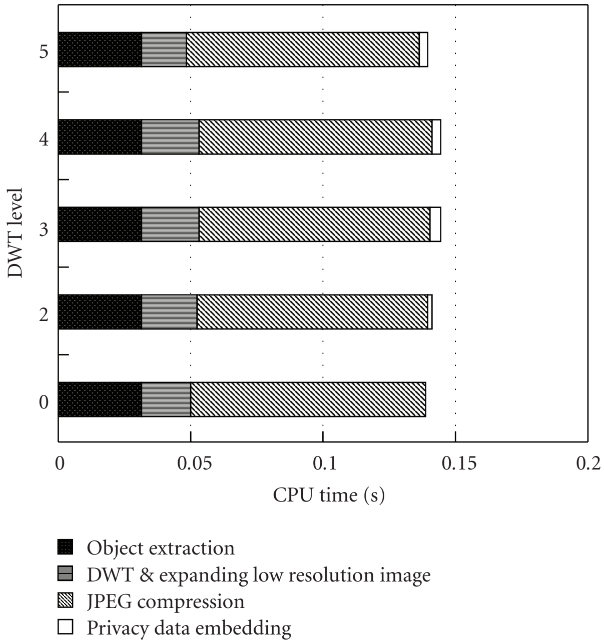


FIGURE 11: Rates of CPU time of each processing for generating recoverable protected image.

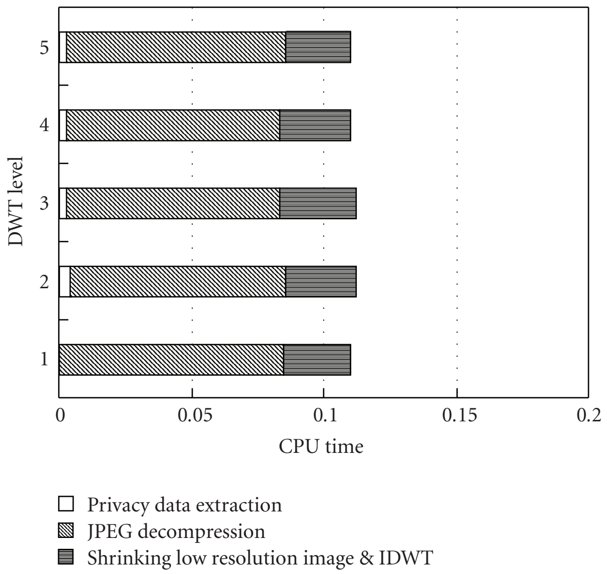


FIGURE 12: Rates of CPU time of each processing for image recovering.

outside (e.g., in a server), and only authorized viewers can access the privacy information. Such a system would be more secure than self-recoverable system, although the system is inferior with respect to the convenience. It is desirable to develop the system that can deal with both methods for protecting privacy information.

Currently, we use the same DWT level for each object region in a single image. However, it is better to change the DWT level adaptively according to the size of object region since the permissible visible detail of each object is not identical. The realization of this function is one of our future work. In our current method, when JPEG2000 is applied for image compression, we have to calculate DWT twice; one is for image compression using 5–3 filter, and another one is for privacy protection using Haar bases. If these two DWTs can be unified, the process of recoverable privacy protection becomes much simpler, and the computational time will be further reduced. The development of such methods is also our future work.

Acknowledgments

This work was supported in part by a Grant-in-Aid for scientific research from the Japan Society for the Promotion of Science and by SCOPE from the Ministry of Internal Affairs and Communications, Japan.

References

[1] E. M. Newton, L. Sweeney, and B. Malin, "Preserving privacy by de-identifying face images," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 2, pp. 232–243, 2005.

[2] I. Kitahara, K. Kogure, and N. Hagita, "Stealth vision for protecting privacy," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, vol. 4, pp. 404–407, Cambridge, UK, August 2004.

[3] J. Wickramasuriya, M. Datt, S. Mehrotra, and N. Venkatasubramanian, "Privacy protecting data collection in media spaces," in *Proceedings of the 12th ACM International Conference on Multimedia (ACM Multimedia '04)*, pp. 48–55, New York, NY, USA, October 2004.

[4] M. Boyle, C. Edwards, and S. Greenberg, "The effects of filtered video on awareness and privacy," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 1–10, Philadelphia, Pa, USA, December 2000.

[5] J. L. Crowley, J. Coutaz, and F. Babaguchi, "Things that see," *Communications of the ACM*, vol. 43, no. 3, pp. 54–64, 2000.

[6] K. Chinomi, N. Nitta, Y. Ito, and N. Babaguchi, “PriSurv: privacy protected video surveillance system using adaptive visual abstraction,” in *Proceedings of the 14th International Multimedia Modeling Conference (MMM ’08)*, vol. 4903 of *Lecture Notes in Computer Science*, pp. 144–154, Kyoto, Japan, January 2008.

[7] X. Yu, K. Chinomi, T. Koshimizu, N. Nitta, Y. Ito, and N. Babaguchi, “Privacy protecting visual processing for secure video surveillance,” in *Proceedings of the International Conference on Image Processing (ICIP ’08)*, pp. 1672–1675, San Diego, Calif, USA, October 2008.

[8] F. Dufaux and T. Ebrahimi, “Scrambling for video surveillance with privacy,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, New York, NY, USA, June 2006.

[9] F. Dufaux, M. Ouaret, Y. Abdeljaoued, A. Navarro, F. Vergnenègre, and T. Ebrahimi, “Privacy enabling technology for video surveillance,” in *Mobile Multimedia/Image Processing for Military and Security Applications*, vol. 6250 of *Proceedings of SPIE*, Kissimmee, Fla, USA, April 2006.

[10] W. Zhang, S.-C. S. Cheung, and M. Chen, “Hiding privacy information in video surveillance system,” in *Proceedings of the International Conference on Image Processing (ICIP ’05)*, vol. 3, pp. 868–871, Genova, Italy, September 2005.

[11] X. Yu and N. Babaguchi, “Privacy preserving: hiding a face in a face,” in *Proceedings of the 8th Asian Conference on Computer Vision (ACCV ’07)*, vol. 4844 of *Lecture Notes in Computer Science*, pp. 651–661, Tokyo, Japan, November 2007.

[12] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active appearance models,” in *Proceedings of the European Conference on Computer Vision*, vol. 2, pp. 484–498, 1998.

[13] G. Li, Y. Ito, X. Yu, N. Nitta, and N. Babaguchi, “A discrete wavelet transform based recoverable image processing for privacy protection,” in *Proceedings of the International Conference on Image Processing (ICIP ’08)*, pp. 1372–1375, San Diego, Calif, USA, October 2008.

[14] G. K. Wallace, “The JPEG still picture compression standard,” *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. 18–34, 1992.

[15] A. Skodras, C. Christopoulos, and T. Ebrahimi, “The JPEG 2000 still image compression standard,” *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 36–58, 2001.

[16] W. E. L. Grimson, C. Stauffer, R. Romano, and L. Lee, “Using adaptive tracking to classify and monitor activities in a site,” in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 22–29, Santa Barbara, Calif, USA, June 1998.

[17] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.

[18] M. Wu, *Multimedia data hiding*, Ph.D. dissertation, Princeton University, 2001.