ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Image Steganography-based Copyright and Privacy-Protected Image Trading Systems

Wannida Sae-Tang[1] and Adisorn Sirikham[2]

ABSTRACT

This paper proposes steganography-based copyright- and privacy-protected image trading systems using image transformation, i.e., either discrete cosine transform (DCT) or Hadamard transform (HT). In the systems, there are a content provider (CP), a consumer, the first trusted third party (TTP), and the second TTP. To protect the copyright of the image, the consumer ID is embedded into the amplitude components of the commercial image by the first TTP using the digital fingerprinting technique, and to protect the consumer's privacy against the first TTP and a malicious third party (s), the image steganography is applied to the commercial image by using image transformation. A color dummy image is used instead of a gray dummy image for security purposes. After applying the image transformation to both images, the coefficient signs of the commercial image are replaced by the coefficient signs of the dummy image pixel-by-pixel so that the inversely transformed commercial image looks like the dummy image instead of the commercial image. Once the consumer receives the fingerprinted image from the first TTP and the coefficient signs of the commercial image from the second TTP, the consumer reconstructs the fingerprinted commercial image without losing the hidden fingerprint at all because of the compatibility of the proposed image steganography method and the amplitude-based fingerprinting method. The experimental results confirm that the stego-images generated by the proposed systems do not look suspicious with higher qualities compared with those generated by existing systems. Moreover, the fingerprinted image quality and the correct fingerprint extracting rate have been improved by the proposed systems.

Article information:

Keywords: Image Copyright Protection, Image Data Hiding, Image Encryption, Image Steganography, Image Trading System

Article history:
Received: May 1, 2023
Revised: May 27, 2023
Accepted: July 20, 2023
Published: August 12, 2023 (Online)

DOI: 10.37936/ecti-cit.2023173.252500

1. INTRODUCTION

Nowadays, online image trading is very popular. Copyright protection techniques are used to protect commercial images [1]-[14]. In addition, consumer's privacy is protected in the systems [4]-[14]. To protect the consumer's privacy against a content provider (CP), the commercial image is copyright-protected by a trusted third party (TTP) instead of the CP [4]-[14]. Also, the consumer's privacy should be protected against the TTP. The CP, therefore, encrypts the commercial image before sending it to the TTP. The copyright-protected image is then transmitted from the TTP to the consumer simultaneously with sending the decryption key from the CP to the consumer. Finally, the copyright-protected commercial image is reconstructed by the consumer. In this system, the consumer's privacy is protected against both the CP and the TTP. Besides considering the CP and the TTP, there is the possibility that a malicious third party (s) attacks/stoles the image from the TTP. For this situation, the image encryption technique also protects the image from a malicious third party (s).

However, using image steganography instead of image encryption has a benefit. That is the TTP

[1] The author is with The Sirindhorn International Thai-German Graduate School of Engineering, King Mongkut's University of Technology North Bangkok, Thailand, E-mail: wannida.s@tggs.kmutnb.ac.th
[2] The author is with Department of Electrical and Telecommunication Engineering, Rajamangala University of Technology Krungthep, Thailand, E-mail: adisorn.s@mail.rmutk.ac.th

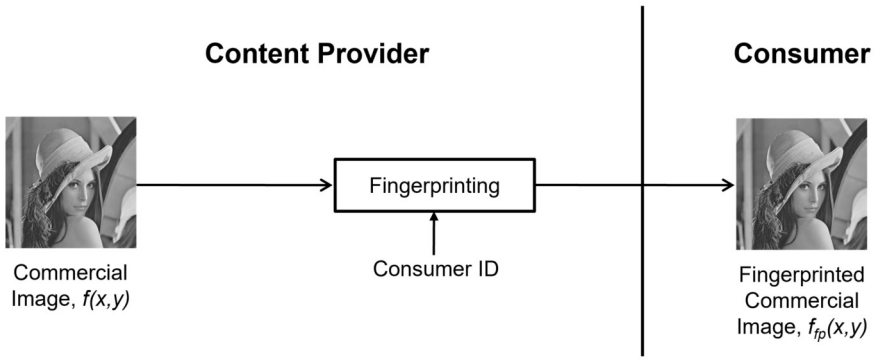

Fig.1: *Copyright-protected image trading systems [1]-[3].*

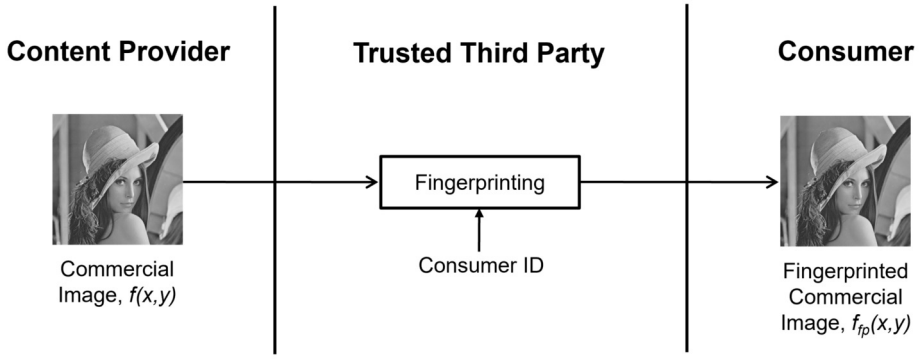

Fig.2: *Copyright- and privacy-protected image trading system with a trusted third party [4]-[13].*

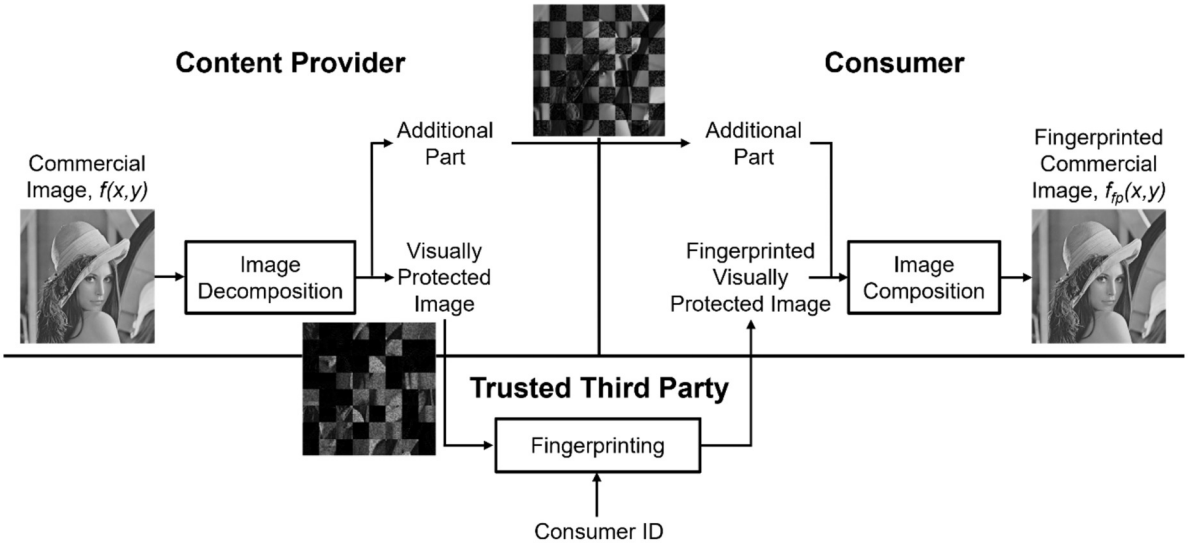

Fig.3: *Copyright- and privacy-protected image trading system using image decomposition [7].*

and the malicious third party (s) do not notice something's wrong with the image. This paper then proposes steganography-based copyright- and privacy-protected image trading systems using image transformation, i.e., either discrete cosine transform (DCT) or Hadamard transform (HT) is used. In the systems, there are a CP, a consumer, the first TTP, and the second TTP. To protect the copyright of the image, the consumer ID is embedded into the amplitude components of the commercial image by the first TTP using the digital fingerprinting technique, and to protect the consumer's privacy against the first TTP, the image steganography is applied to the commercial image by using image transformation (Either the DCT or the HT is used.). A color dummy image is used instead of a gray image to increase the difficulty of the attack. After applying the image transformation to both images, the coefficient signs of the commercial image are replaced by the coefficient signs of the dummy image pixel-by-pixel so that the inversely

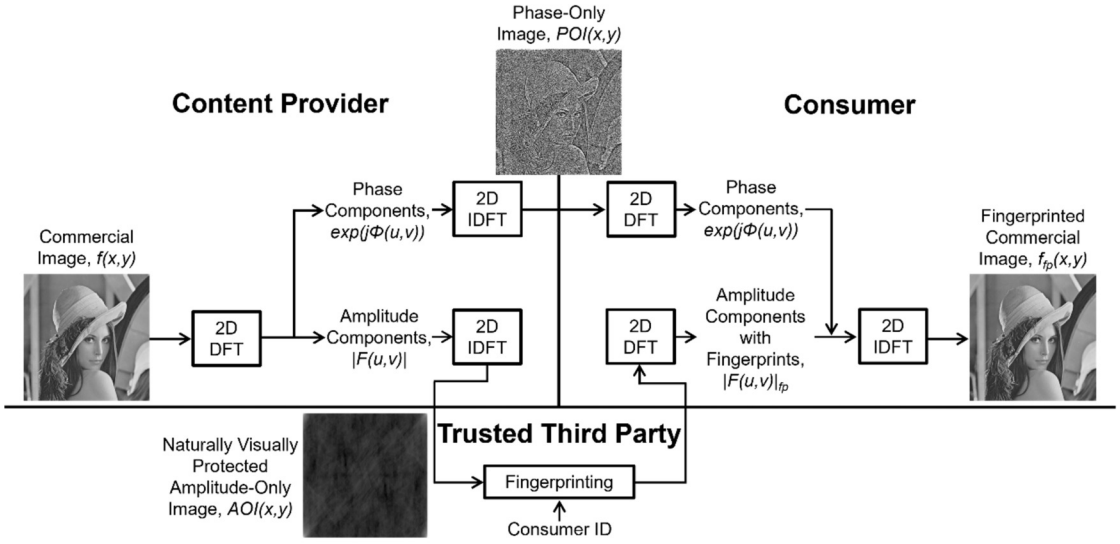

Fig.4: *Copyright- and privacy-protected image trading system using the DFTed AOI [8].*

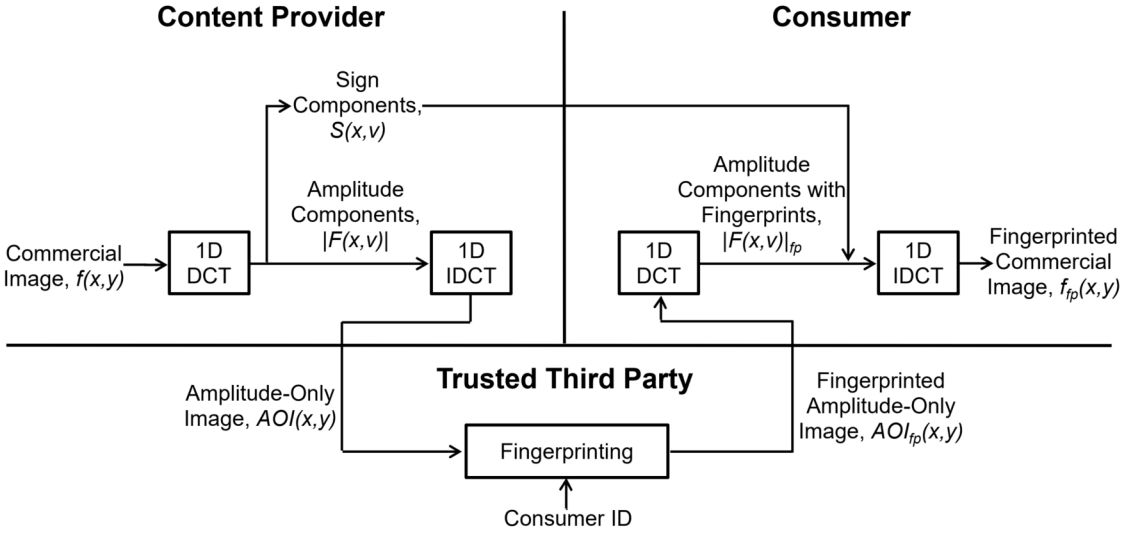

Fig.5: *Copyright- and privacy-protected image trading system using the 1D DCTed AOI [10].*

transformed commercial image looks like the dummy image instead of the commercial image. The coefficient signs of the commercial image are sent to the second TTP before sending them from the second TTP to the consumer. This is because the CP is not allowed to directly contact the consumer for protecting the consumer's privacy. Once the consumer receives the fingerprinted image from the first TTP and the coefficient signs of the commercial image from the second TTP, the consumer reconstructs the fingerprinted commercial image with a hidden fingerprint.

The rest of this paper is organized as follows. Section 2 describes existing copyright- and privacy-protected image trading systems. Section 3 describes the proposed systems. Experimental results and discussions are given in section 4. Finally, section 5 concludes this paper.

2. EXISTING IMAGE TRADING SYSTEMS

Fig. 1 shows the image trading system in which the image is copyright-protected by the CP using the digital fingerprinting technique [1]-[3]. The fingerprinted image is then sent directly from the CP to the consumer. In this system, the consumer's privacy is not protected, while in the system shown in Fig. 2, the consumer's privacy is protected against the CP by introducing the TTP to the system [4]-[13]. The TTP is responsible for the image copyright-protection process instead of the CP. Therefore, the CP does not need to contact the consumer directly. As a result, the consumer's privacy is protected against the CP. To perform the copyright protection process, the TTP uses the digital fingerprinting technique to embed the consumer's information, such as the con-

sumer ID, into the image. However, in this system, the consumer's information could be leaked from the TTP. Using either the image visual encryption technique or the image steganography technique can solve this problem.

2.1 Encryption-based systems

Figs. 3-5 show the copyright- and privacy-protected image trading systems in which the image visual encryption technique is used. Fig. 3 shows the system using image decomposition [7]. The original image is decomposed into 2 parts by the CP. One is sent from the CP to the TTP, and the other is sent directly from the CP to the consumer. Then, the TTP embeds the consumer ID into the first part of the original image using the digital fingerprinting technique and subsequently sends the fingerprinted part to the consumer. The consumer composes both parts to reconstruct the fingerprinted image which looks like the original image. In this system, the fingerprinting capacity is insufficient. Fig. 4 shows the system using discrete Fourier transformed amplitude-only image (DFTed AOI) [8] which allows higher fingerprinting capacity. Firstly, the original image is transformed into the frequency domain using two-dimensional discrete Fourier transform (2D DFT). Then, the amplitude components and the phase components are separated. Finally, the amplitude components and the phase components are inversely transformed into the spatial domain independently. The outputs are called the AOI and phase-only image (POI), respectively. Fig. 5 shows the system using one-dimensional discrete cosine transformed amplitude-only image (1D DCTed AOI) [10]. In this system, one-dimensional discrete cosine transform (1D DCT) is used instead of the 2D DFT and yields better results in terms of image quality, fingerprinting performance, and complexity.

2.2 Steganography-based systems

Although using the image encryption technique can protect both image copyright and consumer's privacy, using image steganography instead of image encryption has a better benefit. That is the first TTP and the malicious third party (s) do not notice anything wrong with the image so that they do not try to attack the image.

Image steganography methods could be grouped into three categories, i.e., traditional methods, convolutional neural network (CNN)-based methods, and general adversarial networks (GAN)-based methods.

1) Traditional methods: Traditional methods are frameworks that use methods that are not related to machine learning or deep learning algorithms. Many traditional methods are based on the least significant bits (LSB) technique. LSB works under the assumption that modifying a few pixel values would not show any visible changes. The secret information is converted into a binary and is then substituted in the LSB of the cover image. The substitution method must be performed cautiously to make sure that no visible changes are leaking the presence of the secret information [15]-[16]. Besides the LSB methods, a combination of DCT and discrete wavelet transformation (DWT) for hiding the secret message in a cover video has been proposed [17]. Pixel value differencing (PVD) was proposed in [18]. It works by taking the difference of consecutive pixels to search for the locations for hiding the secret bits in such a way that the consistency of the cover image is maintained. Another proposed technique is coverless steganography, where the cover image is not given. It is generated based on secret information using an object detection method. Similar coverless steganography was proposed where the local binary patterns (LBP) features of the cover image and the secret images are hashed first. Later, the hashes are matched to create the stego-image [19]. Similarly, instead of LBP, the edges of the color cover images are obtained. Then, the binary bits of the secret information are hidden in the edges of the cover images [20]. DCT coefficients were used for medical JPEG image steganography by taking the difference in the DCT coefficients of the pixels between two consecutive blocks by considering the pixels at the same positions in the two blocks [21]. A novel method called the pixel density histogram (PHD) was proposed for halftone images [22]. The pixel density of the images is calculated, and a pixel density histogram is formed to hide the secret information. Another method using STC coding [23] utilizes the Poisson distribution to get the burst error and the reconstruction of the images that are compression-resistant. For more clarifications and explanations on the traditional steganography methods, [24] can be referred to.

2) CNN-based methods: CNN-based methods are a type of deep CNNs. The cover image and the secret image are fed as input to the encoder to generate the stego-image, and the stego-image is fed as input to the decoder to output the embedded secret image. Wu et al. proposed an encoder-decoder architecture [25]-[26]. U-Net-based encoder-decoder architecture used for hiding and a CNN with 6 layers for extraction was proposed by Duan et al. [27]. In addition, A U-net-based hiding (H-net) and revealing (R-net) network are used by Van et al. in [28]. A separable convolution with a residual block (SCR) is used to concatenate the cover image and the secret image in [25]. The embedded image is given as the input to the encoder for constructing the stego-image which is fed to the decoder for secret image extraction. In [26], a new cost function reducing the effect of noise in the generated container image called "variance loss" was proposed. While an encoder-decoder architecture was proposed in [29]. This method differs from other

methods in the way the inputs are given. The encoder part consists of two parallel architectures. One is for the cover image, and the other is for the secret image. Features from the cover image and the secret images are extracted through the convolutional layer and concatenated. The concatenated features are used to construct the stego-image. A slightly different approach was proposed in [30]. The created stego-image is converted into the style image given as input. The revealed network is used to decode the secret information from the stego-image created. Similar to other methods, an auto encoder-decoder architecture with VGG-based is used. An arbitrary image size for the secret information and the style image is taken using the adaptive instance normalization (AdaIN) layer, and the output size is the size of the cover image. In [31], the pixel distribution of the cover image is obtained by using a pixel CNN. The secret information is then embedded into the pixel distribution evenly by reduced sampling. In [32], the cover image is converted into a YCrCb image format, and the secret image is embedded into only the Y channel as all the semantic and color information is presented in the Cr and Cb channels. Not only image steganography but also, video steganography has been proposed using CNN. Usually, 2D convolutional layers are used for images, whereas 3D convolutional layers are used for videos. In [33], temporally connected cover and secret video frames are given as the input to the autoencoder to produce the container video.

3) GAN-based methods: GAN-based methods use some of the GAN variants [34]. GANs are known for their good performance in the image generation field. Image steganography can be considered as one such image generation task where two inputs (the cover image and the secret image) are given to generate one output (stego-image). The existing methods used for image steganography using a GAN architecture can be grouped into five categories; a three-network-based GAN model, cycle-GAN-based architectures, sender-receiver architecture using GAN, a coverless model where the cover image is generated randomly instead of being given as input, and an Alice, Bob, and Eve-based model. Volkhonskiy et al. [35]-[36] proposed DCGAN-based steganographic GAN (SGAN) [37], which is a simple DCGAN with three modules (G, D, and S). Similar to [35], a three-component GAN architecture was proposed by Shi et al. [38]. DCGAN is used in [35] and [36], whereas a four fractionally convolutional layer followed by a functional layer with hyperbolic tangent activation with WGAN is used in [38]. Yang et al. [39]-[40] presented a GAN-based image steganography with three modules: Generator, Embedding Simulator, and Discriminatory. The generator is used to create the probability map for the input cover image. U-Net is considered because of its effective performance in pixel-wise segmentation. The embedding simulator uses a double tanh function [40] and a tanh function [39] because the tanh is differentiable and preserves the gradient loss during backpropagation. An automatic steganographic distortion learning framework with GAN (ASDL-GAN) was proposed in [41]. In this architecture, the generator is used to learn the probabilities for each pixel from the input cover image. In addition, a novel activation function ternary embedding simulator (TES) for generating stego-images from the generated probabilities was also proposed in this architecture. The discriminator helps in differentiating between real and fake images. HIDDeN proposed by Zhu et al. [42] is a GAN-based method with four main components: an encoder, a decoder, a discriminator, and a noise layer. CycleGAN [43] is well-suited for image steganography, where the input image is given and the output which is similar to the given input image but with hidden information using the adversarial training is generated. ACGAN [44] can generate realistic images for a given label and also recognize the label of the generated images. Three steps are followed to hide information which is generated by the model and extract the hidden information [45] and [46]. Similar to [45] and [46], a coverless steganographic method using a generative model was proposed in [47]. A model with four parts: Alice, Bob, Dev, and Eve, has been proposed in [48]. The model is an unsupervised generative model and is named self-supervised steganographic GAN (SSteGAN).

Hiding capacity and accuracy consisting of security and robustness factors are considered to evaluate image steganography methods. The hiding capacity is defined as the amount of secret data embedded into an image. While the security is related to embedding, and the robustness is related to the extraction of the secret image. In terms of the hiding capacity, for the focused application, it is expected that a secret image can be embedded into the cover image of the same size. Among the three groups of image steganography methods, the methods with the least hiding capacity are the traditional-based methods, where text is the primary form of secret information. The second is the GAN-based methods, where only text is used as secret information. In contrast, the CNN-based methods can embed a gray image into the cover image of the same size. CNN-based methods clearly outperform other methods and seem to be able to be applied to the focused application. However, other indicators must be considered. In terms of security and robustness, CNN-based methods and GAN-based methods yield higher security than traditional-based methods. Although the traditional-based methods yield lower security, they yield higher robustness than deep learning-based methods. The extraction of the secret image is prone to loss of information in deep learning-based methods. It is not clear where and

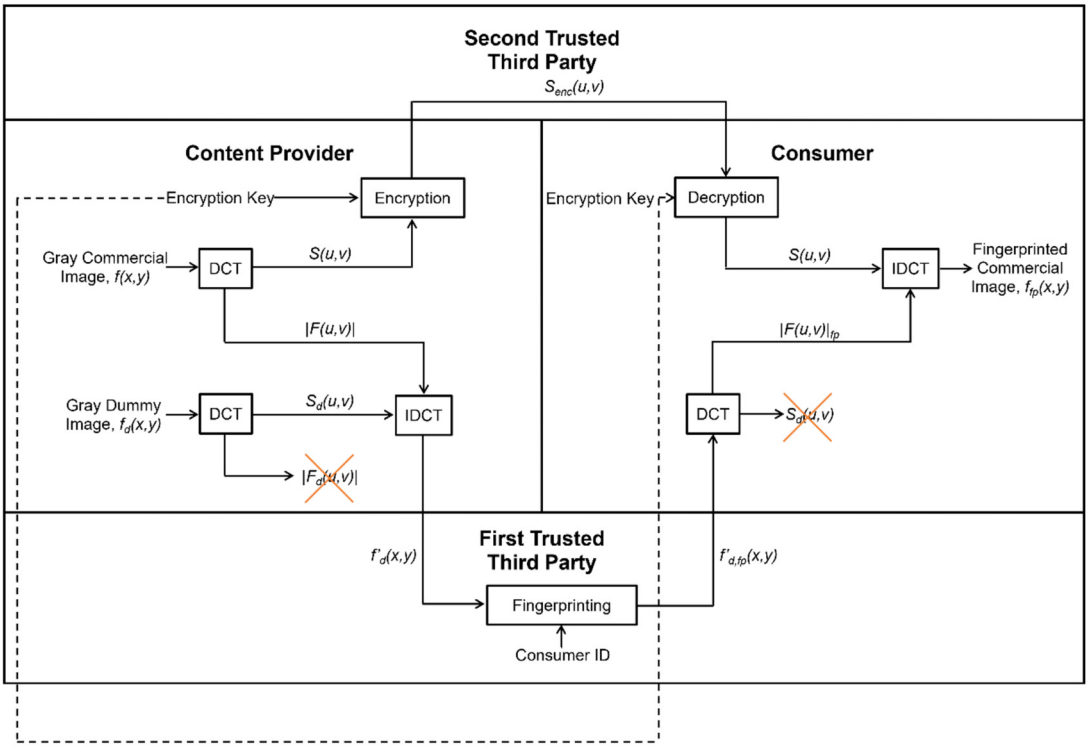

Fig.6: *Image steganography-based copyright- and privacy-protected image trading system using the DCT [14].*

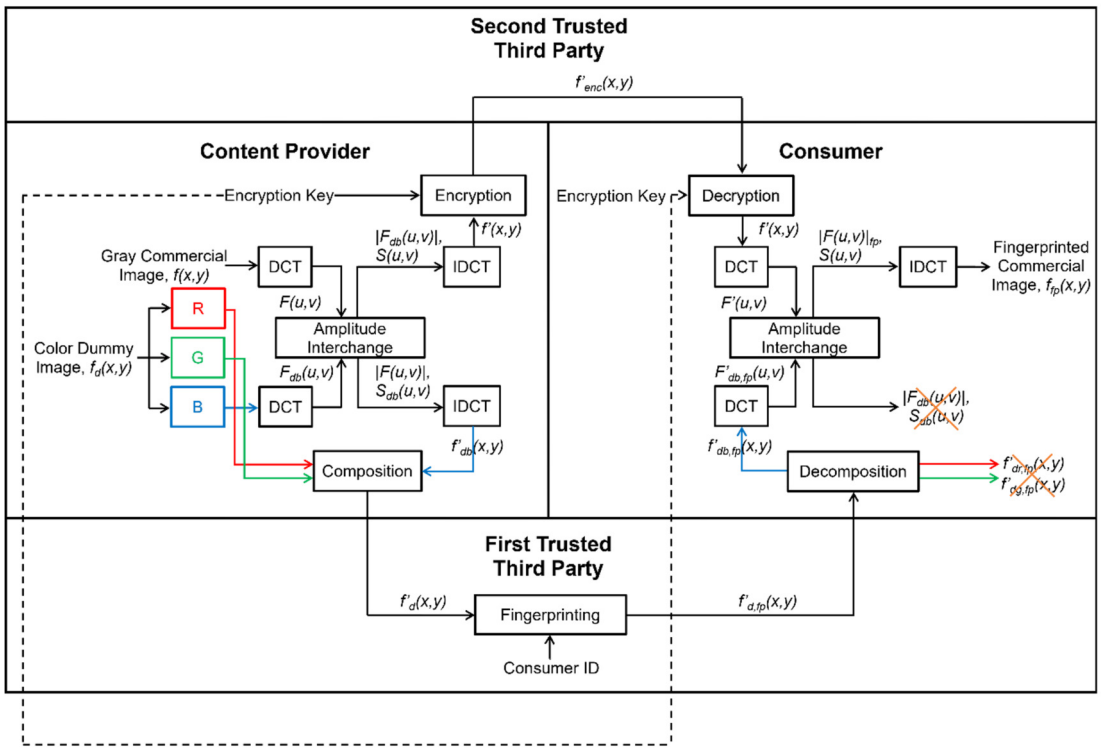

Fig.7: *Proposed image steganography-based copyright- and privacy-protected image trading system using the DCT.*

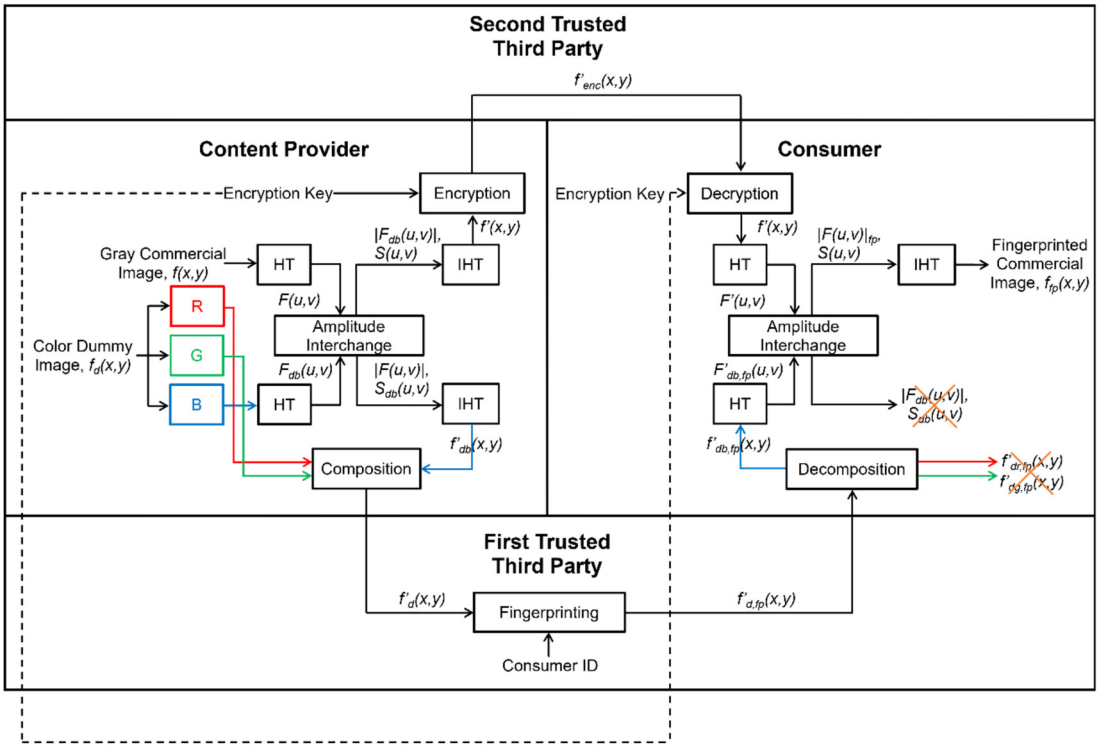

Fig.8: *Proposed image steganography-based copyright- and privacy-protected image trading system using the HT.*

how the secret message pixels are embedded. Therefore, it is difficult to crack the steganography image without the extracting model trained. This increases the security but reduces the robustness, even the secret message cannot be extracted if the extraction model is not working. This is an unexpected situation for the focused application. The robustness is very important in the focused application because the consumer must obtain the secret image 100% correctly in the commercial world. Moreover, in terms of complexity, the deep learning-based methods are worst by the time taken for training, testing, embedding, and extracting. Therefore, deep learning-based methods are not considered for the focused application anymore. In fact, any general image steganography methods are also not suitable for the focused application, because the fingerprint for image copyright protection embedded into the stego-image is disrupted in the steganalysis process unavoidably. By using a general steganography method with a general fingerprinting method, the fingerprint must be embedded into both the cover image and the secret image. If the fingerprint is embedded into both the cover image and the secret image, the fingerprint bits embedded into the cover image are discarded in the steganalysis process. Only fingerprint bits embedded into the secret image are remained. That causes the copyright protection performance degraded. From this reason, a specific image steganography method with a compatible fingerprinting method is then required!

With the limitation mentioned above, image steganography-based copyright- and privacy-protected image trading systems were proposed in [14] as shown in Fig. 6. In the systems, the two-dimensional discrete cosine transform (2D DCT) and the 1D DCT were used with a gray dummy image for image steganography. In terms of hiding capacity, the systems can embed a gray image into the cover image of the same size. In terms of security, the systems generate suspicious stego-image with low PSNRs so that the intruders can recognize that they are steganography images or encrypted images. However, the systems are good in terms of robustness and complexity. They can extract the secret image correctly and are fast. The method is also compatible with the amplitude-based fingerprinting technique. However, the average PSNR of the fingerprinted image quality was less than 47 dB, and the average correct fingerprint extracting rate did not reach 100%. These performances could be improved.

Table 1: *Performance summary of the existing methods/systems.*

	LSB	**CNN**	**GAN**	**[14]**
Hiding Capacity	X	O	X	O
Security	X	O	O	X
Robustness	O	X	X	O
Complexity	O	X	X	O
Fingerprinting Performance	X	X	X	O*

Table 1 concludes the performances of all mentioned methods, where 'O,' 'O*,' and 'X' are defined as 'Good,' 'Good but could be improved,' and 'Bad,' respectively.

3. PROPOSED COPYRIGHT- AND PRIVACY-PROTECTED IMAGE TRADING SYSTEMS

Since there are rooms to enhance the security, the fingerprinted image quality, and the fingerprint extraction performance, this paper proposes steganography-based copyright- and privacy-protected image trading systems using image transformation as shown in Figs. 7-8. In the systems, there are the CP, the consumer, the first TTP, and the second TTP. The first TTP is responsible for embedding the consumer ID into the commercial image. To protect the consumer's privacy against the first TTP, a new image steganography method using image transformation is proposed. Either the 2D DCT, the 1D DCT, two-dimensional HT (2D HT), or one-dimensional HT (1D HT) is used, Therefore, in total, there are four proposed systems. While the commercial image is a gray image, a color dummy image of the same size is used. The reason why a color dummy image is used instead of a gray dummy image is that it could potentially achieve higher security. The commercial image is considered as a secret image, whereas the dummy image is considered as a cover image, in the steganography process.

3.1 Steganography process

The amplitude components of the commercial image are embedded into the blue band of the cover image. Firstly, the $X \times Y$-sized commercial image, $f(x, y)$, and the blue band of the dummy image, $f_{db}(x, y)$, of the same size are transformed independently into the frequency domain by using the 2D DCT. The commercial image, $f(x, y)$, is transformed as

$$F_{2DDCT}(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x,y) \cos\left(\frac{\pi}{X}\left(x + \frac{1}{2}\right)u\right) \cos\left(\frac{\pi}{Y}\left(y + \frac{1}{2}\right)v\right), \quad (1)$$

where $F_{2DDCT}(u, v)$ denotes the two-dimensional discrete cosine transformed (2D DCTed) coefficients of the commercial image,

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{X}}, u = 0 \\ \sqrt{\frac{2}{Y}}, u = 1, 2, ..., X-1 \end{cases}, \quad (2)$$

and

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{Y}}, v = 0 \\ \sqrt{\frac{2}{Y}}, v = 1, 2, ..., Y-1 \end{cases}, \quad (3)$$

and $u = 0, 1, \ldots, X-1$, and $v = 0, 1, \ldots, Y-1$. $F_{2DDCT}(u, v)$ can also be expressed in the polar form as

$$F_{2DDCT}(u,v) = |F_{2DDCT}(u,v)|\, S_{2DDCT}(u,v), \quad (4)$$

where $|F_{2DDCT}(u, v)|$ and $S_{2DDCT}(u, v)$ denote the amplitude components and the sign components, respectively. In the same way, the blue channel of the dummy image, $f_{db}(x, y)$, is transformed by

$$F_{db,2DDCT}(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f_{db}(x,y) \cos\left(\frac{\pi}{X}\left(x + \frac{1}{2}\right)u\right) \cos\left(\frac{\pi}{Y}\left(y + \frac{1}{2}\right)v\right), \quad (5)$$

where $F_{db,2DDCT}(u, v)$ denotes the 2D DCTed coefficients of the blue channel of the dummy image,

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{X}}, u = 0 \\ \sqrt{\frac{2}{X}}, u = 1, 2, ..., X-1 \end{cases}, \quad (6)$$

and

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{Y}}, v = 0 \\ \sqrt{\frac{2}{Y}}, v = 1, 2, ..., Y-1 \end{cases}, \quad (7)$$

and $u = 0, 1, \ldots, X-1$, and $v = 0, 1, \ldots, Y-1$. $F_{db,2DDCT}(u, v)$ can also be expressed in the polar form as

$$F_{db,2DDCT}(u,v) = |F_{db,2DDCT}(u,v)|\, S_{db,2DDCT}(u,v), \quad (8)$$

where $|F_{db,2DDCT}(u, v)|$ and $S_{db,2DDCT}(u, v)$ denote the amplitude components and the sign components, respectively. The amplitude components of the commercial image, $|F_{2DDCT}(u, v)|$, are then combined with the 2D DCTed coefficient signs of the blue channel of the dummy image, $S_{bd,2DDCT}(u, v)$. Finally, the two-dimensional inverse discrete cosine transform (2D IDCT) is used to obtain the stego-image, $f'_{db,2DDCT}(x, y)$, as described by Eq. (9).

$$f'_{db,2DDCT}(x,y) = \alpha(u)\alpha(v)\sum_{x=0}^{X-1}\sum_{y=0}^{Y-1}|F_{2DDCT}(u,v)|S_{db,2DDCT}(u,v)\cos\left(\frac{\pi}{X}\left(x+\frac{1}{2}\right)u\right)\cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right). \quad (9)$$

After that, the stego-image, $f'_{db,2DDCT}(x,y)$, is combined with the original red channel and the original green channel of the dummy image to increase the image visual quality for security purposes. This image looks like the dummy image and is sent to the first TTP. Therefore, the dummy image is considered as the cover image. This is the reason for using a color image as the cover image, and the idea to use the blue channel for steganography is that the change in the blue channel is not so sensitive to human eyes. The green and red cones are mostly packed into the fovea centralis. By population, about 64% of the cones are red-sensitive, about 32% are green-sensitive, and about 2% are blue-sensitive [49]. In the same way, the amplitude components of the dummy image, $|F_{db,2DDCT}(u,v)|$, are then combined with the 2D DCTed coefficient signs of the commercial image, $S_{2DDCT}(u,v)$. The 2D IDCT is used to obtain another stego-image, $f'_{2DDCT}(x,y)$, as described by Eq. (10).

$$f'_{2DDCT}(x,y) = \alpha(u)\alpha(v)\sum_{x=0}^{X-1}\sum_{y=0}^{Y-1}|F_{db,2DDCT}(u,v)|S_{2DDCT}(u,v)\cos\left(\frac{\pi}{X}\left(x+\frac{1}{2}\right)u\right)\cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right). \quad (10)$$

This image looks like the commercial image and is sent to the consumer as a key for image reconstruction. Since the CP is not allowed to directly contact the consumer to protect the consumer's privacy against the CP, the second TTP is then introduced to the systems. The CP sends the stego-image, $f'_{2DDCT}(x,y)$, to the second TTP, and then the second TTP sends it to the consumer. With relying on the encryption technique, to send the coefficient signs to the second TTP, the CP can send either the encrypted version of the stego-image, $f'_{2DDCT}(x,y)$, or the encrypted version of the sign components, $S_{2DDCT}(u,v)$. That is the amplitude components of the dummy image, $|F_{db,2DDCT}(u,v)|$, can be discarded, and the 2D DCTed coefficient signs of the commercial image, $S_{2DDCT}(u,v)$, can be sent to the second TTP as a binary image. (2D DCTed coefficient signs consist of plus and minus which can be stored as a binary image.) The difference in terms of security of both options is shown in Section 4.1.

The processes are performed in the same way for other transforms, i.e., the 1D DCT, the 2D HT, and the 1D HT. For one-dimensional transformations, the images are transformed column by column. For the proposed 1D DCT-based system, the $X \times Y$-sized commercial image, $f(x,y)$, and the blue channel of the dummy image, $f_{db}(x,y)$, are transformed independently into the frequency domain by using the 1D DCT. The commercial image, $f(x,y)$, is transformed as

$$F_{1DDCT}(x,v) = \alpha(v)\sum_{y=0}^{Y-1} f(x,y)\cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right), \quad (11)$$

where $F_{1DDCT}(x,v)$ denotes the column-wise one-dimensional discrete cosine transformed (1D DCTed) coefficients of the commercial image, where

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{Y}}, v = 0 \\ \sqrt{\frac{2}{Y}}, v = 1,2,...,Y-1 \end{cases}, \quad (12)$$

and $v = 0,1,\ldots,Y-1$. $F_{1DDCT}(x,v)$ can also be expressed in the polar form as

$$F_{1DDCT}(x,v) = |F_{1DDCT}(x,v)|S_{1DDCT}(x,v), \quad (13)$$

where $|F_{1DDCT}(x,v)|$ and $S_{1DDCT}(x,v)$ denote the amplitude components and the sign components, respectively. In the same way, the blue channel of the dummy image, $f_{db}(x,y)$, is transformed by

$$F_{db,1DDCT}(x,v) = \alpha(v)\sum_{y=0}^{Y-1} f_{db}(x,y)\cos\left(\frac{\pi}{Y}(y+\frac{1}{2})v\right), \quad (14)$$

where $F_{db,1DDCT}(x,v)$ denotes the 1D DCTed coefficients of the blue channel of the dummy image,

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{Y}}, v = 0 \\ \sqrt{\frac{2}{Y}}, v = 1,2,...,Y-1 \end{cases}, \quad (15)$$

and $v = 0,1,\ldots,Y-1$. $F_{db,1DDCT}(x,v)$ can also be expressed in the polar form as

$$F_{db,1DDCT}(x,v) = |F_{db,1DDCT}(x,v)|S_{db,1DDCT}(x,v), \quad (16)$$

where $|F_{db,1DDCT}(x,v)|$ and $S_{db,1DDCT}(x,v)$ denote the amplitude components and the sign components, respectively. The amplitude components of the commercial image, $|F_{1DDCT}(x,v)|$, are then combined with the 1D DCTed coefficient signs of the blue channel of the dummy image, $S_{db,1DDCT}(x,v)$. Finally, the one-dimensional inverse discrete cosine transform (1D IDCT) is used to obtain the stego-image, $f'_{db,1DDCT}(x,y)$, as described by Eq. (17).

$$f'_{db,1DDCT}(x,y) = \alpha(v)\sum_{y=0}^{Y-1}|F_{1DDCT}(x,v)|S_{db,1DDCT}(x,v)\cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right) \quad (17)$$

After that, the stego-image, $f'_{db,1DDCT}(x,y)$ is combined with the original red channel and the original green channel of the dummy image. The amplitude components of the dummy image, $|F_{db,1DDCT}(x,v)|$, can then be combined with the 1D DCTed coefficient signs of the commercial image, $S_{1DDCT}(x,v)$. The 1D IDCT is used to obtain another stego-image, $f'_{1DDCT}(x,y)$, as described by Eq. (18).

$$f'_{1DDCT}(x,y) = \alpha(v)\alpha(v)\sum_{y=0}^{Y-1}|F_{db,1DDCT}(x,v)|S_{1DDCT}(x,v)\cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right) \quad (18)$$

As mentioned above, the CP can send either the encrypted version of the stego-image, $f'_{1DDCT}(x,y)$, or the encrypted version of the sign components, $S_{1DDCT}(u,v)$, to the second TTP.

Besides the DCT-based systems, this paper also proposes HT-based systems. The HT works sim-ilarly to the Fourier transform: it takes a vector and maps it to its frequency components, which are the Walsh functions. Instead of sine waves in the Fourier transform, the Walsh functions are dis-crete "square waves." This makes it easier to com-pute. The description should start with the 1D HT. Firstly, the $X \times Y$-sized commercial image, $f(x, y)$, and the blue channel of the dummy image, $f_{db}(x, y)$, are transformed independently into the frequency domain by using the 1D HT. The transform is performed on each column of the images. Let $f_x = [f(x, 0)\ f(x, 1)\ f(x, 2)\ ...\ f(x, Y-1)]$ be column vec-tors of $f(x, y)$, where $x = 0, 1, 2, ..., X-1$, and $y = 0, 1, ..., Y-1$, and $F_{1DHT,u}$ be $1 \times Y$-sized Hadamard transformed (HTed) coefficients of f_x. Thus,

$$F_{1DHT,u} = H_m f_x, \quad (19)$$

where H_m denotes a $Y \times Y$-sized Hadamard matrix which can be defined recursively as

$$H_m = \frac{1}{\sqrt{2}}\begin{pmatrix} H_{m-1} & H_{m-1} \\ H_{m-1} & -H_{m-1} \end{pmatrix} = H_1 \otimes H_{m-1}, \quad (20)$$

where H_m represents a $2^m \times 2^m$ Hadamard matrix, and $2^m = Y$. $\otimes$ denotes the Kronecker product, and the 1×1 Hadamard transform H_0 is defined by the identity $H_0 = 1$. The $\frac{1}{\sqrt{2}}$ is a normalization that is sometimes omitted. Thus, other than this normalization factor, the Hadamard matrices are made up entirely of 1 and -1. When the transform is applied across the entire X columns, we obtain

$$F_{1DHT}(u,v) = H_m f_{x,y}, \quad (21)$$

where $F_{1DHT}(u, v)$ denotes the one-dimensional Hadamard transformed (1D HTed) coefficients of the commercial image, and $u = 0, 1, ..., X-1$, $v = 0, 1, ..., Y-1$, $x = 0, 1, 2, ..., X-1$, and $y = 0, 1, ..., Y-1$. $F_{1DHT}(u,v)$ can also be expressed in the polar form as

$$F_{1DHT}(u,v) = |F_{1DHT}(u,v)|S_{1DHT}(u,v), \quad (22)$$

where $|F_{1DHT}(u,v)|$ and $S_{1DHT}(u,v)$ denote Hadamard amplitude components and Hadamard sign components, respectively. In the same way, the blue channel of the dummy image, $f_{db}(x,y)$, is transformed by applying the 1D HT to all X columns as described by Eq. (23)

$$F_{db,1DHT}(u,v) = H_m f_{db}(x,y), \quad (23)$$

where $F_{db,1DHT}(u,v)$ denotes the 1D HTed coefficients of the blue channel of the dummy image and can also be expressed in the polar form as

$$F_{db,1DHT}(u,v) = |F_{db,1DHT}(u,v)|S_{db,1DHT}(u,v), \quad (24)$$

where $|F_{db,1DHT}(u,v)|$ and $S_{db,1DHT}(u,v)$ denote amplitude components and sign components of the blue channel of the dummy image, respectively. The 1D HTed amplitude components of the commercial image, $|F_{1DHT}(u,v)|$, are then combined with the 1D HTed coefficient signs of the blue channel of the dummy image, $S_{db,1DHT}(u,v)$. Finally, the one-dimensional inverse Hadamard transform (1D IHT) is used to generate the stego-image, $f'_{db,1DHT}(x,y)$, as described by Eq. (25).

$$f'_{db,1DHT}(x,y) = \frac{1}{y}H_m^T(|F_{1DHT}(u,v)|S_{db,1DHT}(u,v)), \quad (25)$$

where H_m^T denotes the transpose of H_m. The transpose of H_m is closely related to its inverse. In fact: $H_m H_m^T = Y I_Y$, where I_Y is the $Y \times Y$ identity matrix. After that, the stego-image, $f'_{db,1DHT}(x,y)$ is combined with the original red channel and the original green channel of the dummy image. Another stego-image, $f'_{1DHT}(x,y)$, can also be calculated by Eq. (26).

$$f'_{1DHT}(x,y) = \frac{1}{Y}H_m^T|F_{db,1DHT}(u,v)|S_{1DHT}(u,v) \quad (26)$$

As well as the DCT-based systems, the CP can send either the encrypted version of the stego-image, $f'_{1DHT}(x,y)$, or the encrypted version of the sign components, $S_{1DHT}(u,v)$, to the second TTP.

In order to apply the Hadamard transform to two-dimensional images, the 1D HT is applied across X columns and then Y rows as described by Eq. (27).

$$F_{2DHT}(u,v) = (H_n(H_m f(x,y))^T)^T, \quad (27)$$

where $F_{2DHT}(u,v)$ denotes the two-dimensional

Hadamard transformed (2D HTed) coefficients of the commercial image, $2^n = X$, and $(\ )^T$ defines the matrix transpose operator. The couple matrix-transpose operation is needed for row transformation. However, by considering the matrix property: $(BA^T)^T = AB$, Eq. (28) can be rewritten as

$$F_{2DHT}(u,v) = (H_m f(x,y))H_n. \quad (28)$$

$F_{2DHT}(u,v)$ can also be expressed in the polar form as

$$F_{2DHT}(u,v) = |F_{2DHT}(u,v)|S_{2DHT}(u,v), \quad (29)$$

where $|F_{2DHT}(u,v)|$ and $S_{2DHT}(u,v)$ define amplitude components and sign components, respectively. In the same way, the blue channel of the dummy image, $f_{db}(x,y)$, is transformed by using 2D HT as described by Eq. (30)

$$F_{db,2DHT}(u,v) = (H_m f_{db}(x,y))H_n, \quad (30)$$

where $F_{db,2DHT}(u,v)$ denotes the 2D HTed coefficients of the blue channel of the dummy image which can also be expressed in the polar form as

$$F_{db,2DHT}(u,v) = |F_{db,2DHT}(u,v)|S_{db,2DHT}(u,v), \quad (31)$$

where $|F_{db,2DHT}(u,v)|$ and $S_{db,2DHT}(u,v)$ denote amplitude components and sign components of the blue channel of the dummy image, respectively. The 2D HTed amplitude components of the commercial image, $|F_{2DHT}(u,v)|$, are then combined with the 2D HTed coefficient signs of the blue channel of the dummy image, $S_{db,2DHT}(u,v)$. Finally, the two-dimensional inverse Hadamard transform (2D IHT) is used to obtain the stego-image, $f'_{db,2DHT}(x,y)$, as described by Eq. (32).

$$f'_{db,2DHT}(x,y) = \tfrac{1}{XY}(H_m^T(|F_{2DHT}(u,v)|S_{db,2DHT}(u,v)))H_n^T, \quad (32)$$

where H_n^T denotes the transpose of H_n, where $H_n H_n^T = XI_X$, and I_X is the $X \times X$ identity matrix. After that, the stego-image, $f'_{db,2DHT}(x,y)$ is combined with the original red channel and the original green channel of the dummy image. Another stego-image, $f'_{2DHT}(x,y)$ can also be calculated by Eq. (33).

$$f'_{2DHT}(x,y) = \tfrac{1}{XU}(H_m^T(|F_{db,2DHT}(u,v)|S_{2DHT}(u,v)))H_n^T, \quad (33)$$

Again and again, the CP can send either the encrypted version of the stego-image, $f'_{2DHT}(x,y)$, or the encrypted version of the sign components, $S_{2DHT}(u,v)$, to the second TTP.

It is clearly seen that the hiding capacity of the proposed systems is as high as that of the existing systems proposed in [14], i.e., a whole gray image can be embedded into a cover image of the same size.

By using the proposed systems, even if intruders do not notice that the algorithm is being used, but the hidden information in the image is stolen accidentally, the intruders still cannot recognize the secret image, because what the intruders see are only amplitude components of the secret image which do not reveal any recognizable details of the image. The recognizable details of the image are only in the phase components (the discrete cosine transformed (DCTed)/HTed coefficient signs) of the image which are transferred to the consumer by the second TTP as a decryption key. In other words, the proposed method can be considered as a hybrid of steganography and cryptography. That is the proposed systems have double security. Moreover, the coefficient signs are encrypted by the CP before sending. Therefore, the proposed systems could even be considered triple-locked systems. Anyway, to encrypt the coefficient signs, any existing binary encryption method could be used.

Another concern is the complexity. A real-time system is expected for the focused application because it is commercial. The complexity of the column-wise 1D DCT/1D HT for an $X \times Y$-sized image is $O((Y^2)X)$ or $O((Y\log_2 Y)X)$ with fast algorithms, whereas that of the 2D DCT/2D HT is $O((XY)^2)$ or $O((XY)log_2(XY))$ with fast algorithms. While the complexities of the deep learning-based methods are drastically higher. For example, the complexity of training a general 2D CNN for each layer and each epoch is $O(N^2K^2XY)$, where N denotes the channel size of the convolution, K denotes the kernel size, and X and Y denote the output width and height, respectively. It is clearly seen that the proposed systems are much faster than the deep learning-based methods.

3.2 Fingerprint embedding process

To protect the copyright of the image, the consumer ID is embedded into the stego-image by the first TTP using the digital fingerprinting technique. Since there are rooms to enhance the fingerprinting performances from [14], this paper proposes a fingerprinting method that only embeds the fingerprint into only the "amplitude components." The fingerprint is embedded into the blue channel of the stego-image, and the transform used is the corresponding transform used in the steganography process. Since the fingerprint is embedded into the amplitude components without changing the coefficient signs. Once the consumer reconstructs the fingerprinted commercial image, the coefficient signs of the stego-image (actually the coefficient sign of the blue channel of the dummy image) are discarded, and then the am-

plitudes with the embedded fingerprint are combined with the signs of the commercial image received from the second TTP. Since the fingerprint is not embedded into the coefficient signs at all, the fingerprint is as well not degraded by the image reconstruction process at all. Although the fingerprinting method used can be any method that embeds the fingerprint into only the "amplitude components" of the commercial image, the digital fingerprinting method should still be considered and chosen cautiously in terms of fingerprinted image quality for the consumer purpose. This paper then proposes a steganography compatible-and-high performance fingerprinting method. Firstly, two M-sequences that differ from each other used for different single bits of the fingerprint are generated. For example, if $W = 0$ then M-sequence = 0101, and if $W = 1$ then M-sequence = 1100, where W denotes the fingerprint bit string. Each M-sequence is embedded into an 8×8-pixel divided block of the DCTed/HTed coefficients of the blue channel of the stego-image by linearly scaling and adding it to four bottom-right coefficients of the block as

$$C'_l = sgn(C_l)(|C_l| + \delta W_{b,l}), \tag{34}$$

where C'_l denotes the l^{th} coefficient containing the fingerprint in the block, C_l denotes the l^{th} original efficient, δ is a scaling factor for data hiding, and $W_{b,l} \in \{0, 1\}$ denotes the l^{th} chip of the M-sequence for b^{th} bit in a coefficient block, and $l = 1, 2, 3, ..., L$, where L denotes the number of chip of the M-sequence.

3.3 Steganalysis-and-reconstruction process

After the image fingerprinting process is done, the fingerprinted stego-image is sent from the first TTP to the consumer. The consumer can reconstruct the fingerprinted commercial image by combining what he/she receives from the first TTP and the second TTP. Firstly, the consumer transforms the image received from the first TTP using the corresponding transform. The coefficient signs of the image (the coefficient signs of the blue channel of the dummy image) are subsequently discarded. The extracted amplitude components with the embedded fingerprint are then combined with the coefficient signs of the commercial image received from the second TTP. Finally, the fingerprinted commercial image is reconstructed by using the inverse transformation.

For image steganalysis performance, it can be easily understood by mathematics that the DCT and the HT are "reversible." Therefore, the amplitude components of the commercial image can be extracted from the stego-image with 100% correctness. This confirms that the proposed systems have high steganalysis robustness. However, the commercial image quality is still degraded by the fingerprinting process which is evaluated in Section 4.2.

3.4 Fingerprint extraction process

In cases where the CP would like to check the consumer ID from the suspicious image, the CP requests the first TTP to extract the fingerprint from that image. Firstly, the fingerprinted commercial image is transformed into the frequency domain using the 2D DCT/1D DCT/2D HT/1D HT by the CP. The amplitudes are then extracted. Subsequently, the coefficient signs of the dummy image are combined with the extracted amplitudes. After that, the inverse transformation is performed. The output image of this process which is a stego-image is then sent from the CP to the first TTP. Finally, the first TTP extracts the fingerprint from the blue channel of the image. Therefore, fingerprint extraction performance is very important. If the fingerprint is extracted incorrectly, the CP cannot know the correct consumer ID to sue that consumer. The proposed fingerprinting method is then evaluated in terms of fingerprint extraction performance in Section 4.2 as well.

4. RESULTS AND DISCUSSION

Four proposed systems are compared with two existing systems [14] in terms of security or stego-image visual quality, fingerprinted image quality, and fingerprint extraction performance. Five 512 × 512-pixel gray test images (for commercial images) and five 512 × 512-pixel color test images (for dummy images) were used in the experiment. That is there are 25 pairs of the cover image and the secret image. Basically, the cover image and the secret image should not be the same, but we also show those cases just for reference.

4.1 Security

The stego-image quality does not affect the quality of the reconstructed commercial image. However, it helps in protecting the consumer's privacy against the first TTP and the malicious third party (s), which is called "security." Peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used as indicators for security measurement of each dummy image-and-stego-image pair. The PSNR can be calculated by

$$PSNR = 20 \log_{10} \left(\frac{MAX_{I_d}}{\sqrt{MSE}} \right), \tag{35}$$

where $PSNR$, MAX_{I_d}, and MSE denote the PSNR, the maximum intensity of the dummy image, and the mean square error (MSE) which can be calculated by Eq. (36).

$$MSE = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [I_d(x, y) - I_s(x, y)]^2, \tag{36}$$

where $I_d(x,y)$ and $I_s(x,y)$ denote the dummy image and the stego-image, respectively. In cases of the existing systems using gray cover images, the color dummy images were converted into gray cover images. Therefore, Eq. (36) can then be used instantly. For the proposed systems using color cover images, the MSE can be calculated by using the Euclidean distance [50] instead which is described by Eq. (37).

$$MSE_{Euclidean} = \frac{1}{XY}\sum_{x=0}^{X}\sum_{y=0}^{Y}(R_d(x,y)-R_s(x,y))^2 + (G_d(x,y)-G_s(x,y))^2 + (B_d(x,y) - B_s(x,y))^2, \quad (37)$$

where $MSE_{Euclidean}$ denotes the MSE calculated from the Euclidean distances of red, green, and blue values between the dummy image and the stego-image. R_d, G_d, and B_d denote red, green, and blue values of the dummy image, respectively, whereas R_s, G_s, and B_s denote red, green, and blue values of the stego-image, respectively. By the way, SSIM can be calculated by

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (38)$$

where $SSIM$ denotes the $SSIM$, μ_x denotes the average of x, μ_y denotes the average of y, σ_x^2 denotes the variance of x, σ_y^2 denotes the variance of y, σ_{xy} denotes the covariance of x and y, and c_1 and c_2 are two variables to stabilize the division with weak denominator. $c_1 = (k_1R)^2$ and $c_2 = (k_2R)^2$, where R denotes the dynamic range of the pixel values (typically $2^{\#bitsperpixel}$-1), and $k_1 = 0.01$ and $k_2 = 0.03$ by default. The SSIM index then ranges from -1 to +1.

The calculated PSNRs and SSIMs are shown in Table 2 and Table 3, respectively. The results show that the proposed systems are superior to the existing systems proposed in [14] in terms of both PSNR and SSIM. Even though the PSNRs achieved by the proposed systems are not quite high with drastic changes in the blue channels of the cover images, the visual qualities of the stego-images are still high in terms of SSIM. The SSIM indicator is a newer measurement tool that is designed based on three factors, i.e., luminance, contrast, and structure, which is more suitable for the workings of the human visual system [51]. While the PSNR indicator has been shown to perform poorly compared to other quality indicators when it comes to estimating the qualities of images and particularly videos as perceived by humans [52]-[53]. However, the PSNR value for an RGB image could be calculated using other formulas such as 1) averaging the errors of R, G, and B channels, and then calculating the PSNR, or 2) averaging the PSNRs of R, G, and B channels. By using the first choice, the PSNR value would be increased by about 10 dB for our cases, and by using the second choice, it might be difficult to calculate the PSNR if there is no error in some channel (s) causing the PSNR value (s) of that (those) channel (s) to infinite. Table 3 compares the SSIM values achieved by the proposed systems and the existing systems. The proposed 2D DCT, 1D DCT, 2D HT, and 1D HT systems have achieved the average SSIMs at 0.8398, 0.8319, 0.8384, and 0.8307, respectively. These high values imply that the stego-images generated by the proposed systems look like the dummy images instead of the commercial images or the secret images. While the existing 2D DCT and 1D DCT systems have achieved the average SSIMs at 0.2581 and 0.2100, respectively. These low values imply that the stego-images generated by the existing systems do not look like the dummy images. Furthermore, a subjective evaluation could also be done. Fig. 9 shows the results of the first image pair, 'Pepper-Couple.' The results show that the stego-images generated by both proposed systems and existing systems cannot be recognized easily as the commercial images by human eyes. However, it is clearly seen that the stego-images generated by the proposed systems almost perfectly look like the dummy image, and there are no visual artifacts found in the stego-images (See Fig. 9 (e)-(h).), while the stego-images generated by the existing systems [14] look suspicious with low image qualities and are quite difficult to be recognized as the dummy image (See Fig. 9 (c)-(d).).

Table 2: *Peak signal-to-noise ratio (PSNR) [dB].*

Cover Image-Secret Image	2D DCT [14]	1D DCT [14]	Proposed 2D DCT	Proposed 1D DCT	Proposed 2D HT	Proposed 1D HT
Pepper-Couple	14.7855	14.9690	21.6934	21.7388	21.8602	21.7982
Pepper-Boat	15.5347	15.7138	20.9292	21.0050	20.7884	20.9645
Pepper-Rose	14.6278	14.1117	21.1142	20.5870	21.1847	20.5833
Pepper-Home	13.6476	13.6982	19.3260	19.1697	19.4276	19.1183
Pepper-Car	15.3696	15.3745	19.5049	19.5528	19.5473	19.5412
Fruit-Couple	14.1735	14.3859	21.6189	21.5394	21.8373	21.4558
Fruit-Boat	14.6355	14.7793	21.0103	20.8052	21.0734	20.7292
Fruit-Rose	13.9836	13.2604	21.5089	21.0076	21.7548	21.0435
Fruit-Home	12.9170	12.5515	19.3852	19.2762	19.7282	19.2176
Fruit-Car	14.9239	14.8177	19.9497	19.9062	20.1634	19.9330
Splash-Couple	12.1432	12.1186	20.9528	20.9507	21.3635	20.9670
Splash-Boat	12.6466	12.5940	20.6152	20.7844	20.8041	20.8994
Splash-Rose	12.5992	12.8787	19.2120	19.1870	19.1968	19.1216
Splash-Home	13.4409	13.0254	19.5594	18.6003	19.4517	18.5815
Splash-Car	12.2001	11.8189	17.2216	17.0341	17.2639	17.0438
Cat-Couple	14.9464	13.9641	22.1448	21.3477	22.1191	21.3466
Cat-Boat	14.9881	14.4587	21.2126	20.8739	21.1685	20.8311
Cat-Rose	13.3188	12.4645	20.7314	20.2172	21.1125	20.1920
Cat-Home	13.0269	12.4729	19.3277	18.9940	19.5681	18.9555
Cat-Car	15.2354	13.8638	19.7433	19.2479	19.7526	19.2251
Lena-Couple	14.7880	13.8030	21.6039	20.4285	21.6598	20.6263
Lena-Boat	15.4820	14.8231	20.6202	20.1612	20.6094	20.1429
Lena-Rose	16.3137	14.9695	20.7010	20.0922	21.1003	20.2975
Lena-Home	14.3354	12.8889	19.0128	18.5256	19.0416	18.5421
Lena-Car	16.6420	14.0913	19.0647	18.4808	19.0610	18.5185
Min	12.1432	11.8189	17.2216	17.0341	17.2639	17.0438
Max	16.6420	15.7138	22.1448	21.7388	22.1191	21.7982
Average	14.26822	13.7559	20.31056	19.98054	20.42553	19.98702

Table 3: *Structural similarity index (SSIM).*

Cover Image-Secret Image	2D DCT [14]	1D DCT [14]	Proposed 2D DCT	Proposed 1D DCT	Proposed 2D HT	Proposed 1D HT
Pepper-Couple	0.2466	0.2090	0.8837	0.8840	0.8862	0.8840
Pepper-Boat	0.3574	0.3589	0.8599	0.8614	0.8519	0.8595
Pepper-Rose	0.1955	0.1496	0.8792	0.8734	0.8809	0.8732
Pepper-Home	0.2588	0.2565	0.7263	0.7385	0.7178	0.7320
Pepper-Car	0.2374	0.2177	0.8699	0.8714	0.8691	0.8701
Fruit-Couple	0.2506	0.1786	0.8808	0.8741	0.8848	0.8716
Fruit-Boat	0.3805	0.3110	0.8577	0.8495	0.8555	0.8476
Fruit-Rose	0.2055	0.1415	0.8763	0.8643	0.8782	0.8647
Fruit-Home	0.2682	0.2171	0.7222	0.7092	0.7133	0.7027
Fruit-Car	0.2641	0.1959	0.8760	0.8780	0.8805	0.8778
Splash-Couple	0.1716	0.1273	0.8901	0.8833	0.8922	0.8832
Splash-Boat	0.3585	0.2769	0.8685	0.8672	0.8672	0.8663
Splash-Rose	0.1318	0.0955	0.8388	0.8399	0.8364	0.8388
Splash-Home	0.2565	0.1986	0.7359	0.7152	0.7306	0.7115
Splash-Car	0.1849	0.1451	0.7995	0.7931	0.7958	0.7906
Cat-Couple	0.2007	0.1658	0.8885	0.8733	0.8867	0.8716
Cat-Boat	0.3813	0.3181	0.8672	0.8527	0.8611	0.8522
Cat-Rose	0.1477	0.1112	0.8673	0.8531	0.8712	0.8559
Cat-Home	0.2512	0.2242	0.7128	0.7006	0.7094	0.6965
Cat-Car	0.2117	0.1754	0.8680	0.8630	0.8704	0.8599
Lena-Couple	0.2841	0.2072	0.8891	0.8694	0.8888	0.8727
Lena-Boat	0.3233	0.2961	0.8524	0.8438	0.8508	0.8438
Lena-Rose	0.3089	0.1999	0.8908	0.8663	0.8928	0.8692
Lena-Home	0.2665	0.2324	0.7415	0.7132	0.7319	0.7131
Lena-Car	0.3082	0.2401	0.8535	0.8588	0.8575	0.8583
Min	0.1318	0.0955	0.7128	0.7006	0.7094	0.6965
Max	0.3813	0.3589	0.8908	0.8840	0.8928	0.8840
Average	0.2581	0.2100	0.8398	0.8319	0.8384	0.8307

Fig. 10 shows the second stego-images sent to the second TTP and generated by various proposed systems for the first image pair, 'Pepper-Couple.' These images are gray and may reveal some details of the commercial images because of phase sign characteristics. Instead of sending the stego-images to the second TTP, sending only the signs to the second TTP as a binary image could potentially achieve higher security. Fig. 11 shows the coefficient signs of the image 'Couple' as binary images generated by various proposed transforms. Besides security purposes, using a binary image instead helps in saving memory usage and communication bandwidth consumption.

4.2 Fingerprinting performance

Two 4-bit M-sequences are used in the experiment: if $W = 0$ then M-sequence = 0101, and if $W = 1$ then M-sequence = 1100, and $L = 4$. $\delta = 1, 2, ..., 7$.

Fig. 12 compares the results of the proposed systems with those of the existing systems [14]. Since the different fingerprinting method is used in [14], the graphs of fingerprinted image qualities in terms of PSNR versus correct fingerprint extracting rates are plotted instead of the fingerprinting parameters. The results confirm that the proposed systems yielded much better results than those of the existing systems. In terms of PSNR, the proposed systems have achieved 46.2803 dB and 44.1900 dB where $\delta = 7$, and 63.1823 dB and 61.0900 dB where $\delta = 1$, for the proposed DCT systems and the proposed HT systems,

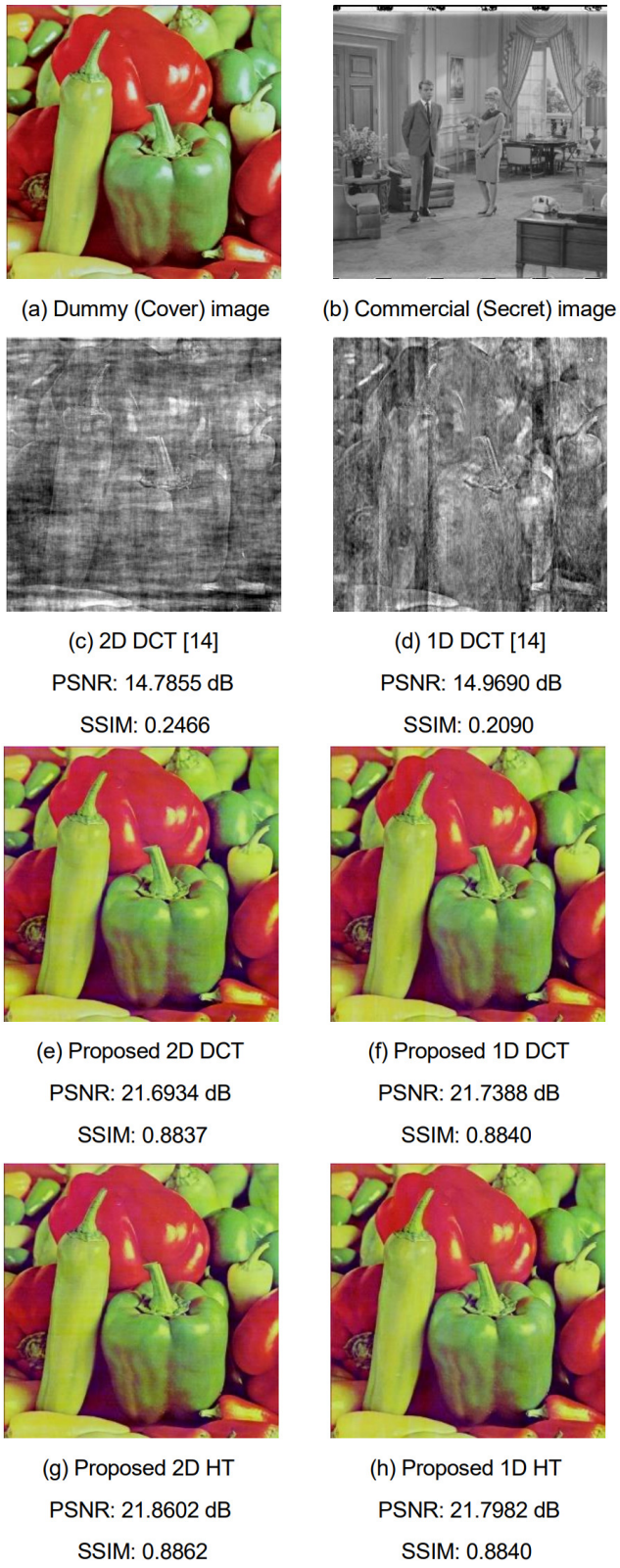

Fig.9: *Stego-images sent to the first TTP and generated by various systems.*

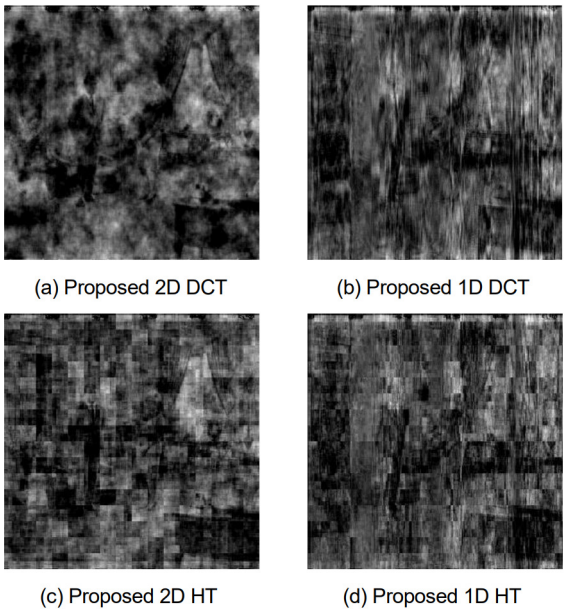

Fig.10: *Stego-images sent to the second TTP and generated by various proposed systems.*

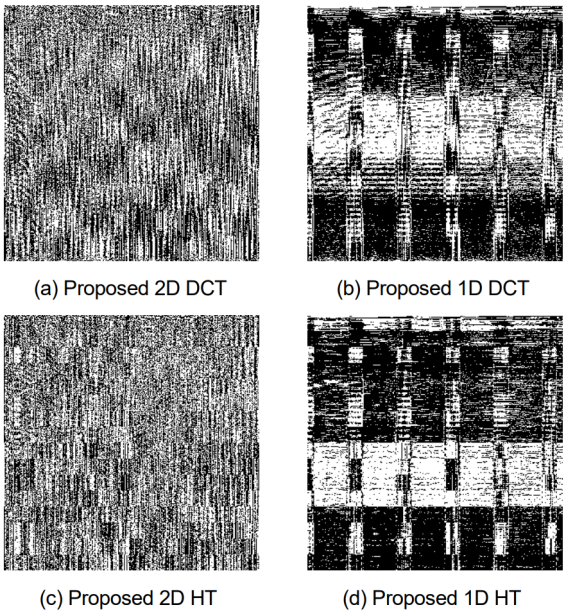

Fig.11: *Binary images sent to the second TTP and generated by various proposed systems.*

respectively. While the existing systems achieved only 25.5165 dB and 26.0655 dB where $\delta = 7$, and 46.3902 dB and 46.0637 dB where $\delta = 1$, for the existing 2D DCT system and the existing 1D DCT system, respectively. Among the four proposed systems, using the DCT achieves better fingerprinted image quality than using the HT, while the correct fingerprint extracting rate is 100%, where $\delta = 1, 2, \ldots, 7$, for all proposed systems. That is the results confirm the compatibility of the proposed image steganography method and the proposed amplitude-based fingerprinting method.

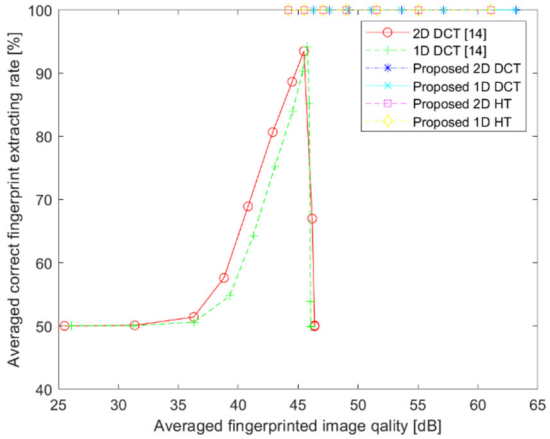

Fig.12: *Fingerprinted image qualities versus correct fingerprint extracting rates.*

4.3 Processing time

The processing time of the proposed systems and the existing systems are not significantly different, because each of them applies either the 2D DCT, the 1D DCT, the 2D HT, or the 1D HT. By using MATLAB toolboxes for transforming, only about 0.2 seconds is spent for a 512 × 512-pixel image for the image steganography process, where Intel® Core™ i5, 256 GB SSD, 8 GB RAM device is used.

5. CONCLUSIONS

This paper proposes image steganography-based copyright- and privacy-protected image trading systems. In the systems, there are the CP, the consumer, the first TTP, and the second TTP. The first TTP is responsible for image copyright protection. He/She embeds the consumer ID into the stego-image generated by the CP. In the steganography process, either the DCT or the HT is used with a color dummy image. After applying image transformation to both images separately, the coefficient signs of the commercial image are replaced by the coefficient signs of the dummy image pixel-by-pixel so that the inversely transformed commercial image looks like the dummy image instead of the original one. To protect the copyright of the image, the consumer ID is embedded into the amplitude components of the commercial image by the first TPP using the digital fingerprinting technique. Once the consumer receives the fingerprinted image from the first TTP and the coefficient signs of the commercial image from the CP, the consumer reconstructs the fingerprinted commercial image without losing the hidden fingerprint. As a result, the commercial image is transferred from the CP to the consumer with both image copyright protection and consumer's privacy protection.

The experimental results show that the commercial images cannot be easily recognized from the stego-images generated by the proposed systems.

Moreover, the stego-images generated by the proposed systems are not suspicious unlike the stego-images generated by the existing systems. The experimental results also confirm the compatibility of the proposed image steganography method and the amplitude-based fingerprinting method. Compared with the existing systems, the proposed systems achieved significantly higher fingerprinted image qualities and 100% correct fingerprint extracting rates. Table 4 concludes the performances of the proposed systems and the existing methods/systems.

The challenging task for future work could be embedding a "color" commercial image into a color dummy image.

Table 4: *Performance summary of the proposed systems and the existing methods/systems.*

	LSB	**CNN**	**GAN**	**[14]**	**Proposed**
Hiding Capacity	X	O	X	O	O
Security	X	O	O	X	O
Robustness	O	X	X	O	O
Complexity	O	X	X	O	O
Fingerprinting Performance	X	X	X	O*	O

ACKNOWLEDGMENT

This research was funded by King Mongkut's University of Technology North Bangkok. Contract no. KMUTNB-64-DRIVE-30.

References

[1] R. L. Lagendijk, Z. Erkin, and M. Barni, "Encrypted signal processing for privacy protection: conveying the utility of homomorphic encryption and multiparty computation," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 82-105, Jan. 2012.

[2] M. Kuribayashi, "Recent fingerprinting techniques with cryptographic protocol," *Signal Processing*, 2010.

[3] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, Digital watermarking and steganography, 2nd ed., Morgan Kaufmann Publishers, 2008.

[4] B. Schneier, *Applied cryptography: protocols, algorithms, and source code in C*, 2nd ed., John Wiley & Sons, 1996.

[5] Y. Sengoku and H. Hioki, "A model of privacy and copyright-aware image trading system based on adaptive image segmentation and digital watermarking," *Proceeding of 27th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC 2012)*, pp. D-W1-02, 2012.

[6] Y. Sengoku and H. Hioki, "An image segmentation method for privacy and copyright-aware image trading system," *IEICE Technical Report*, vol. 111, no. 496, EMM2011-67, pp. 19-24, Mar. 2012.

[7] M. Okada, Y. Okabe, and T. Uehara, "A web-based privacy-secure content trading system for small content providers using semi-blind digital watermarking," *Proceeding of 7th IEEE Consumer Communications & Networking Conference (CCNC 2010)*, 2010.

[8] S. Liu, M. Fujiyoshi, and H. Kiya, "An image trading system using amplitude-only images for privacy- and copyright-protection," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E96-A, pp. 1245-1252, Jun. 2013.

[9] W. Sae-Tang, M. Fujiyoshi, and H. Kiya, "A generation method of amplitude-only images with low intensity ranges," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E96-A, no. 6, pp. 1323-1330, Jun. 2013.

[10] W. Sae-Tang, S. Liu, M. Fujiyoshi, and H. Kiya, "1D Frequency transformation-based amplitude-only images for copyright- and privacy-protection in image trading systems," *ECTI-CIT*, vol. 8, No. 2, Nov. 2014.

[11] W. Sae-Tang, M. Fujiyoshi, and H. Kiya, "Evaluation of amplitude-only images for copyright- and privacy-protected image trading systems," *Proceeding of 29th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC 2014)*, pp. 113-116, 2014.

[12] W. Sae-Tang and H. Kiya, "Hadamard transform-based amplitude-only images for image trading systems," *2016 International Workshop on Advanced Image Technology (IWAIT 2016)*, pp. 3C.5, 2016.

[13] W. Sae-Tang, M. Fujiyoshi, and H. Kiya, "Encryption-then-compression-based copyright- and privacy-protected image trading system," *2017 International Conference on Advances in Image Processing (ICAIP 2017)*, pp. 66-71, 2017.

[14] W. Sae-Tang, M. Fujiyoshi, and H. Kiya, "A new copyright- and privacy-protected image trading system using a novel steganography-based visual encryption scheme," *ECTI-EEC*, vol. 17, no. 1, pp. 95-107, Feb. 2019.

[15] N. F. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," *Computer*, vol. 31, no. 2, pp. 26-34, Feb. 1998.

[16] S. Gupta, G. Gujral, and N. Aggarwal, "Enhanced least significant bit algorithm for image steganography," *Int. J. Comput. Eng. Manage.*, vol. 15, no. 4, pp. 40-42, 2012.

[17] R. J. Mstafa, K. M. Elleithy, and E. Abdelfattah, "A robust and secure video steganography method in DWT-DCT domains based on multiple object tracking and ECC," *IEEE Access*, vol.

5, pp. 5354-5365, 2017.

[18] G. Swain, "Very high capacity image steganography technique using quotient value differencing and LSB substitution," *Arabian J. Sci. Eng.*, vol. 44, no. 4, pp. 2995-3004, Apr. 2019.

[19] A. Qiu, X. Chen, X. Sun, S. Wang, and W. Guo, "Coverless image steganography method based on feature selection," *J. Inf. Hiding Privacy Protection*, vol. 1, no. 2, p. 49, 2019.

[20] R. D. Rashid and T. F. Majeed, "Edge based image steganography: Problems and solution," *Proceeding of 2019 International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, pp. 1-5, Mar. 2019.

[21] X. Liao, J. Yin, S. Guo, X. Li, and A. K. Sangaiah, "Medical JPEG image steganography based on preserving inter-block dependencies," *Comput. Electr. Eng.*, vol. 67, pp. 320-329, Apr. 2018.

[22] W. Lu, Y. Xue, Y. Yeung, H. Liu, J. Huang, and Y. Shi, "Secure halftone image steganography based on pixel density transition," *IEEE Trans. Dependable Secure Comput.*, Aug. 6, 2019.

[23] Y. Zhang, C. Qin, W. Zhang, F. Liu, and X. Luo, "On the fault-tolerant performance for a class of robust image steganography," *Signal Process.*, vol. 146, pp. 99-111, May 2018.

[24] H. M. Sidqi and M. S. Al-Ani, "Image steganography: Review study," *Proceeding of Int. Conf. Image Process., Comput. Vis., Pattern Recognit. (IPCV)*, pp. 134-140, 2019.

[25] P. Wu, Y. Yang, and X. Li, "Image-into-image steganography using deep convolutional network," *Proceeding of Pacific Rim Conf. Multimedia. Cham*, Switzerland: Springer, pp. 792-802, 2018.

[26] P. Wu, Y. Yang, and X. Li, "StegNet: Mega image steganography capacity with deep convolutional network," *Future Internet*, vol. 10, no. 6, p. 54, Jun. 2018.

[27] X. Duan, K. Jia, B. Li, D. Guo, E. Zhang, and C. Qin, "Reversible image steganography scheme based on a U-Net structure," *IEEE Access*, vol. 7, pp. 9314-9323, 2019.

[28] T. P. Van, T. H. Dinh, and T. M. Thanh, "Simultaneous convolutional neural network for highly efficient image steganography," *Proceeding of 19th Int. Symp. Commun. Inf. Technol. (ISCIT)*, pp. 410-415, Sep. 2019.

[29] R. Rahim and S. Nadeem, "End-to-end trained CNN encoder-decoder networks for image steganography," *Proceeding of Eur. Conf. Comput. Vis. (ECCV)*, pp. 1-6, 2018.

[30] Z. Wang, N. Gao, X. Wang, J. Xiang, and G. Liu, "STNet: A style transformation network for deep image steganography," *Proceeding of Int. Conf. Neural Inf. Process. Cham*, Switzerland: Springer, pp. 3-14, 2019.

[31] K. Yang, K. Chen, W. Zhang, and N. Yu, "Provably secure generative steganography based on autoregressive model," *Proceeding of Int. Workshop Digit. Watermarking. Cham*, Switzerland: Springer, pp. 55-68, 2018.

[32] R. Zhang, S. Dong, and J. Liu, "Invisible steganography via generative adversarial networks," *Multimedia Tools Appl.*, vol. 78, no. 7, pp. 8559-8575, Apr. 2019.

[33] S. Islam, A. Nigam, A. Mishra, and S. Kumar, "VStegNET: Video steganography network using spatio-temporal features and microbottleneck," *Proceeding of BMVC*, p. 274, Sep. 2019.

[34] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Proceeding of Adv. Neural Inf. Process. Syst.*, pp. 2672-2680, 2014.

[35] D. Volkhonskiy, B. Borisenko, and E. Burnaev, "Generative adversarial networks for image steganography," *Proceeding of ICRL Conf.*, 2016.

[36] D. Volkhonskiy, I. Nazarov, and E. Burnaev, "Steganographic generative adversarial networks," *Proceeding of 12th Int. Conf. Mach. Vis. (ICMV)*, vol. 11433, Art. no. 114333M, 2020.

[37] D. J. Im, C. D. Kim, H. Jiang, and R. Memisevic, "Generating images with recurrent adversarial networks," 2016., arXiv:1602.05110. [Online]. Available: `http://arxiv.org/abs/1602.05110`

[38] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "SSGAN: Secure steganography based on generative adversarial networks," *Proceeding of Pacific Rim Conf. Multimedia. Cham*, Switzerland: Springer, pp. 534-544, 2017.

[39] J. Yang, K. Liu, X. Kang, E. K. Wong, and Y.-Q. Shi, "Spatial image steganography based on generative adversarial network," 2018., arXiv:1804.07939. [Online]. Available: `http://arxiv.org/abs/1804.07939`

[40] J. Yang, D. Ruan, J. Huang, X. Kang, and Y.-Q. Shi, "An embedding cost learning framework using GAN," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 839-851, 2020.

[41] W. Tang, S. Tan, B. Li, and J. Huang, "Automatic steganographic distortion learning using a generative adversarial network," *IEEE Signal Process. Lett.*, vol. 24, no. 10, pp. 1547-1551, Oct. 2017.

[42] J. Zhu, R. Kaplan, J. Johnson, and L. Fei-Fei, "Hidden: Hiding data with deep networks," *Proceeding of Eur. Conf. Comput. Vis. (ECCV)*, pp. 657-672, 2018.

[43] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks,"

Proceeding of IEEE Int. Conf. Comput. Vis. (ICCV), pp. 2223-2232, Oct. 2017.

[44] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," *Proceeding of Int. Conf. Mach. Learn.*, pp. 2642-2651, 2017.

[45] Z. Zhang, G. Fu, J. Liu, and W. Fu, "Generative information hiding method based on adversarial networks," *Proceeding of Int. Conf. Comput. Eng. Netw. Cham, Switzerland: Springer*, pp. 261-270, 2018.

[46] M.-M. Liu, M.-Q. Zhang, J. Liu, Y.-N. Zhang, and Y. Ke, "Coverless information hiding based on generative adversarial networks," 2017., arXiv:1712.06951. [Online]. Available: `http://arxiv.org/abs/1712.06951`

[47] X. Duan, H. Song, C. Qin, and M. K. Khan, "Coverless steganography for digital images based on a generative model," *Comput., Mater. Continua*, vol. 55, no. 3, pp. 483-493, Jul. 2018.

[48] Z. Wang, N. Gao, X. Wang, X. Qu, and L. Li, "SSteGAN: Self-learning steganography based on generative adversarial networks," *Proceeding of Int. Conf. Neural Inf. Process. Cham*, Switzerland: Springer, pp. 253-264, 2018.

[49] R. K. Pradhan, "Does Rayleigh scattering explain the Blueness of Sky?," *Science*, vol. 28, no. 31, 2015.

[50] L. Liberti, C. Lavor, N. Maculan, and A. Mucherino, "Euclidean distance geometry and applications," *SIAM review*, vol. 56, No. 1, pp. 3-69, 2014.

[51] D. R. I. M. Setiadi, "PSNR vs SSIM: imperceptibility quality assessment for image steganography," *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 8423-8444, 2021.

[52] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electronics Letters*, vol. 44, no. 13, pp. 800-801, 2008.

[53] Q. Huynh-Thu and M. Ghanbari, "The accuracy of PSNR in predicting video quality for different video scenes and frame rates," *Telecommunication Systems*, vol. 49, no. 1, pp. 35-48, 2012.

Wannida Sae-Tang received her BEng in Electronic and Telecommunication Engineering with first-class honours and her MEng in Electrical Engineering from King Mongkut's University of Technology Thonburi, Thailand in 2007 and 2011, respectively, and her PhD in Information and Communication Systems from Tokyo Metropolitan University, Japan in 2014 funded by the Tokyo Metropolitan Governmental Asian Human Resources Fund scholarship. She is currently an assistant professor at the Sirindhorn International Thai-German Graduate School of Engineering, King Mongkut's University of Technology North Bangkok, Thailand. Her research interests include image processing and multimedia communications. She received the Best Paper Award at the IEICE/ITE/KSBE IWAIT in 2014. From 2007 to 2009, she was an IC Packaging Design engineer in the New Product Design and Research and Development team at United Test and Assembly Center Thai Ltd.

Adisorn Sirikham received his PhD in Manufacturing from Cranfield University, U.K. in 2020 and his MEng in Electrical Engineering from King Mongkut's University of Technology Thonburi, Thailand in 2007. He is currently a lecturer at Department of Electrical and Telecommunication Engineering, Rajamangala University of Technology Krungthep, Thailand. His research interests include artificial intelligence, internet of things, computer vision and image processing, and infrared thermography.