IEICE TRANS. INF. & SYST., VOL.E106–D, NO.11 NOVEMBER 2023

LETTER

A DFT and IWT-DCT Based Image Watermarking Scheme for Industry

Lei LI[†], **Hong-Jun ZHANG**[†], **Hang-Yu FAN**[††], ***Nonmembers***, ***and*** **Zhe-Ming LU**[††a)], ***Member***

SUMMARY Until today, digital image watermarking has not been large-scale used in the industry. The first reason is that the watermarking efficiency is low and the real-time performance cannot be satisfied. The second reason is that the watermarking scheme cannot cope with various attacks. To solve above problems, this paper presents a multi-domain based digital image watermarking scheme, where a fast DFT (Discrete Fourier Transform) based watermarking method is proposed for synchronization correction and an IWT-DCT (Integer Wavelet Transform-Discrete Cosine Transform) based watermarking method is proposed for information embedding. The proposed scheme has high efficiency during embedding and extraction. Compared with five existing schemes, the robustness of our scheme is very strong and our scheme can cope with many common attacks and compound attacks, and thus can be used in wide application scenarios.

key words: *digital image watermarking, robust watermarking, high efficient watermarking*

1. Introduction

Nowadays, massive digital images are being produced, and this brings many new problems. One big problem is copyright infringement since digital images are easy to be copied, edited and falsified. If a copyrighted image is attacked or tampered with, the attacked image is difficult to detect and proof of copyright is difficult to give. Therefore, there is an urgent need to utilize technology to protect the rights of image owners.

As a very important branch of information hiding, image watermarking is an effective technique to solve the copyright protection problem for images [1]–[10]. The watermarking technique can embed the copyright information into the image to be protected and the information can be extracted when needed. Image watermarking algorithms also can be used in many other application scenarios, such as covert communication, tamper detection, and content annotation.

Although image watermarking techniques can be used in many scenarios, they have not been large-scale used yet in the industry. There are two reasons, one is that the watermarking efficiency is often ignored; the other is that the scheme cannot resist various attacks.

Currently, image watermarking algorithms can be mainly divided into two categories: traditional schemes and deep learning based methods. Traditional schemes include spatial domain based [2], moment based [3], transform domain based [4], [5] and hybrid domain based methods [6]. Typical deep learning based schemes are HiDDeN [7] and StegaStamp [8]. Traditional methods belong to the white box method, so their process is easy to control. Deep learning based methods belong to the black box method, where the encoder and decoder are trained by setting the appropriate loss, and their computing process is opaque and difficult to control. To control the process and enhance the efficiency, in this paper, we choose the traditional way.

However, for the existing state of art image watermarking techniques in the traditional way, either robustness or efficiency is not high enough to apply widely in industry. Reference [2] proposed a spatial-embedding watermarking method based on an attended just noticeable difference (JND) model with color complexity, but it is not robust to geometrical attacks without knowing the geometrical parameters and the efficiency is not high because of the JND model. Reference [10] presented a watermarking method in the spatial domain with HVS-imperceptibility for high dynamic range images, it is also not robust to geometrical attacks. Although color image watermarking using new fractional-order exponent moments in [3] is very robust to different kinds of geometric distortions and attacks, the complexity is very high and the efficiency is not high enough for industrial use. In [4], a blind watermarking algorithm based on contourlet transform with singular value decomposition was proposed, but this method is not robust enough to rotation and collage attacks and the complexity is high. In [5], an image watermarking scheme by using advantages of both frequency domain and wavelet domain was proposed, which is very robust to Gaussian noise, but the robustness against geometrics is not high and the complexity is not low. In [6], a robust and reversible color image watermarking algorithm in the spatial domain fusing discrete Fourier transform (DFT) was proposed, but the scheme is not robust to rotation and collage, and the complexity is not low.

From above, we can see that many existing state of the art methods based on different domains cannot achieve both high robustness and fast speed, and thus they are not so suitable to industrial circles. In order to apply digital watermarking techniques to industry, we present a multi-domain

Manuscript received December 31, 2022.
Manuscript revised March 11, 2023.
Manuscript publicized August 22, 2023.

[†]The authors are with the Command and Control Engineering College, Army Engineering University, No.88, Houbiaoying Road, Nanjing, 310027, P. R. China.
[††]The authors are with the School of Aeronautics and Astronautics, Zhejiang University, No.38, Zheda Road, Hangzhou 310027, P. R. China.
a) E-mail: zheminglu@zju.edu.cn (Corresponding author)

DOI: 10.1587/transinf.2022EDL8108

image watermarking scheme, including DFT based watermarking and IWT-DCT based watermarking layers. The innovations and features of the proposed scheme can be summarized as follows:

First, a DFT based watermarking layer is proposed for synchronization with high embedding and detection efficiency and this method also has good imperceptibility. In addition, because of the special template embedding mechanism, we are able to replace the DFT operation with spatial addition to speed up the embedding process. To save the computation time, we also design a new method for determining the rotation angle and scale factor during extraction. This method is therefore faster than many existing DFT-based watermarking schemes.

Second, an IWT-DCT based watermarking layer is proposed for resisting translation attack, JPEG attack and some other common attacks. On the one hand, we use fast algorithms to calculate IWT and DCT. On the other hand, we use the information header to resist the translation attack and shear attack. Thus, our embedding scheme is both efficient and robust.

Third, the watermarking scheme integrates many techniques because it is designed for industrial use with high efficiency, good robustness and good imperceptibility. We combine several techniques such as synchronization correction, information header and error-correcting code together during both embedding and extraction process to make our scheme not only fast enough but also robust to most attacks that maybe occur in real applications. Thus, our method is superior to existing methods since they only use fewer techniques in their scheme and don't think over the real industrial application.

2. Proposed DFT-IWT-DCT Watermarking Scheme

In order to resist various attacks, especially rotation, scaling, collage and some compound attacks, this section proposes the watermarking scheme with two layers. The first watermarking layer is based on DFT for image synchronization. The second watermarking layer is based on IWT-DCT for embedding information.

2.1 DFT Based Watermarking for Synchronization

Synchronization correction is a main technique used in robust image watermarking [9]. DFT is a suitable transform that can be used for synchronization. If the image is rotated or scaled, the 2D DFT spectrum of the image will rotate or scale as well. Therefore, by using these features, we can obtain the scaling rate and the rotation angle of the image with a high accuracy.

The implementation process of this method can be concisely described as follows: First, the image should be a square. If the image is not a square, then it should be padded to be a square. Second, DFT is performed on the image to obtain the 2D spectrum. Third, a pseudo-random $\{0, 1\}$-sequence is generated and the spectral amplitude is changed according to the sequence. The amplitudes to be changed are located on a circle. If the bit in the sequence is 0, the amplitude is unchanged; if the bit in the sequence is 1, the amplitude is enhanced. An example can be shown in Fig. 1. Finally, IDFT is performed on the spectrum to get the watermarked image. By computing the cross-covariance value between the embedded pseudo-random sequence and the circle spectral amplitude sequence of the input image, the rotation angle and scaling factor can be obtained.

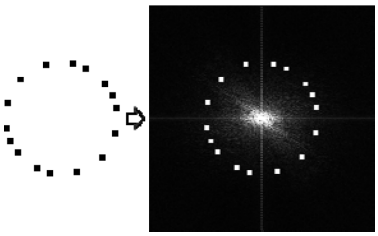

Fig. 1 An embedding example, where a pseudo-random $\{0, 1\}$-sequence on the left, and the embedded spectrum on the right.

The above method can get good results, but the embedding process takes much time, and the embedding strength is not flexible, and the imperceptibility is not good as well. In addition, the extraction process needs exhaustive search, which has low efficiency. In order to deal with these problems, we design a new DFT based method for synchronization.

The proposed method is similar to the above method in principle, but the implementation is different. In our method, before the embedding, a spatial domain template should be constructed. The spatial domain template is a residual result, and the construction process can be summarized as follows:

$$\begin{aligned} spectrum_I &= \mathrm{DFT}(I) \\ spectrum_I^* &= \mathrm{EMBED}(spectrum_I) \\ template &= \alpha(\mathrm{IDFT}(spectrum_I^*) - I) \end{aligned} \tag{1}$$

In Eq. (1), α indicates the weight parameter, I denotes a grayscale image with a large square size. Here, all pixel values in image I are the same. To avoid numerical overflow, we usually choose 128 as the same pixel value. In Eq. (1), the "EMBED" function enhances the corresponding amplitudes of $spectrum_I$ based on the traditional method, but the $\{0, 1\}$-sequence rather than the pseudo-random sequence is embedded. The principles of constructing the $\{0, 1\}$-sequence are as follows: (1) The sequence should not be periodic; (2) The sequence should be asymmetric; (3) The total number of 0s and 1s in the sequence should be balanced. A template construction example is shown in Fig. 2.

The template embedding process can be described as the following equation:

$$I_c^* = I_c + template_{crop} * adaGain \tag{2}$$

where I_c denotes the carrier image, I_c^* denotes the embedded image, $template_{crop}$ denotes the cropped template from the center of the template in Eq. (1). Note that $template_{crop}$ and

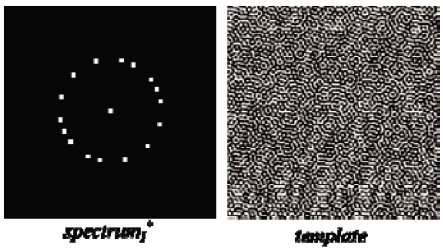

Fig. 2 A template construction example.

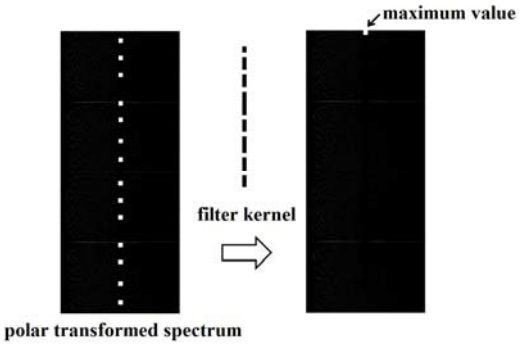

Fig. 3 An example of polar transformed spectrum, constructed filter kernel and result of filtering.

the carrier image have the same size. And *adaGain* denotes the adaptive gain designed for controlling the local weight of the template. Generally, smooth areas are weighted lower and rough areas are weighted higher.

As we know, the rotation angle and scaling factor can be found by computing the cross-covariance value, but this process takes too much time. To save the computation time, we design a new method for determining the rotation angle and scale factor. Before detection, the image to be detected can be cropped for faster detection if the image has a large size. Then DFT is performed on the cropped square image to get spectrum. To reduce possible noises in the DFT spectrum, the spectrum needs to be filtered by the Wiener filter, and then the filtered spectrum needs to be thresholded. If a point value exceeds a certain predefined limit, the point value is divided with the local mean. If the point value does not exceed the limit, the point value is replaced with zero. According to the $\{0, 1\}$-sequence, a filter kernel should be constructed as shown in Fig. 3. And the polar transformation needs to be performed on the clean spectrum. Then, the polar transformed spectrum is filtered by the constructed filter kernel, and the maximum value should be found in the filtered map. If the maximum value is larger than the settled threshold, we can say that this image contains the DFT based watermark; else, there is no DFT based watermark in the image. And the coordinates of the maximum value are the rotation angle and the circle radius. According to the detected radius, the scaling factor can be computed by:

$$scalingFactor = \frac{r_{embedded}}{r_{detected}/\mathrm{int}(l/2)} \tag{3}$$

where $r_{embedded}$ denotes the circle radius in the embedding process, $r_{detected}$ denotes the detected circle radius, and l denotes the width of the square image.

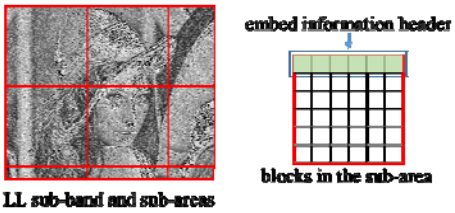

Fig. 4 An example of LL sub-band, sub-areas and blocks in the sub-area. The position for embedding information header in the sub-area is shown as well.

Different from traditional methods, our innovations and contributions are: first, we convert the DFT based transform domain method into a spatial domain based method with better imperceptibility and robustness; second, the template is prepared in advance to greatly reduce the required embedding time; third, the new principles for constructing the $\{0, 1\}$-sequence are given; finally, the required detection time is reduced by cropping the image and using the polar transformation and filter.

The proposed DFT based watermarking method is robust, but it is not easy to embed enough information. Therefore, the IWT-DCT based watermarking method is proposed to embed the customized information.

2.2 IWT-DCT Based Watermarking for Embedding Information

We embed the customized information based on IWT and DCT. Here, IWT is used for down-sampling, although it will reduce the embedding capacity, the robustness and imperceptibility are improved, and the watermarking efficiency is improved as well.

The embedding process is simple. IWT is first performed on the carrier image, and then the LL sub-band is selected for information embedding. The LL sub-band is divided into blocks, and DCT is performed on each block. The information is embedded by modifying DCT coefficients. Before modification, the independent sub-area should be settled in the LL sub-band as shown in Fig. 4. In each independent sub-area, the encoded information will be embedded once. Additionally, the information header should be embedded in each sub-area to resist the translation attack. Then a pair of DCT coefficients of the block is chosen, the difference between two coefficients is computed, and then QIM (Quantization Index Modulation) is used to modify the difference for embedding the information.

The extraction process is similar to the embedding process. IWT is also first performed on the image to get the LL sub-band. The up-left point of the LL sub-band is set as a datum mark to divide the blocks. DCT is performed on each block, and compute the difference between the two selected coefficients, and then QIM is used to extract the watermark bit according to the difference. The information header is searched in the extracted bits, and the position of sub-area can be determined if an information header is found. If there

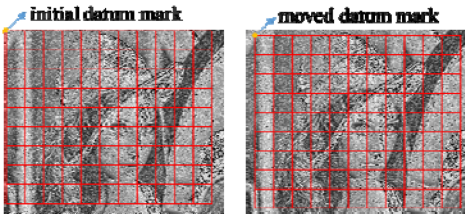

Fig. 5 An example of the datum mark for block dividing.

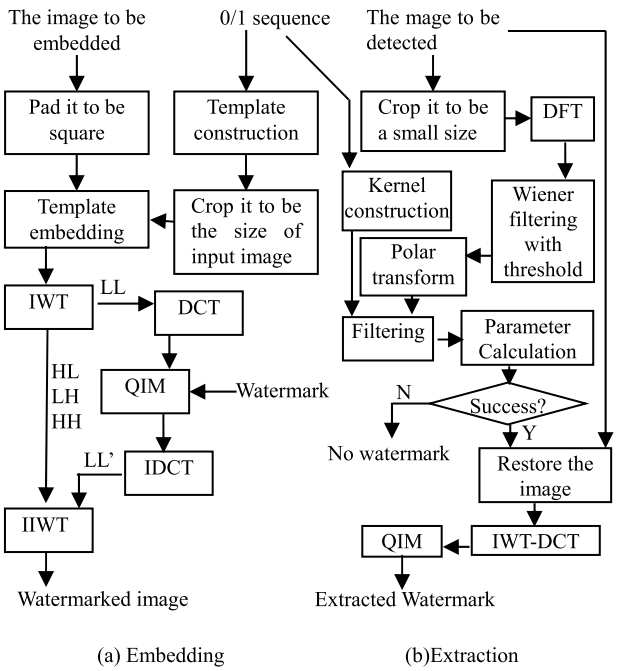

Fig. 6 The block diagram of our scheme.

is no information header in the extracted bits, the position of the datum mark needs to be moved and the blocks need to be divided again, until an information header can be found in the extracted bits. An example is shown in Fig. 5. In this method, we use sub-areas and the search of information headers to resist translation and shear attacks, and the differential quantization embedding method has good robustness and imperceptibility.

2.3 Whole Scheme and Some Tips

The block diagram of the proposed watermarking scheme is shown in Fig. 6.

In the overall embedding process, DFT based watermarking is performed at first, and then IWT-DCT based watermarking is performed. The detailed algorithmic steps are as follows:

Step 0: Construct a spatial domain template based on the equations as given in Eq. (1) in advance as follows:

Step 0.1. Generate a large-size square gray-scale image with each pixel having the same grayscale 128. Perform DFT on this image

Step 0.2. Generate a {0, 1}-sequence and embed it as a circle to obtain the modified spectrum as given in Fig. 2.

Step 0.3. IDFT is performed on the modified spectrum to obtain the final template T to be embedded.

Step 1: Input the image I to be embedded and obtain its image size. If the image is not square, then it should be padded to be square, i.e. I_P.

Step 2: Crop the template to the same size of the padded image I_P, and then the cropped template T_C is added to the padded image to finish the template embedding process based on Eq. (2), obtaining the image I_T.

Step 3: IWT is first performed on I_T, and then the LL sub-band is selected for information embedding. The LL sub-band is divided into blocks, and DCT is performed on each block.

Step 4: The independent sub-area is settled in the LL sub-band and the encoded watermark information is embedded once. Additionally, the information header should be embedded in each sub-area to resist the translation attack. A pair of DCT coefficients of each block is chosen, and then QIM is used to modify the difference between two coefficients to finish embedding the watermark information.

Step 5: Perform the inverse transform to obtain the watermarked image I_W, where the padded part should be removed.

While in the whole extraction process, DFT based extraction is performed first to correct the image, and then IWT-DCT based extraction is performed to extract the information. The detailed algorithmic steps are as follows:

Step 1: Input the image to be detected, it can be cropped for faster detection if the image has a large size.

Step 2: DFT is performed on the cropped square image to get spectrum, which is filtered by the Wiener filter, and then the filtered spectrum is thresholded.

Step 3: Polar transform is performed on the clean spectrum. Based on the {0,1}-sequence, a filter kernel is constructed. Then, the polar transformed spectrum is filtered by the constructed filter kernel.

Step 4: Parameter calculation. If the maximum value in the filtered map is less than the threshold, the image contains no DFT based watermark, and the extraction is terminated. Otherwise, the coordinates of the maximum value are the rotation angle and the circle radius. According to the detected radius, the scaling factor can be computed by Eq. (3). Then, the input image is restored based on the calculated parameters.

Step 5: IWT is first performed on the restored image to get the LL sub-band. The up-left point of the LL sub-band is set as a datum mark to divide the blocks.

Step 6: Perform DCT on each block, and compute the difference between the two selected coefficients, and then QIM is used to extract the watermark bit according to the difference.

Step 7: The information header is searched in the extracted bits as illustrated in Sect. 2.2. Thus, the final extracted watermark bits can be obtained by removing the header bits.

In the practical applications, some techniques or tips can be used to improve the performance: first, to compile the

program faster, with the same algorithm, C++ is more efficient than Python; second, float numbers can be replaced by integer numbers for higher efficiency; third, the information to be embedded needs to be encoded before the embedding; finally, encryption or scrambling can enhance the security, and error-correcting code can be used for error correction.

3. Experimental Results

In this section, the performance of the proposed scheme is demonstrated by experiments. And we also compare our scheme with five state-of-the-art watermarking schemes, including fractional-order exponent moments based method [3], contourlet transform with SVD based method [4], DWT-DCT composition based method [5], DFT in the spatial domain based method [6], and synchronization correction based method [9].

Here, the performances on effectiveness, robustness, and imperceptibility will be illustrated. We embed 8×8 binary image (i.e., 64 bits) redundantly in each image. The number of test images is 20,000, and the images have multiple formats (jpg, bmp, png), and the image size ranges from 700×700 to 2000×2000. First, we illustrate the effectiveness of the proposed scheme. The results are displayed in histograms as shown in Fig. 7. From Fig. 7, we can find that our scheme is very fast, and the real-time requirement can be satisfied.

Second, the imperceptibility of the proposed scheme is demonstrated, and the PSNR (Peak Signal-to-Noise Ratio) is used as an indicator for evaluating the imperceptibility. The results are shown in Fig. 8. From the results, we can find that our scheme has a good imperceptibility (mostly with 38 dB–44 dB).

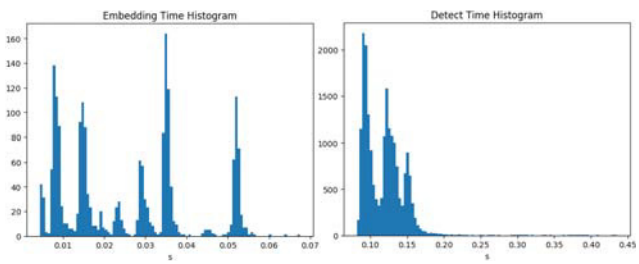

Fig. 7 Embedding and detection time histograms over 20000 images.

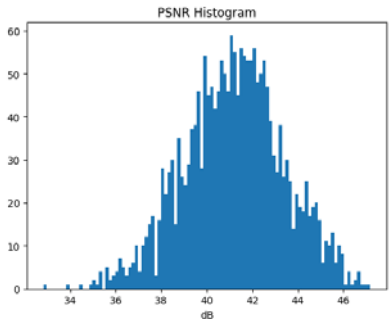

Fig. 8 Watermarked PSNR histogram over 20000 images.

Finally, we compare the robustness of proposed scheme with state-of-the-art methods. First, we compare the invisibility of watermark after embedding. We show the average PSNR and SSIM (Structural Similarity Index) values over 20000 watermarked test images for different algorithms in Table 1. We can see that our algorithm has good invisibility. Second, we compare their robustness. Our scheme can cope with multiple attacks and many joint attacks. The types of attack for test are rotation, scaling, translation, JPEG compression, cropping, collage, Gamma correction, white noise, contrast adjustment, luminance adjustment and median filtering. We use two metrics here. One is average NC (Normalized Cross Correlation) value over 20000 test images under different attacks. The other is accuracy, which is defined as the ratio of the number of images from which we can extract all watermark bits without any error to the total number of images, i.e., 20000. The comparison results are shown in Table 2 and Table 3 respectively, and we can find that the proposed scheme is more robust than other schemes.

Table 1 Comparison of average PSNR and SSIM values over 20000 watermarked test images among different algorithms.

Metrics	our	[3]	[4]	[5]	[6]	[9]
Average PSNR	41.72	44.24	41.65	41.68	41.80	38.42
Average SSIM	0.990	0.993	0.987	0.988	0.990	0.979

Table 2 Comparison of extraction accuracy over 20000 test images.

Attacks	our	[3]	[4]	[5]	[6]	[9]
No	**1.000**	1.000	1.000	1.000	1.000	1.000
JPEG-50	**0.977**	0.973	0.969	0.959	0.971	0.915
Rotate-5^{O}	**0.985**	0.982	0.325	0.175	0.245	0.980
Crop-50%	**0.997**	0.991	0.735	0.850	0.875	0.923
Gamma-0.5	0.893	0.843	**0.934**	0.854	0.812	0.853
Gamma-1.5	**0.961**	0.875	0.943	0.952	0.886	0.885
Scaling-0.6	**0.969**	0.948	0.946	0.792	0.946	0.966
Scaling-1.5	**0.955**	0.952	0.939	0.940	0.950	0.948
Translation	**1.0**	0.990	0.125	0.125	0.234	1.0
Collage	**1.0**	0.924	0.245	0.245	0.158	0.986
White-noise(20dB)	**0.998**	0.994	0.923	0.970	0.958	0.989
Contrast-50%	0.542	0.498	0.515	0.532	0.486	**0.581**
Luminance-50%	**0.770**	0.675	0.698	0.724	0.751	0.743
Median-3×3	**0.970**	0.964	0.910	0.945	0.925	0.956

Table 3 Comparison of average NC values over 20000 test images.

Attacks	our	[3]	[4]	[5]	[6]	[9]
No	**1.000**	1.000	1.000	1.000	1.000	1.000
JPEG-50	**0.981**	0.976	0.972	0.965	0.979	0.931
Rotate-5^{O}	**0.990**	0.988	0.513	0.552	0.610	0.989
Crop-50%	**0.998**	0.997	0.812	0.871	0.883	0.947
Gamma-0.5	0.913	0.879	**0.951**	0.878	0.869	0.871
Gamma-1.5	**0.972**	0.883	0.961	0.968	0.901	0.907
Scaling-0.6	**0.975**	0.962	0.968	0.853	0.965	0.973
Scaling-1.5	**0.971**	0.969	0.954	0.958	0.962	0.960
Translation	**1.000**	0.996	0.523	0.578	0.598	1.000
Collage	**1.000**	0.950	0.590	0.592	0.512	0.991
White-noise(20dB)	**0.999**	0.996	0.950	0.982	0.976	0.994
Contrast-50%	0.732	0.682	0.675	0.686	0.681	**0.772**
Luminance-50%	**0.850**	0.783	0.799	0.813	0.836	0.843
Median-3×3	**0.985**	0.980	0.932	0.957	0.960	0.978

4. Conclusions and Future Works

In this Letter, we propose a multi-domain digital image watermarking method. To make the watermarking scheme suitable for industry, a fast DFT based watermarking layer and a robust IWT-DCT based watermarking layer are designed. Experimental results show that the real-time performance can be satisfied and the proposed scheme can resist a variety of attacks. However, the robustness for print-camera is not good, because we consider the imperceptibility emphatically. Therefore, we are going to study the A/D-D/A conversion attack in the future research.

Acknowledgements

This research is supported in part by the National Key Research and Development Program of China under Grant No. 2020AAA0140004 and the Public Good Research Project of Science and Technology Program of Zhejiang Province under Grant No. LGG21F020005.

References

[1] W. Wan, J. Wang, Y. Zhang, J. Li, H. Yu, and J. Sun, "A comprehensive survey on robust image watermarking," Neurocomputing, vol.488, pp.226–247, March 2022.

[2] W. Wan, K. Zhou, K. Zhang, Y. Zhan, and J. Li, "JND-guided perceptually color image watermarking in spatial domain," IEEE Access, vol.8, pp.164504–164520, Sept. 2020.

[3] K.M. Hosny, M.M. Darwish, and M.M. Fouda, "Robust color images watermarking using new fractional-order exponent moments," IEEE Access, vol.9, pp.47425–47435, 2021.

[4] L. Song, X.-C. Sun, and Z.-M. Lu, "Robust blind watermarking algorithm based on contourlet transform with singular value decomposition," IEICE Trans. Fundamentals, vol.E104-A, no.3, pp.640–643, March 2021.

[5] R.Y. Abadi and P. Moallem, "Robust and optimum color image watermarking method based on a combination of DWT and DCT," Optik, vol.261, 169146, July 2022.

[6] H. Cao, F. Hu, Y. Sun, S. Chen, and Q. Su, "Robust and reversible color image watermarking based on DFT in the spatial domain," Optik, vol.262, 169319, July 2022.

[7] J. Zhu, R. Kaplan, J. Johnson, and F.-F. Li, "HiDDeN: Hiding data with deep networks," Proc. European Conference on Computer Vision, pp.682–697, Sept. 2018.

[8] W.T. Sun and Z.M. Lu, "An improved StegaStamp watermarking scheme by screen-cam robust detector," Journal of Network Intelligence, vol.7, no.1, pp.189–197, 2022.

[9] X.-Y. Wang, H. Xu, S.-Y. Zhang, L.-L. Liang, P.-P. Niu, and H.-Y. Yang, "A color image watermarking approach based on synchronization correction," Fundamenta Informaticae, vol.158, no.4, pp.385–407, 2018.

[10] K.R. Perez-Daniel, F. Garcia-Ugalde, and V. Sanchez, "Watermarking of HDR images in the spatial domain with HVS-imperceptibility," IEEE Access, vol.8, pp.156801–156817, 2020.