Research on the Relevant Methods and Technologies of Digital Watermarking

Ding Yi

College of Design and Engineering National University of Singapore, 4 Engineering Drive 3, Singapore 117583, Singapore

e0973777@u.nus.edu

Abstract. With the diversification of recording equipment and the rapid development of Internet technology, it is easier to steal and spread movies, which makes the copyright protection of movies face great challenges. The technology of tracing the origin of movie piracy in cinemas provides strong technical support for film digital copyright protection by extracting the watermark information hidden in the film carrier, tracking and locating the playback server where the piracy flows out and the theater where the piracy occurs. In addition, the emergence of digital technology has led to the rapid spread and growth of electronic resources, such as photos, music and movies. It is convenient to access online information, but it also poses a challenge to the technical protection of intellectual property. Image watermarking, including watermark embedding and extraction, is a typical technology of copyright protection and ownership identification. In addition, because many software (such as modern browsers and video players) can easily display all these contents, adding image watermarks on the screen and avoiding information leakage in screenshots or screen recording has become an important application scenario of image watermarking. In this paper, we propose and discuss digital watermarking pipeline with some.

Keywords: Digital Watermarking, Image, Video.

1. Introduction

The advancement of digitization technologies, together with the presence and accessibility of the Internet, have facilitated the generation, reproduction, and dissemination of digital content, which has required the development of copy- right protection technology. Many researches and techniques have contributed to data security, such as data encryption and password authentication. How- ever, given that data can be recorded and replicated, there still exist risks in providing the original document for those who have passed the authorization verification. Therefore, there exists a need for not only controlling access but tracking the replication and circulation of electronic records [1, 2].

Digital image watermarking, namely blind watermarking, is a technique in information security that can integrate initial data with critical information, such as data integrity, ownership, and source. Unlike an apparent stamp or sig- nature, digital watermarking utilizes randomness and redundancy in the content of files and ensures each file copy is uniquely and indissolubly marked with authentication and authorization information [3]. The work is also called steganography. Besides, the watermarking is challenging to detect and remove without specific watermarking extraction technology corresponding to the embedding process. Therefore, the discrepancy between digital image watermarking and other data security technology can boil down to two points: Watermarking is invisible and inseparable from data.

Since users will display the digital content on the screen, digital watermarking technology also faces everyday situations like screen capture and recording. Users may use screen capture software or take pictures outside the screen to make copies of the original data. Thus, the digital content may change in scale, intensity and resolution. Under the circumstances, digital watermarking has to maintain intact in an incomplete, blurred, and distorted image. Therefore, in this paper, we comb traditional file-based digital watermarking and consider real-time digital watermarking techniques that prevent screen capture.

The paper is organized as follows: Section 2 shows watermarking embedding and extraction pipeline. Section 3 presents the characteristics often used to evaluate and classify blind watermarking

techniques. In section 4, we introduce traditional and novel methods for digital watermarking. Section 5 summarizes the development of watermarking and its future directions [4, 5].

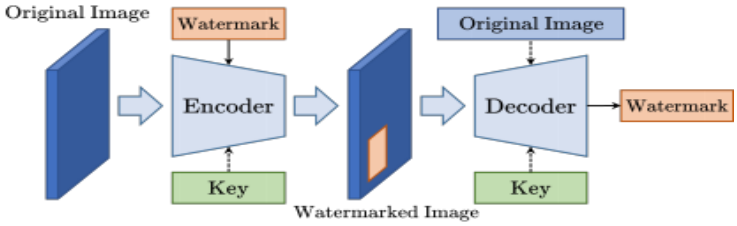

Figure 1. Digital Watermarking Embedding and Extraction Pipeline

2. Correlation theory

The carrier image of the digital watermark is preprocessed. On this basis, the digital watermark algorithm is used to verify whether the proposed carrier image preprocessing method can improve the relationship between the transparency and robustness of the watermark. Discrete cosine transform and singular value decomposition are performed on the preprocessed carrier image, and the binary watermark image to be hidden is embedded into the singular value matrix of the classified carrier image respectively, the specific embedding and extraction process is shown in Figure 1. The carrier image preprocessing has been completed before the watermark embedding. The overall process of watermark embedding is shown in Figure 2. The carrier image preprocessing is completed After that, convert the image corresponding to each type of super-pixel block to RGB color space. The conversion process is the reverse of RGB to CIELab [6]. Calculate the color moments of three types of images in RGB color space, including the color moments of each type of image R, G, B color channels, and then select R, G The color channel with the largest color moment in B is used as the watermark embedding channel of this kind of image [7].

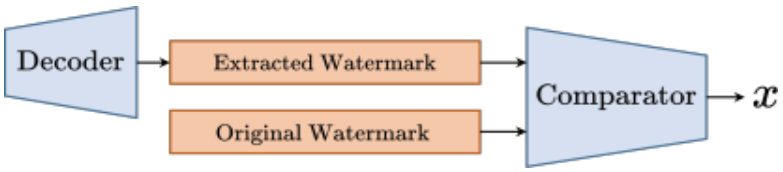

Figure 2. Decoder and Comparator

The embedded watermark image is extracted by the reverse process of the watermark embedding process.

●Step 1: DCT transform the image to be extracted.

●Step 2: Extract embedded information

●Step 3: arrange the extracted binary digits in order to obtain one-dimensional binary. The final sequence is obtained by performing Arnold inverse scrambling on the sequence, which is the extraction result.

3. Basic concepts

Digital watermarking techniques can embed and extract a specific message. Apart from the comparator, there exist many concepts to measure and classify the watermarking algorithms. In this section, we discuss some critical concepts for measuring and classifying different digital watermarking methods [8].

3.1. Steganography and Cryptography

In the cryptography, the main parts include enciphering and deciphering, namely encryption and decryption. The original message, often called plain data, is dis- guised in a distinct form and exposed to all possible receivers. Still, only specific recipients who have the secret key can reconstruct the

initial data. Cryptography can prevent the digital contents from leaking in the transmission from the sender to the receiver but no longer protect the data after the data is deciphered. Unlike cryptography, steganography work focuses more on hiding information than encrypting it. Steganography embeds the information in plain data, and without particular decoding methods, users will not likely notice the blink watermark. The term imperceptibility measures whether the digital watermark is visible or not. In comparison, the concept of steganography is broader than cryptography, because the embedded message can be encrypted in a steganography system. Therefore, the digital watermarking system is independent of and compatible with any encryption algorithm. Steganography and cryptography are also measured differently. The steganography area is often related to the Human Visual System (HVS) research, which is used to model the human's visual perception system and figure out the sensitivity to illumination and color changes of human eyes. The analysis of HVS causes the concept of Just Noticeable Distortion (JND), similar to the notion of password security in cryptography. JND denotes the maximum distortion that the HVS will not perceive [9].

3.2. Blind and Non-blind Techniques

As is mentioned in the pipeline, there exist two types of decoders in digital image watermarking, which are classified by whether the original image has to be used as the input for the decoder. Blind watermark algorithms can remove the embedded information without the original data.

3.3. Robust and Fragile Watermark

Robustness measures the ability of a watermark to resist distortion or attack. Generally, digital watermarking is classified as fragile, semi-fragile and robust techniques. Fragile watermarks are very prone to distortion when the data changes. Any changes appearing in data dissemination and tampering will result in the collapse of the watermarks. Therefore, the fragile watermarks can be used to confirm the data integrity and authenticity. By contrast, robust watermarks can withstand attacks and keep identified in distortions and signal processing, including scaling, compression, and cropping. If the user wants to remove the robust watermarking from the digital content, the quality of protected data will be greatly degenerated. Robust watermarks are often used to record critical information like data ownership and track the distribution of the data [10].

4. Related algorithms

The use of sections to divide the text of the paper is optional and left as a decision for the author. Where the author wishes to divide the paper into sections the formatting. In this section, some main digital watermarking algorithms are introduced. Similar to some other areas in computer vision, digital watermarking can be classified as traditional algorithms and deep-learning-based methods.

4.1. Spatial Domain Watermarking

The watermarking algorithm in the spatial domain alters the pixel values to embed specific information into the image. Mathematically, some spatial-domain watermarking algorithms can be described as simple addition. where I is the initial data. W is the watermark message, and I\ is the watermarked data with embedded information. The Least Significant Bit (LSB) method is one of the most famous watermarking algorithms [11, 12]. In this algorithm, the watermark information is embedded in the least significant bit of the data. Since the changes in the LSB of pixels make nearly no difference to the image, the LSB watermark has good concealment. The watermark information can be extracted just by taking the data modulo 2. Therefore, the method is computationally inexpensive. However, the LSB method is fragile, and the LSB watermark is easy to lose in the transmission and replication of the data, such as screenshot or recapturing. For spatial-domain watermarking algorithm, there always exists a trade- off between the invisibility and the robustness. Since we hope watermarked image I is nearly identical to the original image I, the embedded information W must be numerically small. Therefore, the message is easily corrupted by random

noise. The Patchwork method, namely block-based method, is used to enhance the robustness, which selects n pairs of image regions (x, y). The pixels in region x are changed to larger values and the region y is darkened. The changing brightness level is controlled so that the incremental intensity of illumination keeps invisible. However, the method compromises on the capacity, which means the amount of embedded information decreases [13].

Another spatial-domain based watermarking method is called texture-based watermark, whose idea is that the watermark will be difficult to detect because it is similar to the original figure in texture [14]. As shown in Figure 3. Amplitude Modulation algorithm uses a single channel of an image with RGB channels to embed watermark and remain imperceptible to the HVS model [15].

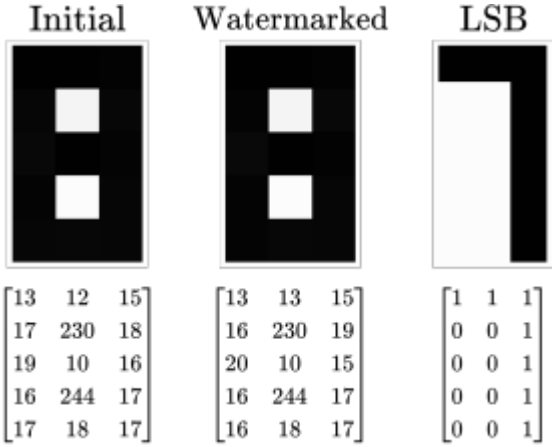

Figure 3. A example of LSB watermark. The original image data is the number 8. The embedded watermark is the number 7

4.2. Transform Domain Watermarking

Transform-domain-based watermarking embeds the information into the frequency domain, mainstream algorithms including Discrete Cosine Transforms (DCT) [16], Discrete Wavelets Transform DWT, Discrete Fourier Transforms (DFT), and Singular Value Decomposition (SVD). As mentioned above, spatial domain algorithms are not resilient against multiple assaults, despite their simplicity to be implemented. Frequency domain approaches are more resistant to several forms of assaults than most spatial domain methods, but they need more computer resources.

DCT watermarking algorithms apply the 1D-DCT in the vertical and horizontal directions and transform the input image into the frequency domain. After the DCT transform, the algorithm embeds the message into DCT coefficients for all the blocks and then performs the inverse transform of DCT. DWT decom- poses the input image into four sub-bands in the frequency domain, namely LL, LH, HL and HH. The DWT decomposition can be iterated on the sub-band LL. Thus the image can be decomposed into k-levels, where high-level sub-bands contain more detailed information about the original image. Compared to the DCT, DWT with multi-resolution information can provide a more detailed de- scription about the HVS. In DFT algorithms, the image can be transformed as a series of complex-valued coefficients, which includes the information of amplitude and phase. DFT watermarks are usually robust because its features can maintain invariant in geometrical changes. Different methods of transform-domain-based watermarking can function simultaneously, which is called hybrid domain. Using multi-level DWT, SVD, and wavelet fusion, the researchers designed an effective watermarking model for colored images. They divided the picture into red, green, and blue (RGB) channels. Each channel was fused with a grayscale picture, and the three layers were combined into a grayscale image. The watermarking in transform domain can be integrated with encryption methods. Another study supposed that watermarking alone is not sufficient for security. When watermarking is used to conceal sensitive information under a cover picture for security purposes, the sensitive information may be attacked, corrupted, or retrieved by third parties. They provided a two-layer security framework for protecting colored watermark pictures. Using multi-level DWT,

SVD, and wavelet fusion, the first level included embedding the color image inside a grayscale picture. The second level was intended to introduce another encryption procedure [16, 17].

4.3. Deep-Learning-based Watermarking

In recent years, deep learning has achieved great success in many branches of computer vision, attracting the attention of many researchers. In digital watermarking technology, the decoder needs to decode the watermarked image into a particular string of specific messages. This process is similar to classification tasks in computer vision, where the watermark can be considered the label. Based on this idea, substantial research has been conducted to associate the watermarking approach with the learning framework. At a glance, the deep-learning-based watermarking technology can be separated into two groups: single-staged network and double-stage network. Double-staged network simultaneously accomplishes the watermark embedding and detection using a single neural network, whereas a single-staged network only focuses on a single task, watermark embedding or detection. Therefore, the single-staged network can be further divided into detection or embedding networks [18].

For the detection network, given that any watermark embedding process often results in imperceptible artifacts in the original image. Therefore, some researchers use deep learning networks to learn the watermark patterns. For example, some methods use a sliding window to traverse the pixels of the image and understand the discrepancy between the central pixel and the pixels around. With the development of a convolutional neural network (CNN), extensive work has been proposed to apply CNN to watermark detection. A CNN-based neural network effectively detects the invisible watermark embedded by the DWT and SVD methods. Furthermore, several studies, motivated by the success of deep learning in image classification, examine the capacity of classification deep neural networks for zero-bit watermarking-based blind watermarking. The watermarking is accomplished in the feature domain, and the resulting watermarked picture is updated repeatedly using the adversarial samples approach.

Some research is executed to encode watermarks with the assistance of a neural network. Since the single- staged encoding networks seldom consider the detection process, they are often non-blind watermarking techniques and compare the watermarked image with the original image to extract the watermark. Some studies encode watermark information using multiple pictures created by the same neural network, assuming that the two watermarked images should be visually indistinguishable but have different embedded information. The watermarked photo is rebuilt based on the outputs of two networks and the original carrier image by comparing the Euclidean distance between the watermark image blocks. Another work pre-trains a deep-learning model with a new watermarking pattern that may be considered the input's designated type. The original coefficients of image information in the transform domain are then replaced with the model's output based on the watermark information. Similarly, the message is reconstructed by comparing the initial figure and the model's production watermarked image [19].

The neural network first computes the image regions' watermarking pro- cess using the stochastic gradient descent method. Then the modified parts of the image are input back into the model many times to determine the optimal domain for watermark representation. The watermarked picture will then be attacked or distorted and sent back into the network for classification to retrieve the watermark message. During this assault and model-updating process, the resilience and precision of the model will be enhanced [20, 21]. Finally, the training process can improve the model's performance by repeating the embedding, attacks, and extraction stages using the revised model. Furthermore, some researches use the auto-encoder-decoder deep-learning framework to replace full-connected network. The application of auto-encoder and decoder has proven effective and is able to withstand a series of attacks [22, 23].

5. Conclusion

In this paper, we briefly review digital watermarking technology. First, we generally introduced the work and application scenarios of digital watermarking. Then, we organized a pipeline framework of digital watermarking encoding and decoding technology. Based on the sorted-out framework, we introduce the essential concepts and commonly used indicators in digital watermarking technology. When sorting out the detailed methods of digital watermarking, we divide the algorithms into traditional and deep learning methods and introduce them, respectively. The rapid development of deep learning allows for new digital watermarking technology. The development of the Internet and digital technology have promoted the production and dissemination of digital content and brought about the problem of intellectual property protection, which provides an opportunity to apply and develop digital watermarking technology. It is hoped to provide a basis for future investigations and research on digital watermarking technology.

References

[1] Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication. Journal of the American society for information science, 51(14), 1306-1320.

[2] Domingos, R. F., Baalousha, M. A., Ju-Nam, Y., Reid, M. M., Tufenkji, N., Lead, J. R., ... & Wilkinson, K. J. (2009). Characterizing manufactured nanoparticles in the environment: multimethod determination of particle sizes. Environmental science & technology, 43(19), 7277-7284.

[3] Ló ez Jiménez, D., Dittmar, E. C., & Vargas Portillo, J. P. (2022). The trusted third party or digital notary in Spain: effect on virtual transactions. International Review of Law, Computers & Technology, 36(3), 453-469.

[4] Wan, W., Wang, J., Zhang, Y., Li, J., Yu, H., & Sun, J. (2022). A comprehensive survey on robust image watermarking. Neurocomputing.

[5] Begum, M., & Uddin, M. S. (2022). Towards the development of an effective image watermarking system. Security and Privacy, 5(2), e196.

[6] Gong, L., Yu, C., Lin, K., & Liu, C. (2022). A Lightweight Powdery Mildew Disease Evaluation Model for Its In-Field Detection with Portable Instrumentation. Agronomy, 12(1), 97.

[7] Feng, L., Li, H., Gao, Y., & Zhang, Y. (2020). The Application of Sparse Reconstruction Algorithm for Improving Background Dictionary in Visual Saliency Detection. Intelligent Automation & Soft Computing, 26(4).

[8] Wu, M., Jin, X., Jiang, Q., Lee, S. J., Liang, W., Lin, G., & Yao, S. (2021). Remote sensing image colorization using symmetrical multi-scale DCGAN in YUV color space. The Visual Computer, 37, 1707-1729.

[9] Jiang, Q., Tawose, O. T., Pei, S., Chen, X., Jiang, L., Wang, J., & Zhao, D. (2019). Weakly-supervised image semantic segmentation based on superpixel region merging. Big Data and Cognitive Computing, 3(2), 31.

[10] Kumar, A. (2022). A cloud-based buyer-seller watermarking protocol (CB-BSWP) using semi-trusted third party for copy deterrence and privacy preserving. Multimedia Tools and Applications, 81(15), 21417-21448.

[11] Abdulrahman, A. K., & Ozturk, S. (2019). A novel hybrid DCT and DWT based robust watermarking algorithm for color images. Multimedia Tools and Applications, 78, 17027-17049.

[12] Fang, H., Chen, D., Huang, Q., Zhang, J., Ma, Z., Zhang, W., & Yu, N. (2020). Deep template-based watermarking. IEEE Transactions on Circuits and Systems for Video Technology, 31(4), 1436-1451.

[13] Yuan, Z., Liu, D., Zhang, X., & Su, Q. (2020). New image blind watermarking method based on two-dimensional discrete cosine transform. Optik, 204, 164152.

[14] Ayubi, P., Jafari Barani, M., Yousefi Valandar, M., Yosefnezhad Irani, B., & Sedagheh Maskan Sadigh, R. (2021). A new chaotic complex map for robust video watermarking. Artificial Intelligence Review, 54, 1237-1280.

[15] Kuraparthi, S., Kollati, M., & Kora, P. (2019). Robust Optimized Discrete Wavelet Transform-Singular Value Decomposition Based Video Watermarking. Traitement du Signal, 36(6).

[16] Salah, E., Amine, K., Redouane, K., & Fares, K. (2021). A Fourier transform based audio watermarking algorithm. Applied Acoustics, 172, 107652.

[17] Zhang, Y., & Sun, Y. (2019). An image watermarking method based on visual saliency and contourlet transform. Optik, 186, 379-389.

[18] Kang, X., Zhao, F., Chen, Y., Lin, G., & Jing, C. (2020). Combining polar harmonic transforms and 2D compound chaotic map for distinguishable and robust color image zero-watermarking algorithm. Journal of Visual Communication and Image Representation, 70, 102804.

[19] Guan, X., Feng, H., Zhang, W., Zhou, H., Zhang, J., & Yu, N. (2020, October). Reversible watermarking in deep convolutional neural networks for integrity authentication. In Proceedings of the 28th ACM International Conference on Multimedia (pp. 2273-2280).

[20] Elshazly, A. R., Nasr, M. E., Fouad, M. M., & Abdel-Samie, F. E. (2021, December). Intelligent high payload audio watermarking algorithm using colour image in dwt-svd domain. In Journal of Physics: Conference Series (Vol. 2128, No. 1, p. 012019). IOP Publishing.

[21] Nazari, M., & Maneshi, A. (2021). Chaotic reversible watermarking method based on iwt with tamper detection for transferring electronic health record. Security and Communication Networks, 2021, 1-15.

[22] Talati, S., & Etezadifar, P. (2020). Providing an Optimal Way to Increase the Security of Data Transfer Using Watermarking in Digital Audio Signals. Majlesi Journal of Telecommunication Devices, 9(1), 35-46.

[23] Wei, Q., Wang, H., & Zhang, G. (2020). A robust image watermarking approach using cycle variational autoencoder. Security and Communication Networks, 2020, 1-9.