Nguyen-Son Vo
Van-Phuc Hoang (Eds.)

334

Industrial Networks and Intelligent Systems

LNICST

6th EAI International Conference, INISCOM 2020
Hanoi, Vietnam, August 27–28, 2020
Proceedings

Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 334

Editorial Board Members

Ozgur Akan
Middle East Technical University, Ankara, Turkey
Paolo Bellavista
University of Bologna, Bologna, Italy
Jiannong Cao
Hong Kong Polytechnic University, Hong Kong, China
Geoffrey Coulson
Lancaster University, Lancaster, UK
Falko Dressler
University of Erlangen, Erlangen, Germany
Domenico Ferrari
Università Cattolica Piacenza, Piacenza, Italy
Mario Gerla
UCLA, Los Angeles, USA
Hisashi Kobayashi
Princeton University, Princeton, USA
Sergio Palazzo
University of Catania, Catania, Italy
Sartaj Sahni
University of Florida, Gainesville, USA
Xuemin (Sherman) Shen
University of Waterloo, Waterloo, Canada
Mircea Stan
University of Virginia, Charlottesville, USA
Xiaohua Jia
City University of Hong Kong, Kowloon, Hong Kong
Albert Y. Zomaya
University of Sydney, Sydney, Australia

Nguyen-Son Vo · Van-Phuc Hoang (Eds.)

Industrial Networks and Intelligent Systems

6th EAI International Conference, INISCOM 2020
Hanoi, Vietnam, August 27–28, 2020
Proceedings

Springer

Contents

Telecommunications Systems and Networks

Intelligent Channel Utilization Discovery in Drone to Drone Networks for Smart Cities 3
Muhammed Raşit Erol and Berk Canberk

Downlink Resource Sharing and Multi-tier Caching Selection Maximized Multicast Video Delivery Capacity in 5G Ultra-Dense Networks 19
Thanh-Minh Phan, Nguyen-Son Vo, Minh-Phung Bui, Quang-Nhat Tran, Hien M. Nguyen, and Antonino Masaracchia

Performance Analysis of Relay Selection on Cooperative Uplink NOMA Network with Wireless Power Transfer 32
Van-Long Nguyen, Van-Truong Truong, Dac-Binh Ha, Tan-Loc Vo, and Yoonill Lee

Convolutional Neural Network-Based DOA Estimation Using Non-uniform Linear Array for Multipath Channels 45
Van-Sang Doan, Thien Huynh-The, Van-Phuc Hoang, and Dong-Seong Kim

An UAV and Distributed STBC for Wireless Relay Networks in Search and Rescue Operations 57
Cong-Hoang Diem and Takeo Fujii

Hardware, Software, and Application Designs

Resolution-Improvement of Confocal Fluorescence Microscopy via Two Different Point Spread Functions 77
Xuanhoi Hoang, Vannhu Le, and MinhNghia Pham

Estimations of Matching Layers Effects on Lens Antenna Characteristics 85
Phan Van Hung, Nguyen Quoc Dinh, Hoang Dinh Thuyen, Nguyen Tuan Hung, Le Minh Thuy, Le Trong Trung, and Yoshihide Yamada

A 3-Stacked GaN HEMT Power Amplifier with Independently Biased Technique 95
Luong Duy Manh, Tran Thi Thu Huong, Bui Quoc Doanh, and Vo Quang Son

Feasibility and Design Trade-Offs of Neural Network Accelerators Implemented on Reconfigurable Hardware 105
Quang-Kien Trinh, Quang-Manh Duong, Thi-Nga Dao, Van-Thanh Nguyen, and Hong-Phong Nguyen

Information Processing and Data Analysis

Adaptive Essential Matrix Based Stereo Visual Odometry with Joint Forward-Backward Translation Estimation 127
Huu Hung Nguyen, Quang Thi Nguyen, Cong Manh Tran, and Dong-Seong Kim

A Modified Localization Technique for Pinpointing a Gunshot Event Using Acoustic Signals 138
Thin Cong Tran, My Ngoc Bui, and Hoang Huy Nguyen

Table Structure Recognition in Scanned Images Using a Clustering Method 150
Nam Van Nguyen, Hanh Vu, Arthur Zucker, Younes Belkada, Hai Van Do, Doanh Ngoc- Nguyen, Thanh Tuan Nguyen Le, and Dong Van Hoang

Distributed Watermarking for Cross-Domain of Semantic Large Image Database 163
Le Danh Tai, Nguyen Kim Thang, and Ta Minh Thanh

Depth Image Reconstruction Using Low Rank and Total Variation Representations 181
Van Ha Tang and Mau Uyen Nguyen

Deep Learning Based Hyperspectral Images Analysis for Shrimp Contaminated Detection 195
Minh-Hieu Nguyen, Xuan-Huyen Nguyen-Thi, Cong-Nguyen Pham, Ngoc C. Lê, and Huy-Dung Han

A Predictive System for IoTs Reconfiguration Based on TensorFlow Framework 206
Tuan Nguyen-Anh and Quan Le-Trung

Industrial Networks and Intelligent Systems

An Optimal Eigenvalue-Based Decomposition Approach for Estimating Forest Parameters Over Forest Mountain Areas 221
Nguyen Ngoc Tan and Minh Nghia Pham

An Improved Forest Height Inversion Method Using Dual-Polarization PolInSAR Data 233
HuuCuong Thieu and MinhNghia Pham

An Attempt to Perform TCP ACK Storm Based DoS Attack on Virtual and Docker Network 243
Khanh Tran Nam, Thanh Nguyen Kim, and Ta Minh Thanh

Identification of Chicken Diseases Using VGGNet and ResNet Models 259
Luyl-Da Quach, Nghi Pham-Quoc, Duc Chung Tran, and Mohd. Fadzil Hassan

Design and Evaluation of the Grid-Connected Solar Power System at the Stage of DC BUS with Optimization of Modulation Frequency for Performance Improvement 270
Nguyen Duc Minh, Quach Duc-Cuong, Nguyen Quang Ninh, Y Nhu Do, and Trinh Trong Chuong

Security and Privacy

An Efficient Side Channel Attack Technique with Improved Correlation Power Analysis 291
Ngoc-Tuan Do and Van-Phuc Hoang

An Optimal Packet Assignment Algorithm for Multi-level Network Intrusion Detection Systems 301
Dao Thi-Nga, Chi Hieu Ta, Van Son Vu, and Duc Van Le

Privacy-Preserving for Web Hosting 314
Tam T. Huynh, Thuc D. Nguyen, Nhung T.H. Nguyen, and Hanh Tan

A Novel Secure Protocol for Mobile Edge Computing Network Applied Downlink NOMA 324
Dac-Binh Ha, Van-Truong Truong, and Duy-Hung Ha

Author Index 337

Distributed Watermarking for Cross-Domain of Semantic Large Image Database

Le Danh Tai, Nguyen Kim Thang, and Ta Minh Thanh(✉)

Le Quy Don Technical University, 239 Hoang Quoc Viet, Cau Giay, Ha Noi, Vietnam
ledanhtai@gmail.com, thangnk1990@gmail.com, thanhtm@mta.edu.vn

Abstract. This paper proposes a new method of distributed watermarking for large image database that is used for deep learning. We detect the semantic meaning of set of images from the database and embed the a part of watermark into such images set. A part of watermark is one shadow generated from the original watermark by using (n, n) secret sharing scheme. Each shadow is embedded into DCT-SVD domain of one image from the dataset. Since the image sets have multiple image and are distributed in the whole of multiple database, we expect that the proposed method is robust against several attacks.

Keywords: Distributed watermarking · Multiple image database · Image sets

1 Introduction

1.1 Overview

With the rapidly increasing use of Internet, various form of digital data is proposed to share digital multimedia for everyone over the wold. Normal users can access to digital contents like electronic advertising, video, audio, digital repositories, electronic libraries, web designing, and so on. Also, users can copy the digital contents and distribute it easily via network. Digital copyright violations happen frequently because of the increased importance of digital contents. That makes the content providers need to focus on the protection that of copyright. The techniques for copyright protection are developed and researched more and more nowadays.

Digital watermarking technique is the promising technique for protecting the copyright of the valuable digital contents. This technique embeds the copyright information (*e.g* digital logo, author's name, identifier number, ...) into the digital contents without quality degradation. The embedded contents, also called cover contents, can be used to distribute to the users who bought the contents. When copyright disputes happened, the embedded information is needed

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2020
Published by Springer Nature Switzerland AG 2020. All Rights Reserved
N.-S. Vo and V.-P. Hoang (Eds.): INISCOM 2020, LNICST 334, pp. 163–180, 2020.
https://doi.org/10.1007/978-3-030-63083-6_13

to extract from the embedded contents to verify the copyright of the related parties. Therefore, the watermarking techniques are required to be robust against the common attacks on the embedded contents such as digital content processing, compression, RST (rotation, scaling, translation) attack, noise attacks, and so on. The copyright information should be successfully extracted even the embedded contents are adjusted by illegal users.

There are two big classifications of watermarking techniques such as spatial domain based techniques [1] and frequency domain based techniques [2] techniques. In general, the spatial domain techniques embed directly the copyright information into the pixels of the original digital contents. These techniques achieve high performance and do not degrade the quality of contents much. However, such techniques are not robust against even simple image processing attacks [3]. On the other hand, the frequency domain based watermarking techniques are employed to embed copyright information on the coefficients' values of the image after applying some frequency transforms (*i.e.* DCT, DFT, DWT, SVD). Frequency domain based watermarking techniques are mostly focused on real applications since it is robust and secure. The application of both techniques is data integrity, authentication, copyright protection, broadcast monitoring [4].

In the frequency domain, several digital watermarking methods are available in literature including Discrete Cosine Transform (DCT) [5], Discrete Wavelet Transform (DWT) [6], Singular Value Decomposition (SVD) [7], and so on. Such frequency domains are normally employed on the digital format that is employed in the real applications such as audio, image, text, and video. In this paper, we focus on the DCT-SVD based watermarking for a proposal of digital watermarking technique. DCT divides carrier signal into low, middle, and high frequency bands. DCT watermarking is classified into two types: Global DCT watermarking and Block-based DCT watermarking. That makes us possible to control the regions for watermark embedding and extraction. On the other hand, SVD transform decompose a information matrix into orthogonal matrices of singular values(eigen values). It is used to approximate the matrix decomposing the data into an optimal estimate of the signal and the noise components. That property is important for watermarking technique to be robust against noise filtering, compression, and forensic attacks. Based on this analytic, we choose DCT-SVD watermarking method to propose a new framework of distributed watermarking.

As mentioned above, general watermarking techniques are always applied on the normal digital format of multimedia contents. Recently, incorporating deep neural networks with image watermarking [8,9] has attracted increasing attention by many researchers. In this framework, researchers mostly focus on how to embed the watermarks into the trained or pre-trained model of deep neural networks (DNN) [10]. They almost do not focus on how to protect the copyright of image dataset provided for training and testing of DNN methods. That means there are not a watermarking method for deep learning image dataset such as

CIFAR-10[1], MNIST[2], MS-COCO[3], and so on. In our understanding, in order to make the dataset for deep learning method, the providers take much more time and effort to gather and to annotate the data. It also is updated to increase amount of dataset year by year. Therefore, the large datasets for deep learning are very valuable datasets to apply on real applications. In our knowledge, there is not a proposal of copyright protection applied on multimedia dataset. That means the copyright protection for large dataset is required as soon as possible.

1.2 Our Contributions

In this paper, we propose a new method of distributed watermarking for cross-domain of semantic large image database. We make a first version of watermarking method for dataset of deep learning algorithms. In order to keep the accuracy of deep learning model, we distribute the copyright information (watermark) on whole of dataset. The original watermark is separated into many scramble shadows to keep the secure of copyright information. Each shadow will be embedded into one image of dataset. The embeddable set of images are selected from the dataset by checking the semantic relationship of images. In summary, we briefly introduce our contributions as follows:

1. We propose a new distributed watermarking method for semantic image sets extracted from dataset published for deep learning applications. The semantic image set is defined as a set of images belonging one class from deep learning dataset. In case of video format, the semantic image can be defined as a set of frames from on shot video. This is the first consideration of copyright protection for deep learning publish dataset.
2. We have an idea to distribute the separated shadows over all image sets to keep the security of original watermark. Only the owners have the secret keys, can extract the original watermark to prove the ownership of dataset.
3. We try to apply the proposed method on the shots of video file in order to verify the efficiency of solution. Shot detection process can be simply performed by using some conventional method[4]. All frames from one shot are used to embed all scrambled shadows in order to improve the security.

We hope that our proposed method can be used for copyright protection of large deep learning dataset. Datasets providers can embed their information into whole of datasets and then publish it for everyone.

1.3 Roadmap

The rest of this paper is organized as follow: in Sect. 2, a brief overview of related techniques using in the paper. In additional, we define the concept of semantic

[1] http://www.cs.toronto.edu/~kriz/cifar.html.
[2] https://datahack.analyticsvidhya.com/contest/practice-problem-identify-the-digits/.
[3] http://cocodataset.org/.
[4] https://github.com/albanie/shot-detection-benchmarks.

large image dataset. In Sect. 3, the detailed steps of embedding and extraction watermark are explained. In Sect. 4, the simulation experimental results and discussion are shown. Section 5 gives conclusions of this paper.

2 Preliminary

In this study, we employ the combination of DCT and SVD transformation and generate the frequency domain, called DCT-SVD domain. After that, the scrambled shadows are embedded into DCT-SVD domain. Such kind of transformation is briefly explained as follows:

2.1 Discrete Cosine Transform

Discrete cosine transform (DCT) is a most popular linear transform domain that is used in processing of multimedia contents [11]. DCT is always applied on compression format of digital contents to remove statistical correlation. In addition, the frequency bands of DCT including high-frequency, middle-frequency, and low-frequency can be selected appropriately for data embedding. The DCT is defined as:

$$\hat{I}(u,v) = \frac{1}{\sqrt{M \times N}} C(u)C(v) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i,j) \cos\left(\frac{(2i+1)u\pi}{2M}\right) \cos\left(\frac{(2j+1)v\pi}{2N}\right) \tag{1}$$

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}}, & u = 0; \\ 1, & u > 0, \end{cases} \tag{2}$$

where $i, u = 0, 1, 2, \ldots, M-1$ and $j, v = 0, 1, 2, \ldots, N-1$. The inverse discrete cosine transform (iDCT) is defined as follows:

$$I(i,j) = \frac{1}{\sqrt{M \times N}} C(u)C(v) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \hat{I}(u,v) \cos\left(\frac{(2i+1)u\pi}{2M}\right) \cos\left(\frac{(2j+1)v\pi}{2N}\right) \tag{3}$$

where $I(i,j)$ is the intensity of image and $\hat{I}(u,v)$ is the DCT coefficients. The DCT low-frequency information contains the main information of image. Therefore, the visual quality of an image will be degraded significantly if the watermark is embedded into DCT low-frequency. In the other hand, the DCT high-frequency information contains the details edges of image, then, it is easily removed in lossy data compression. Based on analysis above, the DCT middle-frequency is more suitable for watermark embedding to maintain the visual quality of image and to keep the robustness of watermark [12].

2.2 Singular Value Decomposition

The singular value decomposition (SVD) is an important factorizing technique to decompose one matrix into three matrices. The SVD is defined as:

$$A = USV^T, \tag{4}$$

where A is an $M \times N$ matrix, U and V are two orthonormal matrices, S is a diagonal matrix consisting of the singular values (SVs) of A. The singular values (SVs) satisfy $s_1 \geq s_2 \geq s_3 \ldots \geq s_N \geq 0$ and the superscript T denotes matrix transposition.

If matrix A is a pixel of image, the SVs can keep stable with a slight perturbation. SVs are almost keep stable even the image is adjusted the pixel value under some attacks. Therefore, in order to improve the robustness of watermark after embedding, we choose to SVs for guaranteeing the copyright information.

2.3 Semantic Large Image Dataset

The concept of semantic large image dataset is not defined beforehand. In this paper, we define the semantic large image dataset as a set of images belonging one class of one object. In case of video file, the semantic large image dataset can be defined as a set of frames from one shot.

For instance, as the explanation of Kaggle[5], "CIFAR-10 is an established computer-vision dataset used for object recognition. It is a subset of the 80 million tiny images dataset and consists of 60,000 32×32 color images containing one of 10 object classes, with 6000 images per class.". That means that CIFAR-10 contains 10 object classes, with 6000 images per class. Therefore, we can extract ten semantic large image datasets from CIFAR-10, then we can embed the copyright information into 6000 images for each class.

2.4 Distributed Watermark

In order to keep the secret of watermark information, we choose the (n, n) secret sharing scheme [13] to distribute the original watermark. That implies that, n-shadows S are generated from the original watermark W XOR-ing with n-secret key K. In order to reveal the original watermark, we need to collect all n-shadows S with XOR-ing it.

The (n, n) secret sharing scheme is described in a pseudo-code style below in terms of its input, output, the construction procedure, and revealing procedure. In the construction procedure, our algorithm shows the way to compute the shadow watermark. In the revealing procedure, the algorithm that reveals the original watermark, is explained how to reconstruct the original watermark from the shadows.

According to the Algorithm 1, we can distribute the original watermark W into n shadows S_i. That can keep the secret of the original watermark (copyright information). Only the owners, who has the secret key K_i, can reveal the original W after collecting all shadows S_i.

[5] https://www.kaggle.com/c/cifar-10.

Algorithm 1: Distributed watermark - (n, n) secret sharing scheme

INPUT: Number of shadows n, the original watermark W
OUTPUT: n distinct matrices $\{S_1, ..., S_n\}$, called shadows watermark.
CONSTRUCTION: Generate $n + 1$ random secret keys $K = \{K_1, ..., K_n, K_{n+1}\}$. Compute n shadow watermark $\{S_1, ..., S_n\}$ with

$last.shadow = W$
for all random secret keys in K_i in K **do**
$S_i = last.shadow \oplus K_i$
$last.shadow = S_i$, for $i = 1, 2, ..., n$
end for

REVEALING: Reveal original watermark W from shadows $S = \{S_1, ..., S_{n-1}\}$.
$last.watermark = S_n$
for all random secret keys in K_i in K **do**
$last.watermark = last.watermark \oplus K_i$, for $i = 1, 2, ..., n - 1$
end for

3 Our Proposed Method

Our proposed watermarking method can be applied on the large image dataset for deep learning applications. However, the semantic large images meaning of one class from dataset can be considered as a shot in the video format. Therefore, in order to verify our idea simply, we apply our method on video format. Our method consists of three processes: shadows construction from watermark, video embedding, video extraction, and watermark revealing.

3.1 Shadows Construction Process

To keep the secret of watermark information, we scramble the original watermark by using the (n, n) secret sharing scheme described in Sect. 2.4. We use the process CONSTRUCTION in Algorithm 1 to generate n shadows S_i from the original watermark W with the secret keys K_i, where $i = 1, 2, ..., n$.

The shadows S_i can be seen as Fig. 1. All shadows are randomized based on the secret keys, therefore, only the owner, who keeps the key, can reveal the watermark information.

3.2 Video Embedding Method

Our proposed watermarking method is shown in Fig. 2. The process to embed the shadows into the frames of one shot can be described as follows:

Step 1: The shadows watermark image S_i of size $p \times p$ is converted to a $1 \times p^2$ binary watermark sequence S_i^o. In our paper, $p = 32$.

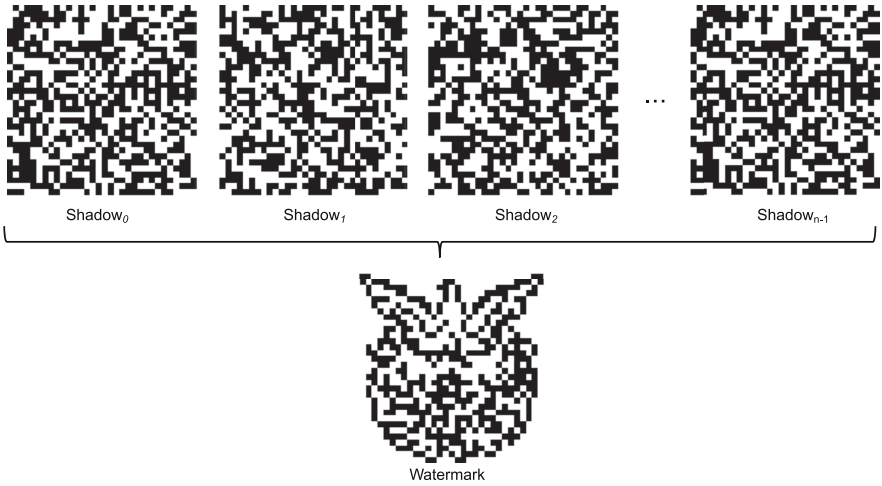

Fig. 1. (n, n) secret sharing scheme.

Step 2: Original video V is separated into multiple shots by using prepared algorithm [14] beforehand.
Step 3: DCT is performed on each frame from one shot to create the DCT frequency domain for all R, G, B plane. To improve the robustness and high quality of the proposed watermarking scheme, the low-frequency region is selected for watermark insertion.
Step 4: The low-frequency region is segmented into non-overlapping blocks C_i of size $4 \times 4, i = 1, 2, ..., N$.
Step 5: For each block:

(1) DCT is performed on each block C_i. The coefficient $C_i(0, 0)$ of each block is collected into the matrix A.
(2) The matrix A is segmented into non-overlapping blocks A_i of size $4 \times 4, i = 1, 2, ..., M$.
(3) SVD is applied on the matrix A_i and the largest singular value is extracted as follows:

$$[U_k, S_k, V_k] = SVD(A_k), k = 1, 2, ...M/4 \tag{5}$$

where U_k, V_k and S_k are the results of SVD operation, respectively. Let $x = S_k(0, 0)$ represents the largest SV of matric A_k.
(4) The shadows watermark sequence S_i^o is embedded by modifying the values of x. If $S_i^o = 1$, then $x = (\lfloor (x * Q)/2 \rfloor * 2)/Q$. If $S_i^o = 0$, then $x = (\lfloor (x * Q + 1)/2 \rfloor * 2 - 1)/Q$. In this term, Q is the embedding strength factor.
(5) The modified matrix S'_k is generated by altering their largest singular values with x.

$$S'_k(0, 0) = x \tag{6}$$

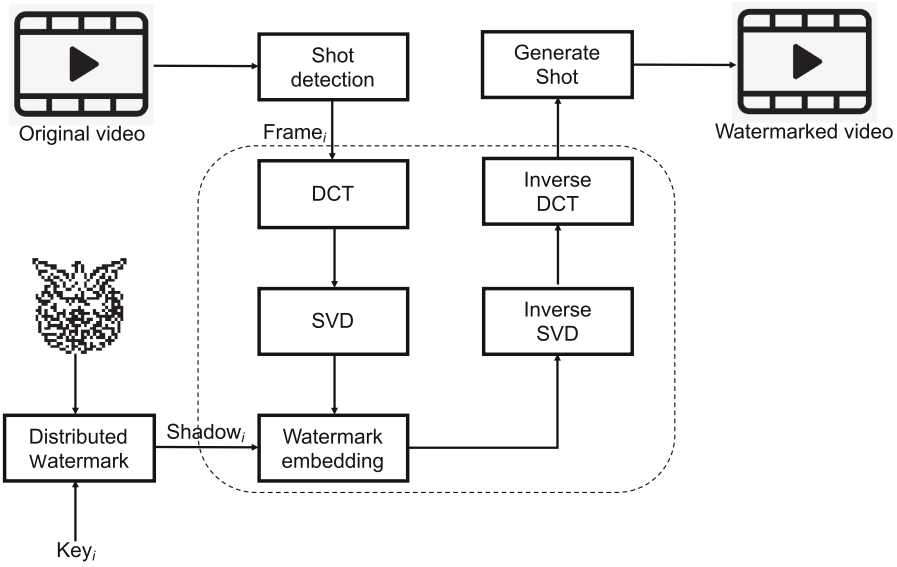

Fig. 2. Video embedding method

(6) Modified matrices A'_k are constructed by the inverse SVD operation.

$$A'_k = U_k S'_k V_k, k = 1, 2, ...M/4 \tag{7}$$

All elements of matrices A'_k are mapped back to their original positions in coefficient matrix A. Then a watermarked block C'_i is produced by the inverse DCT.

Step 6: After Step 5, the embedded frame are generated. Collect all frames and re-create all shots to generate a watermarked video V'.

3.3 Video Extraction Method

To extract the watermark information from a watermarked video V', we do not need the original video V. The secret key K_i, and the number of shadows n are required. The process of video extraction is shown in Fig. 3.

Step 1: The watermarked video V^* is separated into multiple shots by using prepared algorithm [14] beforehand.

Step 2: By applying the DCT operation on each frame from one shot, the DCT frequency domain for all R, G, B plane is obtained.

Step 3: The low-frequency region is segmented into non-overlapping blocks C_i^* of size $4 \times 4, i = 1, 2, ..., N$.

Step 4: For each block:

(1) By applying DCT operation on each block C_i^*, the matrix A^* is generated by collecting all the coefficient $C_i^*(0, 0)$ of each block.

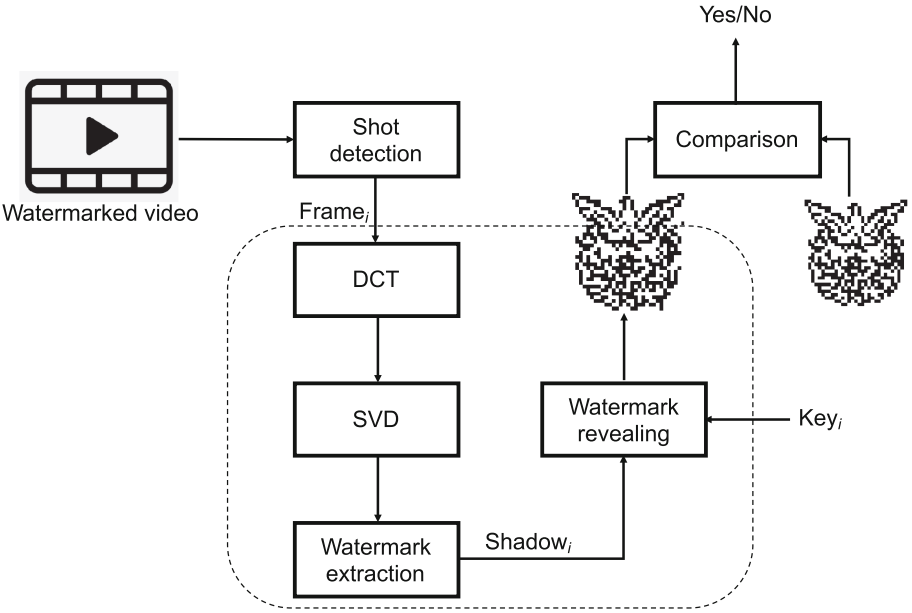

Fig. 3. Video extraction method

(2) The matrix A^* is segmented into non-overlapping blocks A_i^* of size $4 \times 4, i = 1, 2, ..., M$.
(3) SVD operation is applied on the matrix A_i^* and the largest singular value is extracted as follows:

$$[U_k^*, S_k^*, V_k^*] = SVD(A_k^*), k = 1, 2, ...M/4 \tag{8}$$

where U_k^*, V_k^* and S_k^* are the results of SVD operation, respectively. The values $x^* = S_k^*(0, 0)$ represents the largest SV of matrix A_k^*.
(4) The embedded watermark bits can be extracted as follows:

$$votes = \begin{cases} votes + 1, & \text{if } \lfloor x^* * Q \rfloor \% 2 = 0; \\ 0, & \text{other,} \end{cases} \tag{9}$$

where $votes$ is the number of even value of x^*. This method called "voting method". Therefore, the shadows watermark sequence $S_i^{o'}$ can be extracted from $votes$ as follows:

$$S_i^{o'} = \begin{cases} 1, & \text{if } votes > T; \\ 0, & \text{other,} \end{cases} \tag{10}$$

where T is threshold value that is predefined beforehand.

Step 5: The watermark bits from all watermarked blocks of all frames are extracted by repeating Step 4.

Fig. 4. Test video and watermark

Step 6: From all extracted $S_i^{o'}$, we can construct all shadows watermark S_i'. By using REVEALING process in the Algorithm 1, we can reveal the watermark W'.

3.4 Watermark Revealing Process

To obtain the watermark from the extracted shadows $S_i^{o'}$, we employ the REVEALING process in Algorithm 1. We also use the information n of (n, n) secret sharing scheme and the secret keys K_i that is used in CONSTRUCTION process. If we apply on all extracted shadows $S_i^{o'}$, we can obtain the watermark W' which is shown in Fig. 1.

4 Experimental Results and Analysis

4.1 Experimental Environment

Our proposed method is implemented in Python version 3.7.6. All experimental results are obtained in MacBook Pro, macOS version 10.13.

To evaluate the performance of the proposed watermarking method, some video "Akiyo", "Container", and "Foreman" [6] with 300 frames are chosen as the test videos. It is shown in Fig. 4 (a)∼(c). A binary watermark W of size 32×32 shown in Fig. 4 (d) is selected as the watermark image. In order to make all frames are similar to deep learning dataset, we scale the size of all frames with the same size as 1024×1024.

In order to generate the shadows watermark S_i^o, we employ the (4, 4) secret sharing scheme. That means $n = 4$. To generate the secret key K_i, we use the function $randint(0, 255)$[7] with secret $seed = 101$. The embedding strength factor Q is set at 15. The threshold T is set at 2.

4.2 Imperceptibility Measure

To evaluate the imperceptibility of watermaked video, the peak signal-to-noise ratio (PSNR) is adopted. The PSNR can measure the quality of all watermarked frames [15], then average PSNR value for the embedded video can be calculated.

[6] http://trace.eas.asu.edu/yuv/.
[7] https://docs.python.org/3/library/random.html.

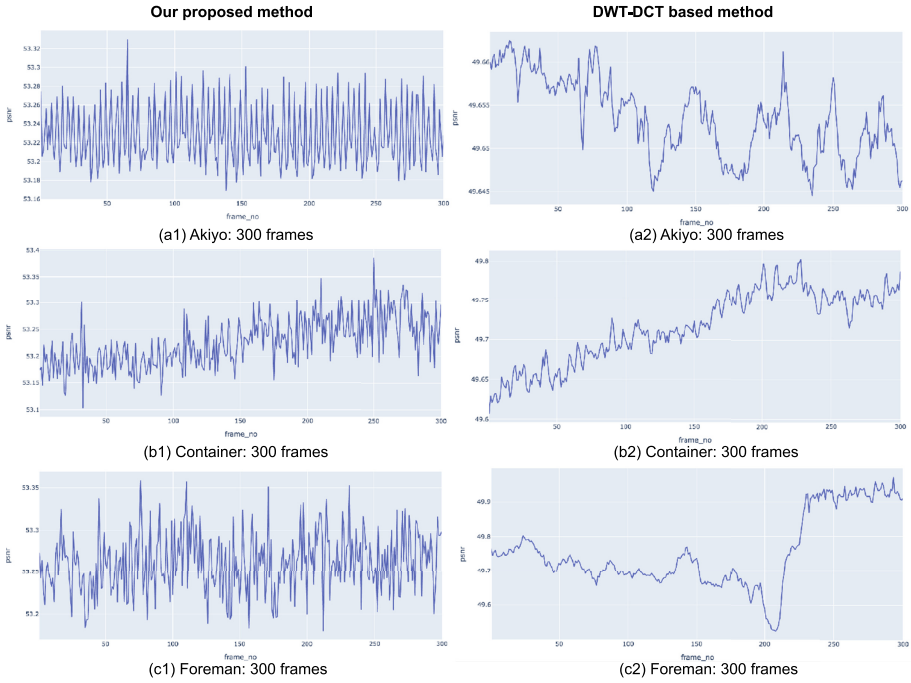

Fig. 5. All PSNR values of all frames

$$PSNR = 10 \log_{10} \frac{MAX^2}{MSE}, \tag{11}$$

where MAX is the maximized value of pixel, *e.g.* $MAX = 255$. And MSE is mean squared error.

$$MSE = \frac{1}{H * W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} [I(i,j) - \hat{I}(i,j)]^2, \tag{12}$$

where H and W are the height and width of frame. I and $\hat{I}$ are the original frame and the watermarked frame, respectively.

In order to compare the efficiency of our proposed method, we also implement the similar embedding process and extraction process on the DWT-DCT domain, which is usually employed for image or video watermarking field [18–20], called **DWT-DCT based method**.

We calculated all PSNR values of all frames from test videos and obtained the average PSNR values. All PSNR values of all frames from three test videos are shown in Fig. 5. The average PSNR values from those of all frames are shown in Fig. 6. Based on our experimental results, our proposed method achieved better results than that of DWT-DCT based method.

To show the visualized quality of one frame from test videos, we extracted 1st frame and calculate the PSNR value from those. Such results are given in Fig. 7. It is clear that our results (Fig. 7(a1) (b1), (c1)) is better than the results (Fig. 7(a3), (b3), (c3)) of DWT-DCT based method.

4.3 Robustness Measure

To evaluate the robustness of extraction method, we calculate the NC (Normalized Correlation) [16] value of each frame.

$$NC = \frac{\sum_{i=0}^{31}\sum_{j=0}^{31}[W(i,j)W'(i,j)]}{\sum_{i=0}^{31}\sum_{j=0}^{31}[W(i,j)]^2}, \tag{13}$$

where W and W' are the original watermark and the revealed watermark that is reconstructed in Sect. 3.4. If the value of NC is close to 1, it means that the robustness is better. Normally, the NC value is acceptable if it is 0.75 or higher.

To verify the robustness of our proposed method, various attacks including median blur, Gaussian noise, Salt and Pepper noise, JPEG compression, overlay attack, color attack, scaling, and rotation, are implemented. The details about these common attacks are described as follows:

Filtering Attack. In this attack, the watermarked frames are corrupted with Gaussian filtering and median filtering, respectively. Gaussian filtering removes the high frequency information of watermarked frames by using blurring based on Gaussian function. Median filtering replaces the original pixel value of watermarked frames with the median value in the 3×3 window. The results of extracted watermark are given in Fig. 8. As can be seen, the performance of the proposed algorithm under filtering attacks is better than DWT-DCT based method.

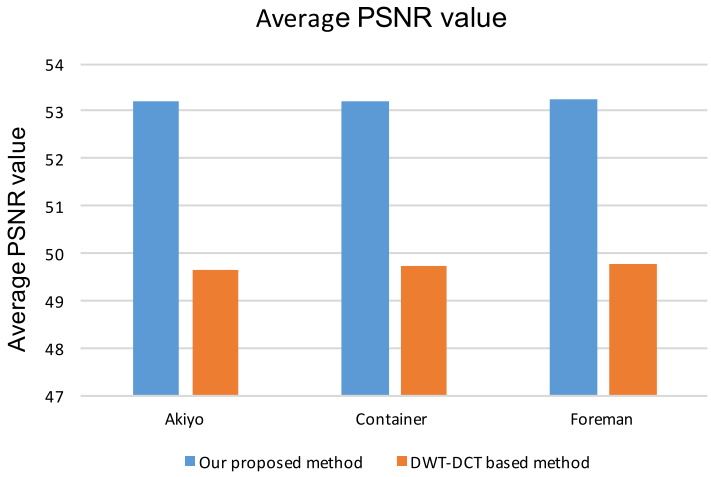

Fig. 6. Average PSNR value of three test videos

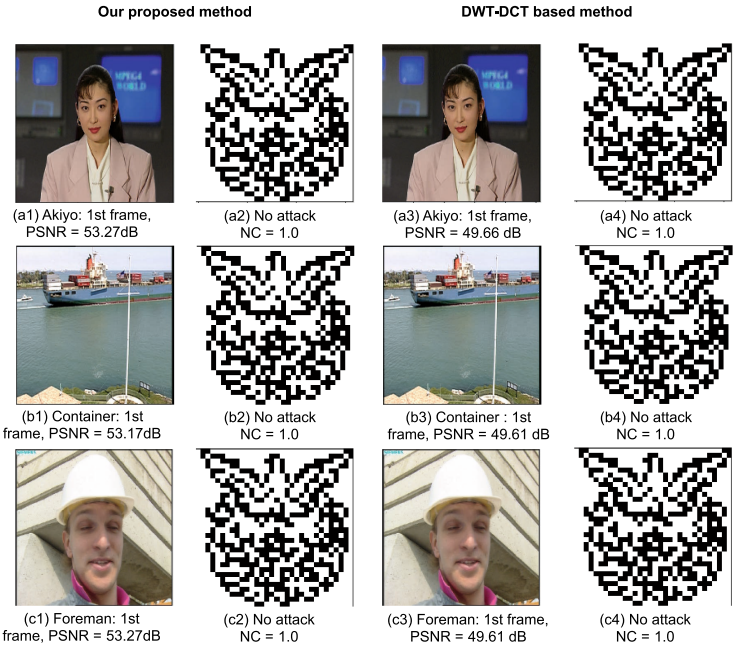

Fig. 7. The PSNR value and NC value of 1st frame from three test videos

Noise Attack. In the noise attack experiment, the watermarked frames are degraded by two kinds of noises attack such as Gaussian noise with variance 0.005, Salt and Pepper with variance 0.01. The results of extracted watermarks are shown in Fig. 9. As can be seen, both methods are not robust against Salt and Pepper attack. However, those are robust against Gaussian noise attack.

Geometric Attacks. Two geometric attacks are employed in this paper. For scaling operation, the watermarked frames are scaled down to 50%, then are scaled up to 100%. In the rotation experiment, the watermarked frames are rotated by 5 in the counterclockwise direction, then are re-rotated by 5 in the opposite direction. The results of our experiment are shown in Fig. 10. As can be seen from Fig. 10, our proposed method can extract exactly the watermark information. On the other hand, DWT-DCT based method is not robust against strong scaling attacks.

JPEG Attack. In general, the JPEG (Joint Photographic Experts Group) compression[8] is a most popular image compression technique in digital watermarking. In this experiment, the watermarked frames are compressed with quality factor set 75 that is set for normal JPEG.

[8] https://jpeg.org/.

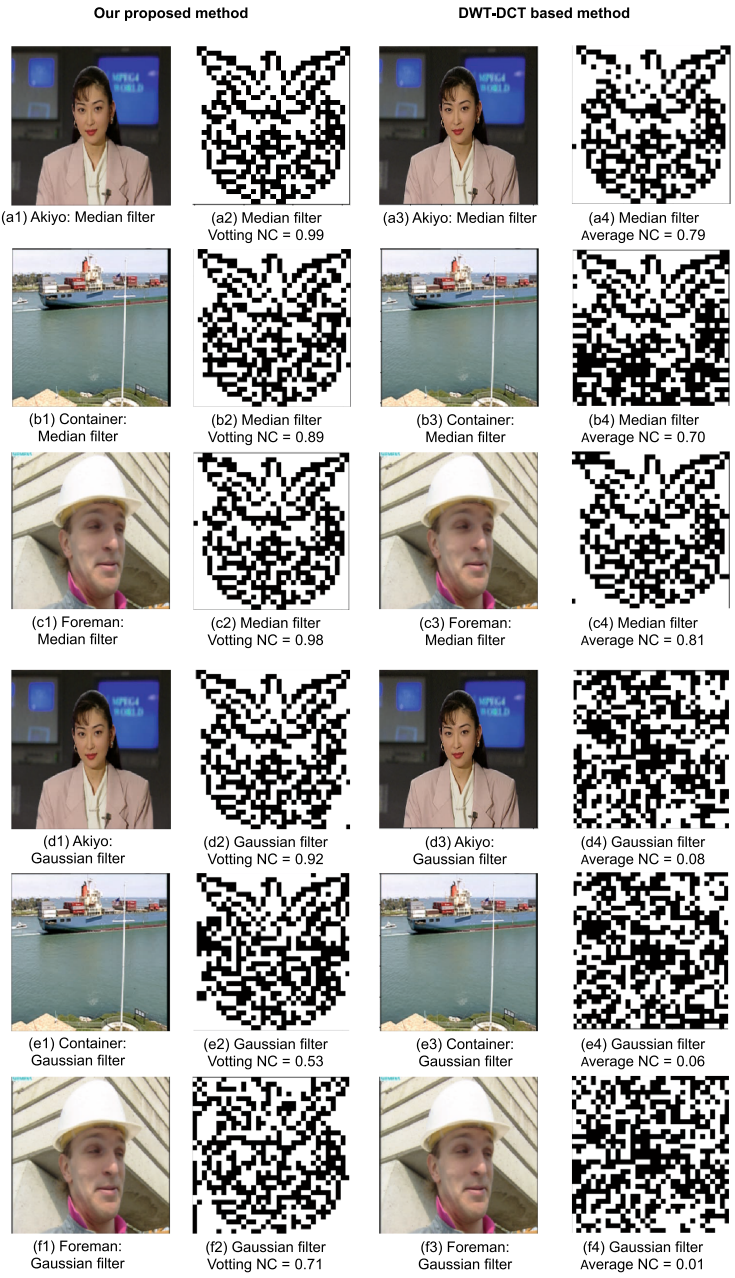

Fig. 8. Filtering attacks: Median and Gaussian filtering

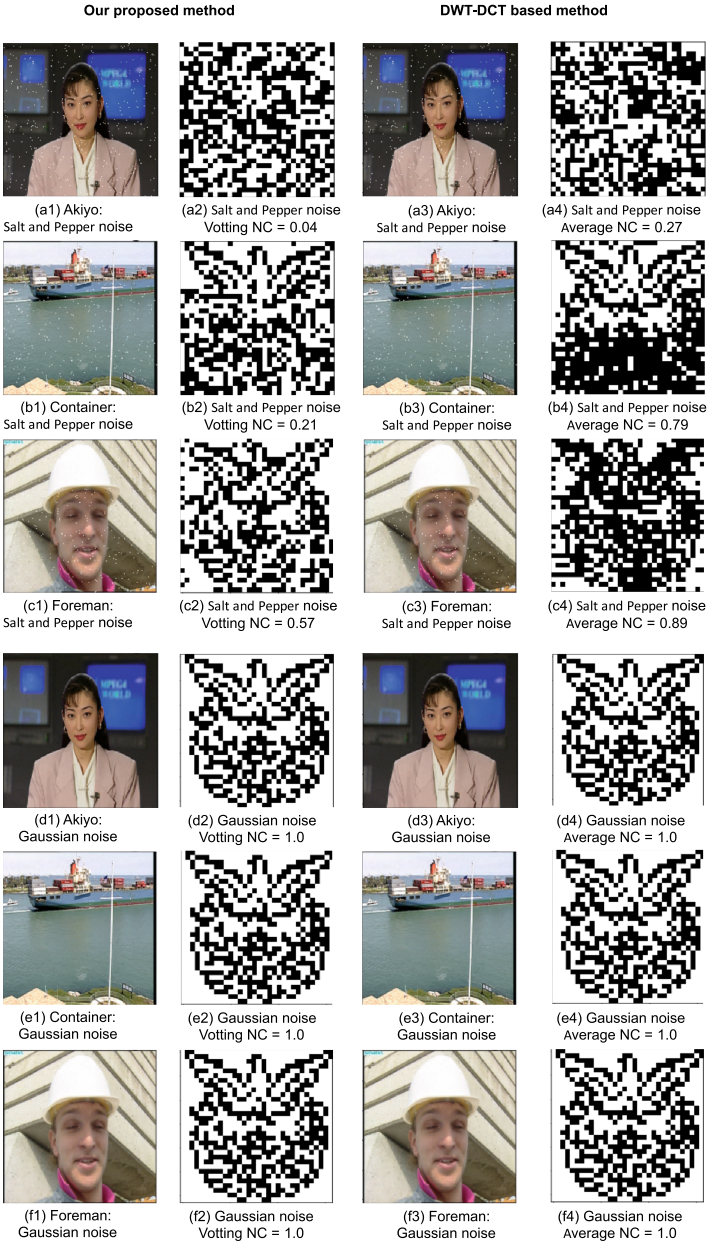

Fig. 9. Noise attacks: Salt & Pepper and Gaussian noise

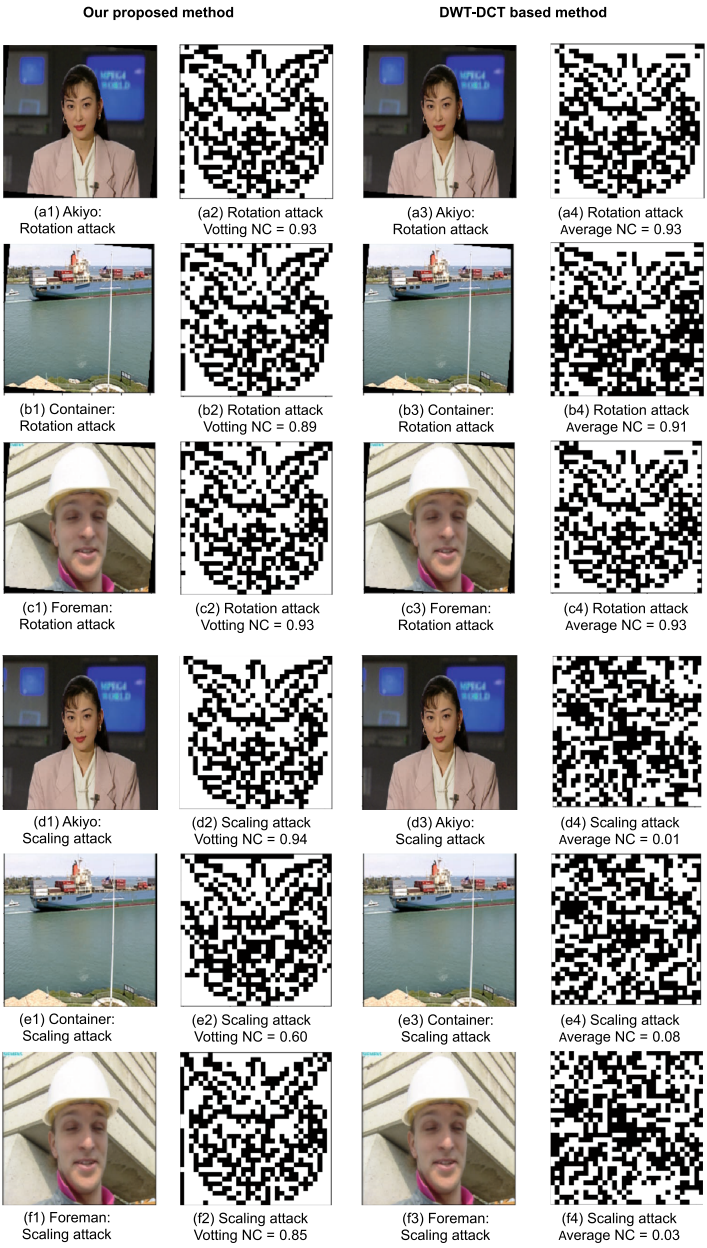

Fig. 10. Geometric attacks: Scaling and Rotation attacks

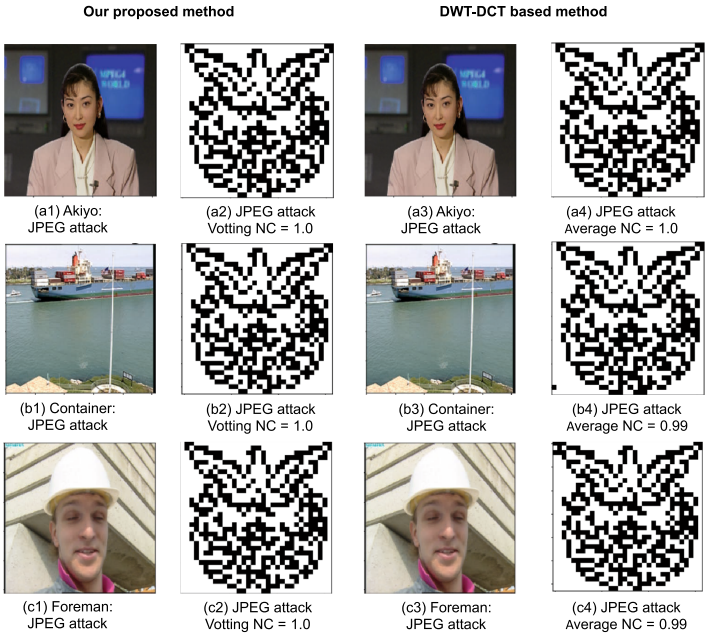

Fig. 11. JPEG attacks

Figure 11 shows that both methods robust against the JPEG compression. Based on these results, we can conclude that our proposed method and DWT-DCT based method are acceptable for JPEG compression.

5 Conclusions

In this paper, we have proposed a method to distribute original watermark to generate scrambled shadows, then embed such shadows into the set of frames from shots. We embed the watermark information into DCT-SVD domain, therefore, it is robust against some attacks such as median blur, Gaussian noise, Salt and Pepper noise attacks, JPEG compression, and geometric attacks. Compared with other related works, our proposed watermarking method performs better in terms of invisibility and robustness. Our proposed watermarking algorithm can be extended for audio signal processing, distributed database.

Acknowledgement. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.01-2019.12.

References

1. Mukherjee, D.P., Maitra, S., Acton, S.T.: Spatial domain digital watermarking of multimedia objects for buyer authentication. IEEE Trans. Multimedia **6**(1), 1–15 (2004)

2. Lin, S.D., Chen, C.: F, "A robust DCT-based watermarking for copyright protection,". IEEE Trans. Consum. Electron. **46**, 415–421 (2000)
3. Ghadi, M., Laouamer, L., Nana, L., Pascu, A.: A blind spatial domain-based image watermarking using texture analysis and association rules mining. Multimedia Tools Appl. **78**(12), 15705–15750 (2018). https://doi.org/10.1007/s11042-018-6851-2
4. Vasudev, R.: A review on digital image watermarking and its techniques. J. Image Graph. **4**(2), 150–153 (2016)
5. Das, C., Panigrahi, S., Sharma, V.K., Mahapatra, K.K.: A novel blind robust image watermarking in DCT domain using inter-block coefficient correlation. AEU-Int. J. Electron. Commun. **68**(3), 244–253 (2014)
6. Srivastava, A., Saxena, P.: DWT? DCT? SVD based semiblind image watermarking using middle frequency band. IOSR J. Comput. Eng. **12**(2), 63–66 (2013)
7. Lagzian, S., Soryani, M., Fathy, M.: A new robust watermarking scheme based on RDWT-SVD. Int. J. Intell. Inf. Process. **2**(1), 22–29 (2011)
8. Sabah, H., Haitham, B.: Artificial neural network for steganography. Neural Comput. Appl. **26**(1), 111–116 (2015)
9. Alexandre, S.B., David, C.J.: Artificial neural networks applied to image steganography. IEEE Latin Amer. Trans. **14**(3), 1361–1366 (2016)
10. Uchida, Y., Nagai, Y., Sakazawa, S., Satoh, S.: Embedding watermarks into deep neural networks. In: ICMR (2017)
11. Leng, L., Zhang, J., Xu, J.: Dynamic weighted discrimination power analysis in DCT domain for face and palmprint recognition. In: 2010 International Conference on Information and Communication Technology Convergence (ICTC), pp. 467–471. IEEE (2010)
12. Kaur, B., Kaur, A., Singh, J.: Steganographic approach for hiding image in DCT domain. Int. J. Adv. Eng. Technol. **1**(3), 72–78 (2011)
13. Wang, D., Zhang, L., Ma, N., Li, X.: Two secret sharing schemes based on Boolean operations. Pattern Recogn. **40**(10), 2776–2785 (2007)
14. https://github.com/albanie/shot-detection-benchmarks
15. Thanh, T.M., Heip, P.T., Tam, T.M., Tanaka, K.: Robust semi-blind video watermarking based on frame-patch matching. AEU Int. J. Electr. Commun. **68**(10), 1007–1015 (2014)
16. Thanh, T.M., Tanaka, K.: An image zero-watermarking algorithm based on the encryption of visual map feature with watermark. J. Multimedia Tools Appl. **76**, 13455–13471 (2017)
17. Thanh, T.M., Iwakiri, M.: A proposal of digital rights management based on incomplete cryptography using invariant Huffman code length feature. Multimedia Syst. **20**(2), 127–142 (2013). https://doi.org/10.1007/s00530-013-0327-z
18. Abdulrahman, A.K., Ozturk, S.: A novel hybrid DCT and DWT based robust watermarking algorithm for color images. Multimedia Tools Appl. **78**(12), 17027–17049 (2019). https://doi.org/10.1007/s11042-018-7085-z
19. Chow, Y.-W., Susilo, W., Tonien, J., Zong, W.: A QR code watermarking approach based on the DWT-DCT technique. Faculty of Engineering and Information Sciences - Papers: Part B. 389 (2017)
20. Hu, H.-T., Hsu, L.-Y.: Collective blind image watermarking in DWT-DCT domain with adaptive embedding strength governed by quality metrics. Multimedia Tools Appl. **76**(5), 6575–6594 (2016). https://doi.org/10.1007/s11042-016-3332-3