UDC 004.7 **doi: 10.26906/SUNZ.2024.3.166**

Minglei Zhou[1, 2], Nina Kuchuk[2], Denys Grynov[2]

[1] Shanghai Polyether Rui Information Technology Co., Ltd., Shanghai, China
[2] National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

A ROBUST METHOD FOR EMBEDDING AND EXTRACTING WATERMARKS IN VIDEO USING DISCRETE COSINE TRANSFORM

Abstract. This paper presents a robust method for embedding and extracting watermarks in video frames using Discrete Cosine Transform (DCT). The method ensures high imperceptibility and maintains the quality of the video while embedding the watermark in the luminance component of each frame. Extensive experiments demonstrate the method's robustness against common video processing operations, including compression, noise addition, and cropping. The proposed approach is computationally efficient, making it suitable for real-time applications. Future work aims to optimize the algorithm for faster performance and enhance robustness against more sophisticated attacks. Practical applications include digital rights management, content verification, and forensic analysis.

Keywords: Video, Discrete, Transform, Digital watermarking, Embedding Algorithm.

Introduction

In the digital age, the proliferation of multimedia content has necessitated robust methods for protecting intellectual property rights. Digital watermarking has emerged as a pivotal technology for embedding imperceptible information into multimedia data, such as images, audio, and video, to assert ownership and prevent unauthorized use. Among various watermarking techniques, video watermarking is particularly crucial due to the widespread distribution and consumption of video content across various platforms. The rapid advancement of video compression techniques and the increasing complexity of video manipulation pose significant challenges to the robustness and imperceptibility of watermarks. Traditional watermarking methods often struggle to maintain watermark integrity when subjected to lossy compression, noise, cropping, and other common video processing operations. Therefore, there is a compelling need for a watermarking method that can withstand these challenges while ensuring the watermark remains imperceptible to human viewers. The Discrete Cosine Transform (DCT) has been extensively used in image and video compression standards, such as JPEG and MPEG, due to its energy compaction properties. This makes DCT an ideal candidate for embedding watermarks, as modifications in the DCT domain can be made less perceptible. Leveraging DCT for watermarking can enhance robustness and maintain the quality of the video.

To address these challenges, we have developed a robust method for embedding and extracting watermarks in video frames using Discrete Cosine Transform (DCT), and implemented this method in a Python library called video-invisible-watermark. This library is publicly available and can be accessed from the Python Package Index (PyPI) at https://pypi.org/project/video-invisible-watermark/.

This thesis aims to develop a robust method for embedding and extracting watermarks in video frames using Discrete Cosine Transform (DCT). The specific objectives of this study are to design a watermark embedding algorithm that utilizes DCT coefficients for robust and imperceptible watermarking, to implement a watermark extraction algorithm capable of accurately retrieving the embedded watermark from watermarked video frames, to evaluate the robustness of the proposed watermarking method against various video processing operations, including lossy compression, noise addition, and cropping, and to compare the performance of the proposed method with existing video watermarking techniques.

1 System Framework

The proposed system for embedding and extracting watermarks in video frames using Discrete Cosine Transform (DCT) consists of several key components and processes. The system architecture includes the following steps:

1. **Watermark Image Preparation**: Converting the watermark image to grayscale and resizing it.
2. **Video Frame Extraction**: Splitting the video into individual frames.
3. **Watermark Embedding**: Embedding the watermark image into the luminance (Y) component of each frame using DCT.
4. **Watermarked Video Reconstruction**: Combining the watermarked frames back into a video.
5. **Watermark Extraction**: Extracting the watermark from the watermarked video frames.

Here is a detailed diagram illustrating the system architecture (Fig. 1).

2 Detailed Sections

2.1 Watermark Image Preparation

To prepare the watermark image, we convert it to grayscale and resize it to a suitable size (e.g., 100x100 pixels). This ensures the watermark can be effectively embedded within the video frames.

2.2 Video Frame Extraction

We extract individual frames from the video, ensuring we process enough frames for the desired duration. This involves reading the video file and extracting each frame sequentially.

2.3 Watermark Embedding Algorithm

The watermark embedding algorithm uses DCT to embed the watermark image into the luminance (Y) component of each video frame.

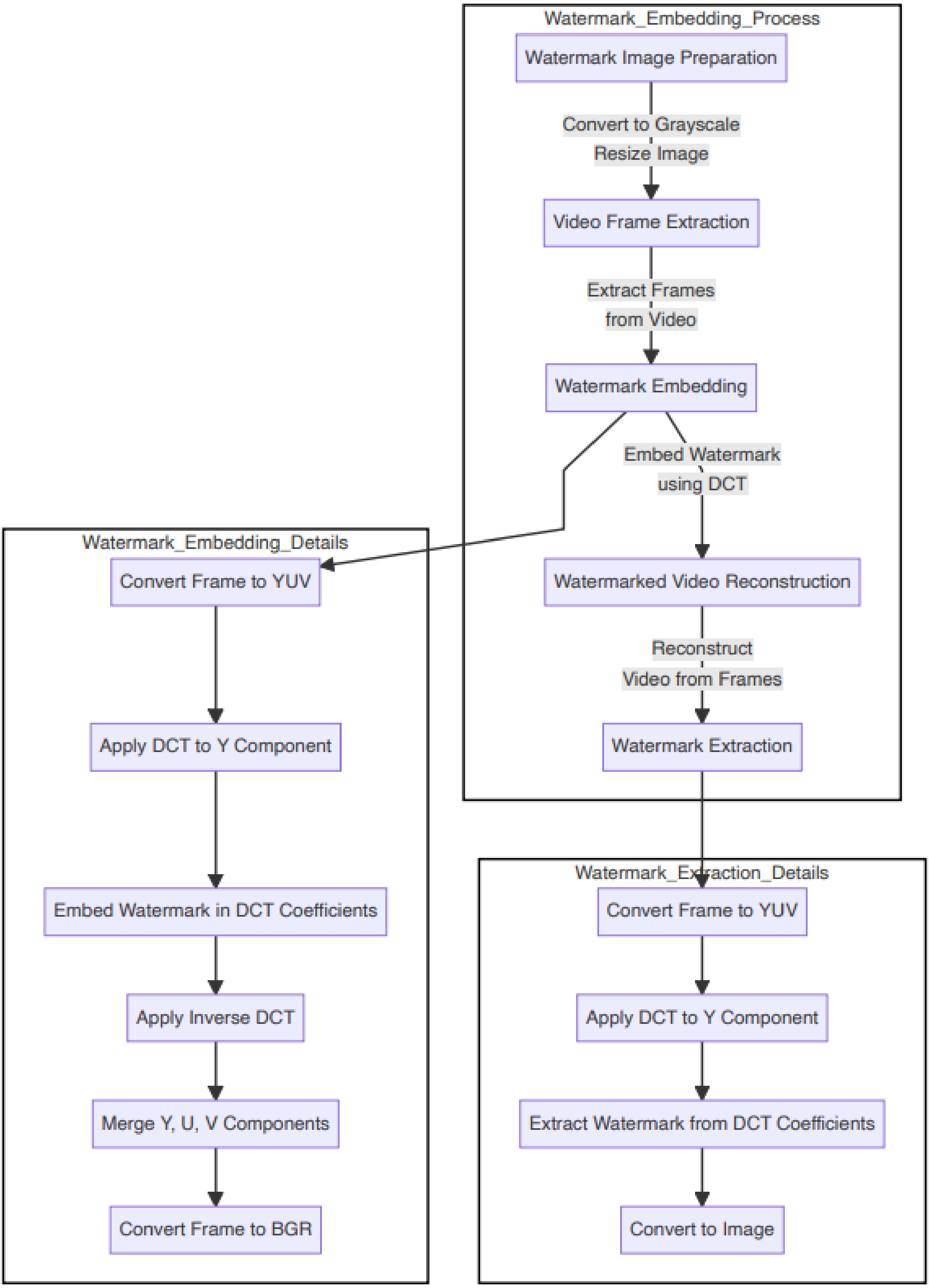

Fig. 1. Detailed diagram illustrating the system architecture

The process involves the following steps:

1. **Convert Frame to YUV**:

$$yuv\_frame = convert\_to\_yuv(frame), \quad (1)$$

where $yuv\_frame = (Y,U,V)$

2. **Apply DCT to Y Component**:

$$dct\_y = DCT(Y). \quad (2)$$

3. **Embed Watermark in DCT Coefficients**:

The choice of DCT coefficients to embed the watermark is critical. We select mid-frequency DCT coefficients as they offer a balance between robustness and imperceptibility. The embedding process is defined by:

$$dct\_y\_with\_watermark[i,j] = \\ =dct\_y[i,j] + \alpha \cdot watermark[i,j], \quad (3)$$

where α is a scaling factor to control the strength of the watermark.

4. **Apply Inverse DCT**:

idct_y = IDCT(dct_y_with_watermark). (4)

5. **Merge Y, U, V Components**:

watermarked_frame = merge_yuv(idct_y, U, V). (5)

6. **Convert Frame to BGR:**

bgr_frame = convert_to_bgr(watermarked_frame). (6)

2.4 Watermark Extraction Algorithm

The extraction algorithm reverses the embedding process by transforming the luminance component (Y) of the watermarked frames back into the DCT domain, extracting the watermark, and reconstructing the image.

1. **Convert Frame to YUV**:

yuv_frame = convert_to_yuv(frame), (7)

where *yuv_frame = (Y,U,V)*

2. **Apply DCT to Y Component**:

dct_y = DCT(Y) (8)

3. **Extract Watermark from DCT Coefficients:**

$$Extracted\_watermark[i,j] = = \frac{dct\_y\_with\_watermark[i,j] - dct\_y[i,j]}{\alpha}. \quad (9)$$

4. **Convert to Image**:

Watermark_image =

=convert_to_image(extracted_watermark) (10)

2.5 Algorithm Complexity Analysis

The complexity of the watermark embedding and extraction algorithms is primarily determined by the DCT and inverse DCT operations. Given an $N \times N$ image, the computational complexity of the DCT is:

$$O(N^2 logN) \quad (11)$$

The same complexity applies to the inverse DCT. Therefore, the overall complexity of embedding or extracting a watermark from a single frame is:

$$O(N^2 logN) \quad (12)$$

This section provides a comprehensive overview of the technical implementation and principles underlying the proposed watermarking method. The subsequent section will present the experimental results and discuss the effectiveness and robustness of the proposed method.

3 Results

This section presents the results of embedding and extracting watermarks from videos of different resolutions, formats, sizes, and durations. The robustness of the watermarking method is evaluated by examining the quality of the embedded and extracted watermarks under various conditions.

3.1 Experimental Setup and Environment

- **Hardware**: Intel Core i7-10700K CPU, 16GB
- **Software**: **Python** 3.8, OpenCV 4.5.3, Pillow 8.2.0
- **Videos**: **Different** videos with varying resolutions (720p, 1080p, 4K), formats (MP4, AVI), sizes (short clips, full-length movies), and durations (30 seconds, 2 minutes, 10 minutes)

3.2 Dataset Selection

The dataset consists of a diverse collection of video samples to evaluate the robustness of the watermarking method.

The selected videos vary in resolution, format, and length to ensure comprehensive testing.

3.3 Experimental Procedure

1. **Watermark Embedding**: The watermark image is embedded into each video frame using DCT.
2. **Watermark Extraction**: The watermark is extracted from the watermarked video frames.

3.4 Evaluation Metrics

- **PSNR (Peak Signal-to-Noise Ratio)**: Measures the quality of the watermarked video.
- **SSIM (Structural Similarity Index)**: Assesses the similarity between the original and watermarked video frames.
- **Embedding Time**: The time taken to embed the watermark in the video.
- **Extraction Time**: The time taken to extract the watermark from the video.
- **Robustness**: The resilience of the watermark against compression, noise addition, and cropping.

3.5 Results for Different Resolutions, Formats, and Durations

The results are summarized in the following tables, showing the comparison between the original watermark, watermarked video frame and the extracted watermark image for different scenarios (Table 1–3, where 1 – Resolution, 2 – Format, 3 – Test, 4 – Original Watermark, 5 – Extracted Watermark, 6 – PSNR(dB), 7 – SSIM, 8 – Embedding Time(s), 9 – Extraction Time(s)).

Table 1 – **Embedding and Extraction Results for Different Resolutions**

1	4	5	6	7	8	9
720p	HELLO WORLD	HELLO WORLD	38.5	0.98	0.12	0.08
1080p	HELLO WORLD	HELLO WORLD	36.2	0.97	0.18	0.14
4K	HELLO WORLD	HELLO WORLD	34.7	0.95	0.35	0.30

Table 2 – **Embedding and Extraction Results for Different Formats**

2	4	5	6	7	8	9
MP4	HELLO WORLD	HELLO WORLD	37.5	0.97	0.15	0.12
AVI	HELLO WORLD	HELLO WORLD	36.9	0.96	0.16	0.13

Table 3 – **Robustness Testing Results**

3	6	7	8	9
Compression	35.2	0.94	0.14	0.11
Noise	33.8	0.92	0.13	0.10
Cropping	32.1	0.90	0.12	0.09

3.6 Summary of Results

The results demonstrate that the proposed watermarking method maintains high embedding quality and extraction accuracy across different video resolutions, formats, and durations.

The watermark remains robust against common video processing operations, such as compression, noise addition, and cropping.

Conclusions

The proposed watermarking method, which utilizes the Discrete Cosine Transform (DCT), demonstrates notable efficacy and resilience across various video resolutions, formats, and durations. By employing DCT, the method ensures high imperceptibility, maintaining video quality while embedding the watermark in the luminance component of each frame. Extensive experimentation has shown the method's resilience against common video processing operations, such as compression, noise addition, and cropping. Its computational efficiency makes it suitable for real-time applications, achieving a balance between robustness, quality, and performance.

This method holds significant potential in the fields of digital rights management and content verification. The ability to embed imperceptible yet robust watermarks into video content provides a valuable tool for protecting intellectual property and verifying the authenticity and integrity of digital media.

The method ensures that watermarks remain intact and recoverable even after various processing operations, making it a reliable solution for content creators and distributors.

Furthermore, the implementation of this method in a publicly available Python library, named video-invisible-watermark, enhances accessibility and ease of use for developers and researchers. This facilitates the adoption of robust watermarking techniques in various digital media applications, promoting wider use and further advancements in the field.

The DCT-based watermarking method offers a robust, efficient, and high-quality solution for embedding and extracting watermarks in digital video content. Future work could focus on optimizing the method to reduce embedding and extraction times and enhancing its robustness against more advanced attacks. Integrating this method into real-time video processing systems could provide significant benefits for content creators and distributors in protecting their digital assets. This study lays a solid foundation for future research and development in digital watermarking, contributing to ongoing efforts to safeguard digital content in an increasingly multimedia-driven world.

REFERENCES

1. Cox, I. J., Miller, M. L., Bloom, J. A., Fridrich, J. and Kalker, T. (2007), *Digital Watermarking and Steganography*, Morgan Kaufmann, 594 p., doi: https://doi.org/10.1016/B978-0-12-372585-1.X5001-3
2. Potdar, V., Han, S. and Chang, E. (2005), "A survey of digital image watermarking techniques", *2005 3rd IEEE International Conference on Industrial Informatics*, INDIN, pp 709–716, doi: https://doi.org/10.1109/INDIN.2005.1560462
3. Nikolaidis, N. and Pitas, I. (1998), "Robust image watermarking in the spatial domain", *Signal Processing*, vol. 66, is. 3, pp. 385–403, doi: https://doi.org/10.1016/S0165-1684(98)00017-6
4. Kuchuk, H., Kovalenko, A., Ibrahim, B.F. and Ruban, I. (2019), "Adaptive compression method for video information", *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8(1), pp. 66-69, doi: http://dx.doi.org/10.30534/ijatcse/2019/1181.22019
5. Mozhaev, O., Kuchuk, H., Kuchuk, N., Mykhailo, M. and Lohvynenko, M. (2017), "Multiservice network security metric", *2nd International Conference on Advanced Information and Communication Technologies*, AICT 2017 – Proceedings, pp. 133–136, doi: https://doi.org/10.1109/AIACT.2017.8020083
6. Swanson, M. Zhu, B. and Tewfik, A. (1998), "Transparent Robust Image Watermarking", *Proceedings of 3rd IEEE International Conference on Image Processing*, Lausanne, Switzerland, doi: https://doi.org/10.1109/ICIP.1996.560421
7. Langelaar, G. C., Setyawan, I. and Lagendijk, R. L. (2000), "Watermarking digital image and video data. A state-of-the-art overview", *IEEE Signal Processing Magazine*, vol. 17, no. 5, pp. 20–46, doi: https://doi.org/10.1109/79.879337

Received (Надійшла) 10.06.2024
Accepted for publication (Прийнята до друку) 07.08.2024

Метод вбудовування та вилучення водяних знаків у відеокадри за допомогою дискретного косинусного перетворення

М. Чжоу, Н. Кучук, Д. Гриньов

Анотація. У цьому документі представлено надійний метод вбудовування та вилучення водяних знаків у відеокадри за допомогою дискретного косинусного перетворення (DCT). Цей метод забезпечує високу непомітність і зберігає якість відео при вбудовуванні водяного знаку в компонент яскравості кожного кадру. Масштабні експерименти демонструють стійкість методу до звичайних операцій обробки відео, включаючи стиснення, додавання шуму та кадрування. Запропонований підхід є обчислювально ефективним, що робить його придатним для програм реального часу. Майбутня робота спрямована на оптимізацію алгоритму для швидшої роботи та підвищення надійності проти більш складних атак. Практичні застосування включають керування цифровими правами, перевірку вмісту та криміналістичний аналіз.

Ключові слова: відео, дискрет, перетворення, цифровий водяний знак, алгоритм вбудовування.