Using QRCode to Enhance Extraction Efficiency of Video Watermark Algorithm

Zhongze Lv
Institute of Automation, Chinese Academy of Sciences
University of Chinese Academy of Sciences
Beijing, China
lvzhongze2017@ia.ac.cn

Hu Guan
Beijing Engineering Research Center of Digital Content Technology
Beijing, China
hu.guan@ia.ac.cn

Ying Huang
Institute of Automation, Chinese Academy of Sciences
Beijing, China
ying.huang@ia.ac.cn

Shuwu Zhang
Beijing Engineering Research Center of Digital Content Technology
Beijing, China
shuwu.zhang@ia.ac.cn

***Abstract*—Video watermarking can effectively protect the copyright of video contents, but how to improve efficiency of watermarking algorithms is an urgent problem to be solved. In this paper, QRCode is embedded in the scene change frames of videos based on the advantage of QRCode's strong fault tolerance. Combined with the characteristic of high decoding reliability of QRCode, a strategy to terminate the extraction process in advance is proposed to improve the extraction efficiency of the watermark algorithm. Experimental results show that the proposed algorithm has higher extraction efficiency than the algorithm that directly uses character string as watermark.**

***Keywords*—video watermark; QRCode; extraction efficiency**

I. Introduction

With the development of Internet, the spread of digital media contents has become more and more convenient and easy, plagiarism and piracy have also become more and more common, and the rights of authors have been seriously damaged. As a algorithm of embedding specific copyright information into digital media contents, digital watermarking algorithm can effectively protect the copyright of digital media contents, so the research on digital watermarking algorithm is very meaningful. This paper mainly focuses on video watermark algorithm.

According to the different embedding locations, video watermarking algorithm can be divided into compressed-domain video watermarking algorithm and original-domain video watermarking algorithm [1], [2]. The compressed-domain video watermarking algorithm refers to an algorithm that embeds and extracts the watermark in the compressed video stream or partially decoded compressed video. The original-domain video watermarking algorithm, also known as uncompressed-domain video watermarking algorithm, refers to an algorithm that directly embeds and extracts the watermark in uncompressed video data. The advantages of original-domain video watermarking algorithm over the compressed-domain video watermarking algorithm are that it is not affected by the video format and it can learn from the design ideas of the image watermarking algorithm, but it may also has the disadvantage of low execution efficiency.

Extraction efficiency is an important evaluation index of video watermarking algorithm and the main focus of this paper is on how to enhance the extraction efficiency of the original-domain video watermarking algorithm. If the extracting efficiency of a video watermarking algorithm is higher, then the watermark can be obtained in a shorter time, so that the algorithm is more suitable for practical application. Therefore, it is necessary to design a video watermarking algorithm with high extraction efficiency.

Selecting appropriate video frames for watermark embedding and extracting can effectively improve the invisibility and efficiency of video watermarking algorithm. It is a good method to embed the watermark in the scene change frames [3]–[5] of videos. Scene change frame is the first frame of each scene in a video. Scene change frames change very quickly when the video is playing, so embedding the watermark into them can ensure the invisibility of the algorithm. Meanwhile, this method can locate the frames containing the watermark with greater probability during extracting, which can effectively improve the extraction efficiency of the algorithm.

QRCode [6] is a common two-dimensional code, which is also a popular coding method on mobile devices in recent years. QRCode has strong fault tolerance and high decoding reliability, so its application in the watermarking field can effectively improve the robustness and efficiency of watermarking algorithm.

In this paper, we encode text watermark information into QRCode and embed it into videos' scene change frames and in the frame by frame extraction process of watermark algorithm, according to the characteristic of high decoding reliability of QRCode, a strategy to decide when to terminate the extraction process in advance according to the success of the decoder decoding is proposed, which improves the extraction efficiency of the algorithm.

The following contents of this paper is arranged as follows:

978-1-7281-8138-7/20/$31.00 ©2020 IEEE
DOI 10.1109/ICCST50977.2020.00071

Section II mainly introduces the related work. Section III describes the preliminaries of this paper. Section IV discusses the proposed method. Section V presents the experimental results. Section VI concludes.

II. Related Work

Many video watermarking algorithms [7]–[9] embed string as watermark in videos. When extracting, they either need to extract frame by frame and then statistically analyze the final string, or need to know the original string information for comparison. The former has very low extraction efficiency, the latter has the disadvantage that the original watermark information cannot be obtained in many cases.

Although some works [10], [11] embed QRCode as watermark into videos, they do not optimize the watermark extraction process according to the high decoding reliability of QRCode. M.Ketcham et. [10] embed QRCode in the discrete multiwavelet transform domain of all frames of videos, and then extract the watermark from the whole frames. The efficiency of this method is very low. Prabakaran et al. [11] embed QRCode in the composite domain of SVD and DWT of videos' I frames. This method needs to use the original video when extracting, so that it is not a blind watermarking algorithm.

In this paper, QRCode is embedded in the scene change frames of videos. When the watermark is extracted frame by frame, if the QRCode decoder successfully decodes the information extracted from the current frame, then the entire extraction process is terminated and the next frame will not continue to extract.

III. Preliminaries

Some basic concepts of QRCode and Discrete Cosine Transform are given in this section.

A. QRCode

QRCodes can be divided into 40 versions according to different sizes. The size of version 1 is 21*21, and the size of version 40 is 177*177. Each higher version has 4 additional modules on each side.

QRCode is divided into L-level, M-level, Q-level and H-level according to the error-correcting ability. Each level's approximate amount of correction is shown in table I. In this paper, H-level QRCode is selected as the watermark to achieve maximum error-correcting ability.

TABLE I
Error-Correction Level of QRCode

Error-Correction Level	Approximate Amount of Correction
L	7%
M	15%
Q	25%
H	30%

B. Discrete Cosine Transform

For an image $f(x,y)$ with size $M \times N$, the Discrete Cosine Transform(DCT) is defined as (1):

$$F(u,v) = c(u)c(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cdot cos[\frac{(x+0.5)\pi}{M}u] cos[\frac{(y+0.5)\pi}{N}v] \tag{1}$$

where $c(u) = \begin{cases} \sqrt{\frac{1}{M}}, u = 0 \\ \sqrt{\frac{2}{M}}, u \neq 0 \end{cases}$ and $c(v) = \begin{cases} \sqrt{\frac{1}{N}}, v = 0 \\ \sqrt{\frac{2}{N}}, v \neq 0 \end{cases}$.

When $u = 0$ and $v = 0$, $F(0,0)$ is called DC coefficient which is defined as (2):

$$F(0,0) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \tag{2}$$

Meanwhile, the Inverse Discrete Cosine Transform(IDCT) is defined as (3):

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} c(u)c(v)F(u,v) \cdot cos[\frac{(x+0.5)\pi}{M}u] cos[\frac{(y+0.5)\pi}{N}v] \tag{3}$$

In this paper, QRCode is embeded by modifying the DC coefficients obtained from the block-DCT of the scene change frames' Y component.

IV. Proposed Method

A. Choosing Scene Change Frames

We calculates the correlation coefficient between the histogram of the current frame's Y component and the histogram of the previous frame's Y component to judge whether the current frame is a scene change frame or not. The histogram of the previous frame's Y component is denoted as H_0 and the histogram of the current frame's Y component is denoted as H_1. The correlation coefficient denoted as $d(H_0, H_1)$ is calculated as (4):

$$d(H_0, H_1) = \frac{cov(H_0, H_1)}{\sqrt{D(H_0)D(H_1)}} \tag{4}$$

where $cov(H_0, H_1)$ is the covariance of H_0 and H_1, $D(H_0)$ is the variance of H_0 and $D(H_1)$ is the variance of H_1. If $d(H_0, H_1)$ exceeds a threshold T, then H_1 is regarded as a scene change frame. T is an empirical value. In order to ensure that more frames that may contain watermark are located during extracting, the value of T during extracting is usually set higher than that during embedding.

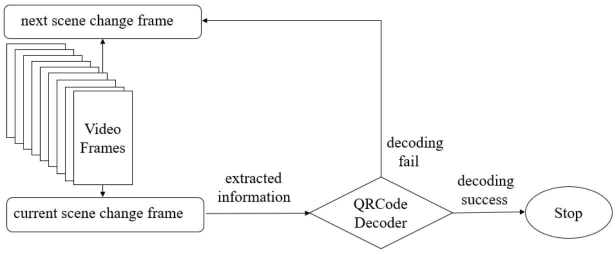

Fig. 1. Process of terminating the extraction process in advance

B. Terminating the Extraction Process in Advance

QRCode has the characteristic of high decoding reliability, and the decoding error rate of each QRCode does not exceed 0.0000001. That means if the QRCode decoder decodes the information successfully, the probability of error is close to zero. According to the advantage mentioned above of QRCode, when extracting the watermark frame by frame, if the information obtained in the current frame can be successfully decoded by the QRCode decoder, then we can directly treat the decoded information of the current frame as the watermark information of the whole video, and terminate the extraction process in advance instead of continuing the extraction of the remaining frames. The process is shown in Fig. 1.

C. Watermark Embedding Process

The flow chart of watermark embedding process is shown in Fig. 2 and the detailed process is described as below:
Step1: Encode the watermark information into QRCode of size W*W, and the QRCode is denoted as E_{ij}, where $i = 0, 1, ..., W-1$ and $j = 0, 1, ..., W-1$.
Step2: According to (4), choose the Y component of scene change frames.

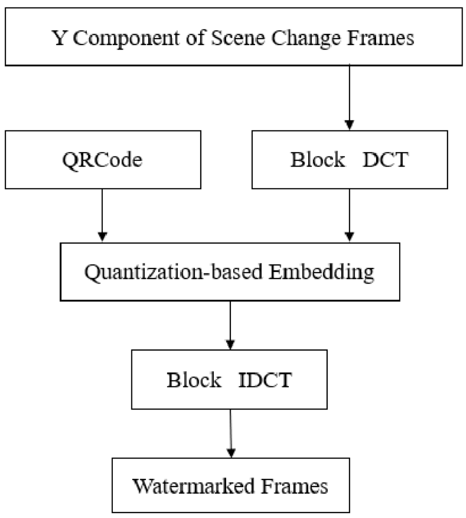

Fig. 2. Process of Watermark Embedding

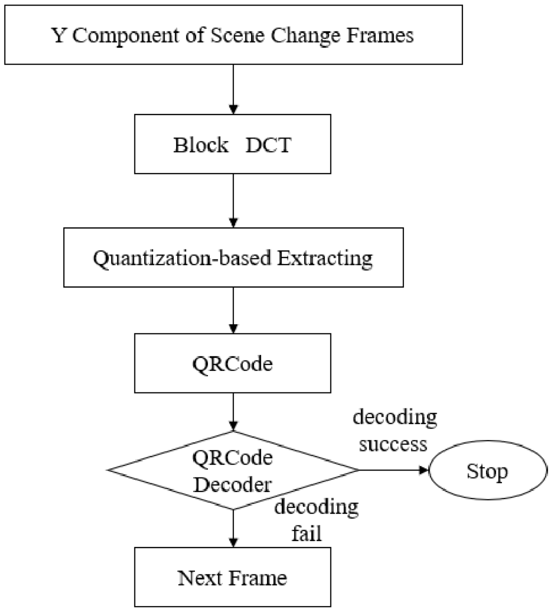

Fig. 3. Process of Watermark Extracting

Step3: Judge whether the width or the height of Y component can be divisible by W or not. If not, enlarge the width or the height with a small scale until it can be divisible by W, and finally get the new Y component denoted as $Y(x, y)$.
Step4: Block $Y(x, y)$, the number of blocks is W*W. DCT is carried out for each block, and the DC coefficient obtained from each block is denoted as DC_{ij}, where $i = 0, 1, ..., W-1$ and $j = 0, 1, ..., W-1$.
Step5: For each DC_{ij}, modify it to embed one bit information of E_{ij} and then get DC'_{ij} according to (5):

$$DC'_{ij} = \begin{cases} (K_{ij} + 0.5)\delta, (K_{ij} + E_{ij})mod2 = 1 \\ (K_{ij} - 0.5)\delta, (K_{ij} + E_{ij})mod2 = 0 \end{cases} \quad (5)$$

where δ refers to the quantization step, $K_{ij} = round(\frac{DC_{ij}}{\delta})$, and $round()$ is a rounding function.
Step6: Replace DC_{ij} with DC'_{ij}, and then perform IDCT on each block to get the Y componenmt containing the watermark.
Step7: Resize the Y component to the original size.

D. Watermark Extracting Process

The flow chart of watermark extracting process is shown in Fig. 3 and the detailed process is described as below:
Step1: According to (4), choose the Y component of scene change frames.
Step2: Judge whether the width or the height of Y component can be divisible by W or not. If not, enlarge the width or the height with a small scale until it can be divisible by W, and finally get the Y component denoted as $Y'(x, y)$.
Step3: Block $Y'(x, y)$, the number of blocks is W*W. DCT is carried out for each block, and the DC coefficient obtained

from each block is denoted as DC_{ij}^*, where $i = 0, 1, ..., W-1$ and $j = 0, 1, ..., W-1$.

Step4: For each DC_{ij}^*, extract one bit watermarking information according to (6):

$$E_{ij}^{'} = \begin{cases} 0, \lfloor \frac{DC_{ij}^*}{\delta} \rfloor mod2 = 1 \\ 1, \lfloor \frac{DC_{ij}^*}{\delta} \rfloor mod2 = 0 \end{cases} \quad (6)$$

Step5: Use the QRCode decoder to decode the information of $E_{ij}^{'}$.

V. Experimental Results and discussion

A. Experemental Setup

In this paper, we use QRCode and two-dimensional 0-1 string as watermarks to embed in the videos, respectively. The two cases are denoted as QRCode-based and String-based. The former uses the extraction termination strategy proposed in this paper when extracting and the latter uses the statistical analysis strategy when extracting. For the sake of fair comparison, the QRCode has the same content and size as the string to ensure that invisibility and robustness of the two cases are consistent. They are just different in format, that means QRCode can be decoded by QRCode decoder and the string cannot. The size(W*W) is set to 25*25. The quantization step δ is set to 80. T is set to 0.6 during embedding process and 0.7 during extracting process.

The test set includes 5 videos. Each video has a resolution of 1280*720. All experiments were performed on a PC with 3.4 GHz Intel Core i7 CPU and 16GB RAM, running in 64-bit Windows 10. The software for simulations was Visual Studio 2010 with OpenCV 2.4.9 and FFmpeg 2.1.

B. Comparison of Extraction Efficiency

In this part, we respectively compared the extraction time of QRCode-based and String-based under the conditions of non-attack, gaussian noise, salt & pepper noise, median filtering and average filtering. The extraction time is the average extraction time of the five videos. The shorter the extraction time is, the higher the extraction efficiency is. The results are shown in Table II. We can find that when invisibility and robustness of the two cases are consistent, the extraction time of QRCode-based is shorter than that of String-based because the strategy of terminating the extraction process in advance proposed in this paper can effectively enhance extraction efficiency.

TABLE II
Comparison of Extraction Time

Attack Types	QRCode-based(s)	String-based(s)
Non-attack	0.899	33.183
Gaussian(0.001)	0.889	33.559
Salt & Pepper(0.001)	0.940	33.542
Median Filtering(3×3)	0.953	33.963
Average Filtering(3×3)	0.971	34.107

VI. Conclusion

In this paper, we proposed a strategy to terminate the extraction process of the video watermark algorithm in advance according to the advantage of QRCode's high decoding reliability. Through experimental verification, our proposed strategy can effectively enhance the extraction efficiency of the watermark algorithm. In future work, we will combine the strategy proposed in this paper to design a more robust and efficient watermarking algorithm.

Acknowledgment

This work was supported by the National Key R&D Program of China (2019YFF0302800) and the Key R&D Program of Shanxi (201903D421007). It was also the research achievement of the Key Laboratory of Digital Rights Services, which is one of the National Science and Standardization Key Labs for Press and Publication Industry.

References

[1] X. Yu, C. Wang, and X. Zhou, "A survey on robust video watermarking algorithms for copyright protection," *Applied Sciences*, vol. 8, no. 10, p. 1891, 2018.

[2] M. Asikuzzaman and M. R. Pickering, "An overview of digital video watermarking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 9, pp. 2131–2153, 2017.

[3] M. Del Fabro and L. Böszörmenyi, "State-of-the-art and future challenges in video scene detection: a survey," *Multimedia systems*, vol. 19, no. 5, pp. 427–454, 2013.

[4] T. R. Singh, K. M. Singh, and S. Roy, "Video watermarking scheme based on visual cryptography and scene change detection," *AEU-International Journal of Electronics and Communications*, vol. 67, no. 8, pp. 645–651, 2013.

[5] B. Reddy and A. Jadhav, "Comparison of scene change detection algorithms for videos," in *2015 Fifth International Conference on Advanced Computing & Communication Technologies*. IEEE, 2015, pp. 84–89.

[6] L. Rosales-Roldan, J. Chao, M. Nakano-Miyatake, and H. Perez-Meana, "Color image ownership protection based on spectral domain watermarking using qr codes and qim," *Multimedia Tools and Applications*, vol. 77, no. 13, pp. 16 031–16 052, 2018.

[7] A. Cedillo-Hernandez, M. Cedillo-Hernandez, M. Garcia-Vazquez, M. Nakano-Miyatake, H. Perez-Meana, and A. Ramirez-Acosta, "Transcoding resilient video watermarking scheme based on spatio-temporal hvs and dct," *Signal Processing*, vol. 97, pp. 40–54, 2014.

[8] A. Cedillo-Hernandez, M. Cedillo-Hernandez, M. N. Miyatake, and H. P. Meana, "A spatiotemporal saliency-modulated jnd profile applied to video watermarking," *Journal of Visual Communication and Image Representation*, vol. 52, pp. 106–117, 2018.

[9] H. Zhao, Q. Dai, J. Ren, W. Wei, Y. Xiao, and C. Li, "Robust information hiding in low-resolution videos with quantization index modulation in dct-cs domain," *Multimedia Tools and Applications*, vol. 77, no. 14, pp. 18 827–18 847, 2018.

[10] M. Ketcham and T. Ganokratanaa, "The evolutionary computation video watermarking using quick response code based on discrete multiwavelet transformation," in *Recent Advances in Information and Communication Technology*. Springer, 2014, pp. 113–123.

[11] G. Prabakaran, R. Bhavani, and M. Ramesh, "A robust qr-code video watermarking scheme based on svd and dwt composite domain," in *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering*. IEEE, 2013, pp. 251–257.