A novel two-dimensional reversible data hiding method with high embedding capacity in H.264/advanced video coding

Yi Chen\textsuperscript{1}*, Hongxia Wang\textsuperscript{2}, Xiaoxu Tang\textsuperscript{1}, Yong Liu\textsuperscript{3}, Hanzhou Wu\textsuperscript{3} and Yanli Chen\textsuperscript{4}

Abstract
Developing the technology of reversible data hiding based on video compression standard, such as H.264/advanced video coding, has attracted increasing attention from researchers. Because it can be applied in some applications, such as error concealment and privacy protection. This has motivated us to propose a novel two-dimensional reversible data hiding method with high embedding capacity in this article. In this method, all selected quantized discrete cosine transform coefficients are first paired two by two. And then, each zero coefficient-pair can embed 3 information bits and the coefficient-pairs only containing one zero coefficient can embed 1 information bit. In addition, only one coefficient of each one of the rest coefficient-pairs needs to be changed for reversibility. Therefore, the proposed two-dimensional reversible data hiding method can obtain high embedding capacity when compared with the related work. Moreover, the proposed method leads to less degradation in terms of peak-signal-to-noise ratio, structural similarity index, and less impact on bit-rate increase.

Keywords
Two-dimensional reversible data hiding, embedding capacity, quantized discrete cosine transform coefficients, H.264/advanced video coding, privacy protection

Date received: 8 October 2019; accepted: 10 February 2020

Handling Editor: Yulei Wu

Introduction
With the introduction of sensor networks, many smart devices are used to collect a large amount of data, including images, videos, speeches, and texts, for smart homes, health monitoring, traffic control, and so on. This indeed makes people’s life more convenient. However, this may lead to the leakage of personal information at the same time. For example, the face information and fingerprint information in public videos are abused. Therefore, it is important to guarantee the key content of video but not make the personal information leakage. Nowadays, distributing digital videos to the global has become easier because of the rapid development of high-speed broadband Internet and video encoding standard. For instance, in practice,
many social applications, such as Skype, Facebook, WhatsApp, WeChat, Blog, and QQ, can be used to spread digital videos. This has brought many concerns and maybe lead to many problems, even criminal activities, such as the illegal distribution of a digital movies and the leakage of personal privacy information in public videos. Therefore, researching and finding out effective ways to solve these problems or prevent them happening has become necessary.

Currently, encryption and digital watermarking are two commonly used techniques in digital multimedia, such as images and videos, to these problems. When applying encryption technique in images or videos (also referred to as motion images), the computational complexity may be high. In addition, it is possible for a video codec to produce format incompatibility. At the same time, encryption usually makes the content of images or videos unavailable. However, in some cases, like copyright protection, the content should be available. Hence, there have been many researchers to start to research digital watermarking in videos not only for solving these problem but also for other purposes, such as broadcast monitoring, copy or playback control, online location, and content filtering. Digital watermarking is a part of data hiding (DH), and it is classified into irreversible and reversible watermarking corresponding to the technologies of irreversible DH and reversible data hiding (RDH), respectively. Compared with irreversible data embedding, RDH has attracted much more attention from many researchers because it can embed additional information into digital media, such as images and videos, and recover the original media content after extracting the embedded information from the marked digital media.

In the past two decades, RDH in images has been rapidly developed that leads to make achievements. For instance, Wu et al. designed an RDH scheme in encrypted palette images since palette images are widely utilized in real life. In their RDH scheme, a color partitioning method is proposed to make use of the palette colors to construct a certain number of embeddable color triples for embedding the secret data. By doing this, their scheme can provide a relatively high data-embedding payload and have a low computational complexity. Recently, Yang et al. propose an adaptive real-time RDH for JPEG images. This RDH scheme is realized by using successive zero coefficients in zig-zag order of discrete cosine transform (DCT) blocks. Their experimental results have verified that their proposed scheme can enhance embedding capacity meanwhile maintaining the image quality. Moreover, Chen and Wang proposed a RDH method with high embedding payload for JPEG images. For this method, each quantized discrete cosine transform (QDCT) coefficients is changed for carrying 1 information bit, thus leading to high embedding payloads.

In videos, there exist many kinds of coding parameters that can be changed for RDH, even DH. Therefore, compared with images, videos have much more research room to develop the techniques of RDH and DH. Recently, as the development of video compression standard, such as H.264/advanced video coding (AVC) and high-efficiency video coding (HEVC), some video DH methods and video RDH methods are reported. For these DH methods, they are proposed for improving embedding capacity, stopping intra-frame drift, and reducing bit-rate increase. For these RDH methods, they are proposed for improving error concealments performance, making embedding capacity larger, protecting privacy, and preventing inter-frame distortion drift. However, compared with the development of RDH technique in images, it is not enough. Thus, this has motivated us to continue to research RDH technique in video. Recently, Xu and Wang proposed a two-dimensional (2D) RDH method, as shown in Figure 1, for error concealment of intra-frame in videos. Compared with one-dimensional RDH method, 2D RDH can keep better performance in terms of peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM). Thus, Xu et al.’s scheme is better than Chung et al.’s scheme. However, Xu et al. did not make full use of and only a part of coefficient-pairs containing zero coefficients are used to map for DH. Based on this, this has motivated us to propose a novel 2D RDH method in this article for improving the embedding capacity. In our experiments, we exploit the method of block selecting in Chen et al.’ method and apply the 2D RDH of Xu and Wang’s method and our proposed 2D RDH method in H.264/AVC reference software JM12.0. In other words, we compare them in identical cases. Experimental results have verified that our proposed method outperforms Xu and Wang’s
method in terms of embedding capacity. Furthermore, our proposed 2D RDH method causes little degradation in visual quality and little impact on coding efficiency in terms of bit-rate increase.

The remainder of this article is organized as follows. In section “proposed method,” we present the proposed 2D RDH method. Some experimental results and analysis are given in section “Experimental results and analysis.” Finally, we draw some conclusions in section “Conclusion”.

**Proposed method**

In common RDH methods, zero QDCT coefficients are not considered and exploited to embed information in compressed images and videos. Therefore, Chen et al. present a video RDH scheme by combining with zero QDCT coefficient-pairs from high-frequency areas. Based on Chen et al.’s work, we propose a novel 2D RDH method in H.264/AVC videos in the following. The histogram modification is shown in Figure 2.

**Data embedding**

During the procedure of data embedding, our proposed 2D RDH method is based on paired QDCT coefficients and thus all coefficients should first be paired two by two. In the following, all coefficient-pairs, each of which is also called as a point, that is, $A(x,y)$ denoted as $(x,y)$ in short in this article, constitute a set denoted as $A$ and then the points in $A$ will be changed for data embedding. In the embedding procedure, $b$ is binary string. For each coefficient-pair, it is changed as follows.

**Shifting**

1. If $y \neq 0$, the point $A(x,y)$ is changed by

   $$(x',y') = \begin{cases} (x,y+1), & \text{if } y > 0 \\ (x,y-1), & \text{if } y < 0 \end{cases}$$  \hspace{1cm} (1)

2. If $A(x,y) = (-1,0)$, that is, $x = -1$ and $y = 0$, the point $A(x,y)$ is changed by

   $$(x',y') = (x,y-1) = (-1,-1)$$  \hspace{1cm} (2)

3. If $A(x,y) = (1,0)$, the point $A(x,y)$ is changed as follows

   $$(x',y') = (x+1,y) = (2,0)$$  \hspace{1cm} (3)

**Data extraction and video recovery**

Corresponding to the procedure of data embedding, the data extraction and the video recovery are addressed as follows.

**Data extraction**

1. For one point $A(x',y')$, if $x' < -1$ or $x' > 2$ and at the same time the absolute value of $y'$ is not greater than 1, the embedded information $b$ is extracted by

   $$b = \begin{cases} 1, & \text{if } y' = -1 \\ 00, & \text{if } y' = 0 \\ 01, & \text{if } y' = 1 \end{cases}$$  \hspace{1cm} (7)
By using equations (10)–(15), the original compressed videos can be restored.

### Video recovery

1. For one point $A(x', y')$, if $x' < -2$ and at the same time the absolute value of $y'$ is not greater than 1, this point is restored by

   $$(x, y) = \begin{cases} 
   (x' + 1, y') & \text{if } y' = 0 \\
   (x', 0) & \text{if } y' \neq 0
   \end{cases}$$

2. For one point $A(x', y')$, if $x' > 1$ and at the same time the absolute value of $y'$ is not greater than 1, this point is restored by

   $$(x, y) = \begin{cases} 
   (x' - 1, y') & \text{if } y' = 0 \\
   (x', 0) & \text{if } y' \neq 0
   \end{cases}$$

3. For one point $A'(x', y')$, if the sum of the absolute values of $x'$ and $y'$ is greater than 2 and $x'$ and $y'$ are not $-1$ at the same time, this point is restored by

   $$(x, y) = (0, 0)$$

4. For one point $A'(x', y')$, if the absolute value of $y'$ is greater than 1, this point is restored by

   $$(x, y) = \begin{cases} 
   (x', y' - 1) & \text{if } y' > 1 \\
   (x', y' + 1) & \text{if } y' < -1
   \end{cases}$$

5. If $A'(x', y') \in \{(2,0)\}$, this point is restored by

   $$(x, y) = (1, 0)$$

6. If $A'(x', y') = (1, -1)$, this point is restored by

   $$A(x, y) = (-1, 0)$$

By using equations (10)–(15), the original compressed videos can be restored.

### Analysis of embedding capacity and distortion

To analyze the embedding capacity and distortion of our proposed method, we first define three sets as follows

- $A_1 = \{(0, 0)\}$
- $A_2 = \{(x, y)|x < -1, y = 0\} \cup \{(x, y)|x > 1, y = 0\}$
- $A_3 = \{(x, y)|x \in Z, y > 0\} \cup \{(x, y)|x \in Z, y < 0\}$
  \[ \cup \{(1, 0)\} \cup \{(-1, 0)\} \]

where $A_1$ and $A_2$ are used for data embedding and $A_3$ is shifted for reversibility. Therefore, the embedding capacity is calculated by

$$EC = k_{d1} \times h(A_1) + k_{d2} \times h(A_2)$$

where $k_{d1}$ and $k_{d2}$ denote average embedding rate per coefficient-pair corresponding to the sets $A_1$ and $A_2$, respectively, and $h(A)$ is defined by

$$h(A) = \#\{A_{set}|A_{set} = A\}$$

where $\#$ is the cardinal number of a set and $A$ denotes all points in the set $A$. Thus, equation (16) can be rewritten as

$$EC = 3 \times h(A_1) + \frac{3}{2} \times h(A_2)$$

In our experiments, we count $h(A_1) = 2375$, $h(A_2) = 149$, and $h(A_3) = 1636$ by setting quantization parameter ($QP$) = 26 and using H.264/AVC encoder to encode Foreman. Therefore

$$EC = 3 \times 2375 + \frac{3}{2} \times 149 = 7348.5$$

It is very close to the result shown in Table 1, that is, 7343 bits. Moreover, embedding distortion can be defined by

$$ED = d_{d1} \times h(A_1) + d_{d2} \times h(A_2) + d_{d3} \times h(A_3)$$

where $d_{d1}$, $d_{d2}$, and $d_{d3}$ denote average modification rate per coefficient-pair corresponding to the sets $A_1$, $A_2$, and $A_3$, respectively. Thereby, equation (19) can be rewritten as

$$ED = \frac{5}{4} \times h(A_1) + h(A_2) + h(A_3)$$

Furthermore

$$ED = \frac{5}{4} \times 2375 + 149 + 1636 = 4753.75$$

In fact, the embedding distortion cannot be measured by equation (19) since H.264/AVC has intra-frame and inter-frame predictions. Equation (19) can
stand for total number of modification on QDCT coefficients but not embedding distortion. Finding a good way to reasonably calculate the embedding distortion is a big challenge and it is also a research direction for us in the future. In this article, we will not address more details about how to find a good way to reasonably calculate embedding distortion.

Experimental results and analysis

This section contains four subsections, that is, setup, embedding capacity, visual quality, and bit-rate variation.

Setup

To evaluate the performance of the proposed 2D RDH method, we applied the proposed 2D RDH method in the H.264/AVC reference software JM12.0.26 Twelve standard video sequences, that is, Akiyo, Claire, Coastguard, Container, Foreman, Miss America, Mobile, Mother–Daughter, News, and Suzie (as shown in Figure 3) downloaded from websites,27 which are with the resolution of 176 × 144, are used in our experiments. Moreover, we give some main configuration parameters of JM12.0 in Table 2, and other parameters not mentioned remain in their default values. The group of picture (GOP) is IBPBPBPBPBPBPBPB. To fairly compare the performance of the proposed 2D RDH method with related method proposed by Xu and Wang,18 we take advantage of the block selecting method proposed by Chen et al.11 to select blocks for data embedding. In other words, the two methods are compared in the identical conditions.

In addition, we exploit embedding capacity, visual quality, and bit-rate variation to measure the performance of our proposed 2D RDH method. PSNR and SSIM28 are used for objectively evaluating visual quality of marked videos. Bit-rate comparisons show the impact of our proposed 2D RDH method on H.264/AVC encoder in terms of coding efficiency. In the following several subsections, the “Original” of PSNR, SSIM, and bit-rate is computed by the original H.264/AVC encoder. Otherwise, they are computed by H.264/AVC encoder with the corresponding DH methods. Otherwise, they are computed by H.264/AVC encoder with the corresponding DH methods. More analyses are given as follows.

Embedding capacity

Table 1 shows the maximum embedding capacities on these 12 video sequences mentioned in section “Setup” by using Xu and Wang’s method18 and our proposed method. In Table 1, QP has three values, that is, 24, 26, and 28 and it determines the quantization step of H.264/AVC encoder.9 According to Table 1, obviously, our proposed method has larger maximum embedding capacities on these 12 video sequences when compared with Xu and Wang’s method.18 For example, on Miss America in Table 1, our proposed method obtains 3204, 2628, and 2543 bits corresponding to QP = 24, 26, and 28, respectively. However, Xu and Wang’s method obtained 2086, 1720, and 1664 bits correspondingly. Moreover, our proposed method obtains average maximum embedding capacities of 6803, 6450, and 3598 bits, which are greater than that Xu and Wang’s method obtains, that is, 4339, 4126, and 2348 bits. These have verified that our proposed 2D RDH method has indeed an advantage in embedding capacity when compared with Xu and Wang’s method.18

| Sequences       | No. | QP = 24 |         | QP = 26 |         | QP = 28 |         |
|-----------------|-----|---------|---------|---------|---------|---------|---------|
|                 | Frame | Xu and Wang18 | Proposed | Xu and Wang18 | Proposed | Xu and Wang18 | Proposed |
| Akiyo           | 300  | 6022    | 9263    | 5226    | 8088    | 4108    | 6291    |
| Carphone        | 382  | 6018    | 9309    | 5296    | 8153    | 5502    | 8422    |
| Claire          | 494  | 3894    | 6043    | 2934    | 4587    | 2412    | 4685    |
| Coastguard      | 300  | 5792    | 9158    | 5780    | 9060    | 5969    | 8831    |
| Container       | 300  | 3046    | 4932    | 3430    | 5448    | 2616    | 4156    |
| Foreman         | 300  | 4510    | 7015    | 4750    | 7343    | 4834    | 7461    |
| Hall Monitor    | 300  | 5466    | 8457    | 5432    | 8289    | 5398    | 8254    |
| Miss America    | 150  | 2086    | 3204    | 1720    | 2628    | 1664    | 2543    |
| Mobile          | 300  | 2908    | 5097    | 3234    | 5561    | 3982    | 6586    |
| Mother–Daughter | 300  | 4688    | 7227    | 4778    | 7345    | 4352    | 6634    |
| News            | 300  | 4862    | 7705    | 4370    | 6955    | 4964    | 7768    |
| Suzie           | 150  | 2786    | 4235    | 2570    | 3946    | 2348    | 3598    |
| Average         | –    | 4339    | 6803    | 4126    | 6450    | 3989    | 6269    |

QP: quantization parameter.

Chen et al.
In this subsection, we will measure the visual quality of marked videos by our proposed method in two sides. On one hand, we give Figures 4 and 5 to subjectively evaluate the visual quality of marked videos. Herein, Figures 4 and 5 represent videos with more smooth areas and more rich areas, respectively. At the same time, Figures 4 and 5 are the 31st frame of Hall Monitor and Mobile, respectively. Figures 4(a)–(c) and 5(a)–(c) are obtained by using the original H.264/AVC encoder. Figures 4(d)–(f) and 5(d)–(f) are generated by H.264/AVC encoder with Xu and Wang’s method\(^\text{18}\) and similarly Figures 4(g)–(i) and 5(g)–(i) are generated by H.264/AVC encoder with our proposed method. According to Figures 4 and 5, we do not observe any distortion of video frame caused by Xu and Wang’s method\(^\text{18}\) and our proposed method when compared with the original video frames. Under this case, our proposed method outperforms Xu and Wang’s method\(^\text{18}\) in terms of embedding capacity.

However, we make use of PSNR and SSIM to objectively evaluate visual quality of marked videos and they are shown in Tables 3 and 4. In Table 3, when \(QP\) is fixed, “Original” PSNR value is greater than PSNR values caused by Xu and Wang’s method\(^\text{18}\) and our proposed method. In addition, Xu and Wang’s method\(^\text{18}\) provides larger PSNR values than our proposed method. For instance, exploiting H.264/AVC encoder with \(QP = 24\) on Mobile, they provide 37.83, 37.55, and 37.47 dB. Average PSNR values also meet this. Noted that, however, Tables 3–5 correspond to Table 1. Therefore, it is not fair to compare the two 2D RDH methods like this. To better compare the two methods, we define PSNR variation (PSNRV) by

\[
PSNRV = \frac{PSNR_{\text{Ori}} - PSNR_{\text{2RDH}}}{EC} \times 100\%
\]

where \(PSNR_{\text{Ori}}\) and \(PSNR_{\text{2RDH}}\) are determined by the H.264/AVC encoder without and with 2D RDH method, respectively. \(EC\) denotes embedding capacity. \(PSNRV\) is used to evaluate embedding distortion. Based on this, we give Figure 6 corresponding to \(QP = 28\). As shown in Figure 6, for these 12 video sequences, the proposed method leads to very close embedding distortion when compared with Xu and Wang’s method\(^\text{18}\). In

![Test video sequences: (a) Akiyo, (b) Carphone, (c) Clair, (d) Coastguard, (e) Container, (f) Foreman, (g) Hall Monitor, (h) Miss America, (i) Mobile, (j) Mother–Daughter, (k) News, and (l) Suzie.](image)
particular, although our proposed method provides greater embedding distortion on Suzie than Xu and Wang’s method,\textsuperscript{18} it provides less embedding distortion on other 11 video sequences. In other words, our proposed method has less impact caused by DH.

| Parameter                  | Configuration |
|----------------------------|---------------|
| Profile                    | Main          |
| Frame rate                 | 30            |
| Rate distortion optimization| On            |
| IntraPeriod                | 8             |
| Symbol mode                | 0: CAVLC      |
| FrameSkip                  | 1             |
| NumberBFrame               | 1             |

CAVLC: context-adaptive variable length coding.

Table 2. Configuration parameters of the JM12.0 software.

In Table 4, for “Original,” Xu and Wang’s method,\textsuperscript{18} and our proposed method, the SSIM values are decreasing with the increase in the $Q_P$ value. Although our proposed method provides least SSIM value of 0.9404 on Coastguard when $Q_P = 28$, the difference between them are little. When considering the embedding capacity (shown in Table 1), the performance of our proposed method is accepted. Totally, our proposed method provides larger embedding capacity and leads to close embedding distortion when compared with the related work.

**Bit-rate variation**

Bit-rate after and before embedding data into video sequences is often used to evaluate the coding efficiency of H.264/AVC codec. In this subsection, we give Table 5 to

![Figure 4. The 31st frame of Hall Monitor (with more smooth areas): (a)–(c) original frames, (d)–(f) marked frames by Xu and Wang’s method,\textsuperscript{18} and (g)–(i) marked frames by the proposed two-dimensional reversible data hiding method. For (a), (d), and (g), $Q_P = 24$. For (b), (e), and (h), $Q_P = 26$. For (c), (f), and (i), $Q_P = 28$.](image-url)
Figure 5. The 31st frame of Mobile (with more rich areas): (a)–(c) original frames, (d)–(f) marked frames by Xu and Wang's method\textsuperscript{18} and (g)–(i) marked frames by the proposed two-dimensional reversible data hiding method. For (a), (d), and (g), $QP = 24$. For (b), (e), and (h), $QP = 26$. For (c), (f), and (i), $QP = 28$.

Table 3. Comparisons of PSNR (dB) between Original, Xu and Wang's method\textsuperscript{18} and the proposed method.

| Sequences     | No. Frames | $QP = 24$ | $QP = 26$ | $QP = 28$ |
|---------------|------------|-----------|-----------|-----------|
|               | Original   | Xu and Wang\textsuperscript{18} | Proposed | Original   | Xu and Wang\textsuperscript{18} | Proposed | Original   | Xu and Wang\textsuperscript{18} | Proposed |
| Akiyo         | 300        | 42.29     | 41.59     | 41.25     | 40.82     | 40.27     | 39.97     | 39.52     | 38.91     |
| Carphone      | 382        | 40.73     | 40.44     | 40.34     | 39.18     | 38.98     | 38.81     | 37.71     | 37.51     |
| Claire        | 494        | 43.83     | 43.52     | 43.31     | 42.47     | 42.25     | 42.07     | 41.16     | 40.88     |
| Coastguard    | 300        | 38.47     | 38.14     | 38.02     | 36.81     | 36.50     | 36.29     | 35.30     | 34.91     |
| Container     | 300        | 39.83     | 39.45     | 39.27     | 38.44     | 38.03     | 37.84     | 37.18     | 36.86     |
| Foreman       | 300        | 39.81     | 39.59     | 39.34     | 38.32     | 38.14     | 37.77     | 36.93     | 36.68     |
| Hall Monitor  | 300        | 40.74     | 40.21     | 40.01     | 39.53     | 38.95     | 38.71     | 38.33     | 37.75     |
| Miss America  | 150        | 43.30     | 43.00     | 42.83     | 42.15     | 41.89     | 41.70     | 41.02     | 40.69     |
| Mobile        | 300        | 37.83     | 37.55     | 37.47     | 36.00     | 35.74     | 35.64     | 34.32     | 34.06     |
| Mother–Daughter | 300       | 41.35     | 40.93     | 40.76     | 39.86     | 39.57     | 39.28     | 38.48     | 38.21     |
| News          | 300        | 41.11     | 40.56     | 40.37     | 39.52     | 38.95     | 38.72     | 38.06     | 37.49     |
| Suzie         | 300        | 40.63     | 40.42     | 40.27     | 39.23     | 39.01     | 38.79     | 37.96     | 37.79     |
| Average       | –          | 40.83     | 40.45     | 40.27     | 39.36     | 39.02     | 38.80     | 38.00     | 37.65     |

PSNR: peak-signal-to-noise ratio; $QP$: quantization parameter.
show bit-rate variation without and with DH method. Likewise, Table 5 also corresponds to Table 1.

According to Table 5, our proposed method leads to close coding efficiency of H.264/AVC encoder with Xu and Wang’s method. Compared with “Original,” our proposed method has a little impact on coding efficiency. In addition, to better compare our proposed method with Xu and Wang’s method, we define bit-rate variation by

$$BVE = \frac{B_{2DRDH} - B_{Ori}}{EC} \times 100\%$$ (22)

where $B_{2DRDH}$ and $B_{Ori}$ are generated by H.264/AVC encoder with and without the RDH method, respectively. Similarly, based on this, we draw Figure 7.

Herein, Figure 7 corresponds Figure 6 and Table 1. As seen from Figure 7, obviously, except Claire, our proposed method leads to less bit-rate increase per embedding information bit when compared with Xu and Wang’s method. That is to say that our proposed method outperforms Xu and Wang’s method in terms of embedding impact per embedding information bit.

**Conclusion**

This article presents a novel 2D RDH method, which is based on H.264/AVC compression standard, with high embedding capacity. In the proposed 2D RDH method, almost all points, referred to as coefficient-pair in this article, containing coefficient with a value of 0 are used
for data embedding. In addition, only one coefficient in other points is changed to vacate room for reversibility. Therefore, the proposed method provides higher embedding capacity compared with the related method. When compared the variation of PSNR and SSIM caused by each embedding information bit, the proposed method keeps better visual quality. Moreover, the increase in bit-rate caused by the proposed method is less.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61972269, the Fundamental Research Funds for the Central Universities under Grant No. YJ201881, and Doctoral Innovation Fund Program of Southwest Jiaotong University under Grant No. DCX201824.

ORCID iD

Yi Chen https://orcid.org/0000-0003-4272-7956

References

1. Asikuzzaman M and Pickering MR. An overview of digital video watermarking. *IEEE T Circ Syst Vid* 2018; 28(9): 2131–2153.
2. Tew Y and Wong K. An overview of information hiding in H.264/AVC compressed video. *IEEE T Circ Syst Vid* 2014; 24(2): 305–319.
3. Ni Z, Shi YQ, Ansari N, et al. Reversible data hiding. *IEEE T Circ Syst Vid* 2006; 16(3): 354–362.
4. Tian J. Reversible data embedding using a difference expansion. *IEEE T Circ Syst Vid* 2003; 13(8): 890–896.
5. Shi YQ, Li X, Zhang X, et al. Reversible data hiding: advances in the past two decades. *IEEE Access* 2016; 4: 3210–3237.
6. Wu HZ, Shi YQ, Wang HX, et al. Separable reversible data hiding for encrypted palette images with color partitioning and flipping verification. *IEEE T Circ Syst Vid* 2017; 27(8): 1620–1631.
7. Yang CN, Kim C and Lo YH. Adaptive real-time reversible data hiding for JPEG images. *J Real-Time Image Pr* 2018; 14(1): 147–157.
8. Chen Y and Wang HX. An improved reversible data hiding scheme by changing modification direction of partial coefficients in JPEG images (arxiv preprint arxiv:1804.06645v3), 2018, https://arxiv.org/abs/1804.06645
9. Wiegand T, Sullivan GJ, Bjontegaard G, et al. Overview of the H.264/AVC video coding standard. *IEEE T Circ Syst Vid* 2003; 13(7): 560–576.
10. Sullivan GJ, Ohm JR, Ohm WJ, et al. Overview of the high efficiency video coding (HEVC) standard. *IEEE T Circ Syst Vid* 2012; 22(12): 1649–1668.
11. Chen Y, Wang HX, Wu HZ, et al. A data hiding scheme with high quality for H.264/AVC video streams. In: *Proceedings of the 4th international conference on cloud computing and security*, Haikou, China, 8–10 June 2016, pp.99–110. Cham: Springer.
12. Fallahpour M, Shirmohammadi S and Ghanbari M. A high capacity data hiding algorithm for H.264/AVC. *Secur Commun Netw* 2015; 8(16): 2947–2955.
13. Liu Y, Liu S, Zhao HG, et al. A new data hiding method for H.265/HEVC video streams without intra-frame distortion drift. *Multimed Tools Appl* 2019; 78(16): 23097–23115.
14. Xu D, Wang R and Zhu Y. Tunable data hiding in partially encrypted H.264/AVC videos. *J Vis Commun Image R* 2017; 45: 34–45.
15. Kumar P and Singh K. An improved data-hiding approach using skin-tone detection for video steganography. *Multimed Tools Appl* 2018; 77(18): 24234–24268.
16. Ma XJ, Li Z, Tu H, et al. A data hiding algorithm for H.264/AVC video streams without intra-frame distortion drift. *IEEE T Circ Syst Vid* 2010; 20(10): 1320–1330.
17. Chen Y, Wang HX, Wu HZ, et al. Reversible video data hiding using zero QDCT coefficient-pairs. *Multimed Tools Appl* 2019; 78(16): 23097–23115.
18. Xu D and Wang R. Two-dimensional reversible data hiding-based approach for intra-frame error concealment in H.264/AVC. *Signal Process: Image* 2016; 47: 369–379.
19. Yao Y, Zhang W and Yu N. Inter-frame distortion drift analysis for reversible data hiding in encrypted H.264/AVC video bitstreams. *Signal Process* 2016; 128: 531–545.
20. Long M, Peng F and Li H. Separable reversible data hiding and encryption for HEVC video. *J Real-Time Image Pr* 2018; 14(1): 171–182.
21. Niu K, Yang X and Zhang Y. A novel video reversible data hiding algorithm using motion vector for H.264/AVC. *Tsinghua Sci Technol* 2017; 22(5): 489–498.
22. Zhao J, Li ZT and Feng B. A novel two-dimensional histogram modification for reversible data embedding into stereo H.264 video. *Multimed Tools Appl* 2016; 75(10): 5959–5980.
23. Ma XJ, Yang LT, Xiang Y, et al. Fully reversible privacy region protection for cloud video surveillance. *IEEE T Cloud Comp* 2017; 5(2): 510–522.
24. Chen Y, Wang HX, Wu HZ, et al. An adaptive data hiding algorithm with low bitrate growth for H.264/AVC video stream. *Multimed Tools Appl* 2018; 77(15): 20157–20175.
25. Chung KL, Huang YH, Chang PC, et al. Reversible data hiding-based approach for intra-frame error concealment in H.264/AVC. *IEEE T Circ Syst Vid* 2010; 20(11): 1643–1647.
26. H.264/AVC reference software JM12.0 (online referencing), http://iphome.hhi.de/suehring/tml/download/old_jm/ (2003, accessed 12 October 2018).
27. Test video sequences (online referencing), http://trace.eas.asu.edu/yuv/ (accessed 12 October 2018).
28. lie A and klaue J. Evalvid-RA: trace driven simulation of rate adaptive MPEG-4 VBR video. *Multimedia Syst* 2008; 14(1): 33–50.