# Robustness Security of Data Hiding for H.265/HEVC Video Streams

**Grace C.-W. Ting, Bok-Min Goi, Sze-Wei Lee**

1

***Abstract:*** ***In the current age of IoT and Industry 4.0, vast amounts of big data are exchanged in the cloud, social media and among sensing devices, at times causing videos to go viral and proliferating fake news. Ultra high quality videos such as in the latest H.265/HEVC standard format have the capacity for data to be embedded in them for different security applications. Security for high quality videos is a concern in our current technological age where netizens often share videos with others through social media and video hosting sites like YouTube. In this paper, we focus on data hiding for H.265/ HEVC and perform detailed security analysis of a recent H. 265/HEVC data hiding scheme. Our security consideration includes the hiding properties of secrecy and undetectability, as well as the robustness security property of tamper resistance. We show specifically an attack against the robustness security of this data hiding scheme. We discuss reasons causing the insecurity against robustness and justify strategies to improve the scheme's protection against these types of attacks.***

***Index Terms: IoT, cloud computing, H.265/ HEVC, video sharing.***

## I. INTRODUCTION

Different network-enabled devices with sensing capabilities proliferate in our current technological internet of things (IoT) world. Industry 4.0 is also revolutionising how the cyber and physical world engage and exchange vast amounts of data. As sensors improve in perceptual quality and systems advance in processing capability, even videos at ultra high resolutions can be exchanged efficiently across the internet and social networks.

H.265/HEVC [1] is the latest video encoding standard for ultra high quality and supports resolutions up to 8192×4320, which includes the 8K UHD standard. As these types of videos are commonly acquired, stored and shared across the social media and video hosting sites on the internet, they could be used for security purposes [2,3] via techniques such as data hiding [4].

Data hiding basically embeds message bits into the videos in a covert manner such that the video quality is preserved, and with the requirements that the message is kept hidden, and to some extent even the existence of such a message is undetectable. One main security property is also that the embedded message should remain intact without modifications, so that the intended recipient is able to extract the original message, even though the videos have undergone common video processing operations.

Only a few data hiding schemes [5,6,7] have been proposed for H.265/HEVC since this is a new video standard.

In this paper, we focus on a very recent (already accepted but not even assigned to an issue yet) data hiding scheme, proposed by Liu et al. which is published in the Multimedia Tools & Applications journal [8], and perform a detailed security analysis of the scheme with respect to fundamental security requirements namely secrecy, undetectability and robustness. This is the first known security analysis of this scheme, i.e. up to now no attacks exist on it.

We first discuss its security against secrecy and undetectability and then show a robustness attack on the scheme that enables attackers to change the original embedded message in the video without affecting the quality. The paper concludes with specific suggestions on how to improve the scheme to prevent these types of attacks.

## II. DATA HIDING SCHEME FOR H.265/HEVC VIDEOS

### *A. Data Embedding*

The H.265/HEVC video coding standard uses integer transforms based on the Discrete Cosine Transform (DCT) and Discrete Sine Transform (DST). An input m×n video frame is essentially divided into multiple non-overlapping blocks of varying sizes ranging from 4×4 to 32×32. For the 4×4 case as an example, let the input 4×4 block be denoted as X. Then the output 4×4 block Y is computed via the DST as:

$$Y = (H\,X\,H^{T}) \tag{1}$$

where

$$H = \begin{bmatrix} a & b & c & d \\ c & c & 0 & -c \\ d & -a & -c & b \\ b & -d & c & -a \end{bmatrix}, \tag{2}$$

where $a$ = 29, $b$ = 55, $c$ = 74, $d$ = 84, and $H^{T}$ denotes the transpose of $H$.

HEVC works like previous video compression standards in the sense that the compression within a frame, i.e. intra-frame, is achieved by having the values of blocks within a frame to be computed (a.k.a. predicted) from the values of adjacent blocks. In more detail, all the 16 pixels of the 4×4 block are predicted using the pixels at the boundary of the adjacent blocks located either above or to the left of the

**Revised Manuscript Received on August 18, 2019.**

**Grace C.-W. Ting**, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Malaysia.

**Bok-Min Goi**, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Malaysia.

**Sze-Wei Lee**, Tunku Abdul Rahman University College, Malaysia.

This research was funded by the Ministry of Education Fundamental Research Grant Scheme (FRGS) project FRGS/1/2016/ICT04/UTAR/01/1.

current block. Fig. 1 illustrates this.

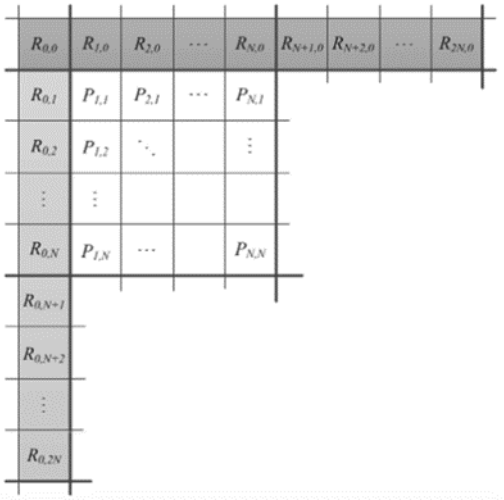


Fig. 1: Prediction of current block from boundary pixels of adjacent blocks above and to the left [8]

Since embedding into a DST coefficient involves changing the value of the DST coefficient, Liu et al. rightly argue that embedding therefore causes distortion over time. This is known as the distortion drift. For instance, recall that the current block's values are predicted from adjacent blocks above or to the left of it, but if the adjacent blocks have values that have been changed from the original due to embedding, then this error propagates as blocks continue to be predicted from adjacent blocks.

The Liu et al. data hiding scheme [8] operates on H.265/HEVC video by embedding the message bits into the 4×4 luminance matrix of the coefficients output from the Discrete Sine Transform (DST). Refer to Fig. 2 for illustration.

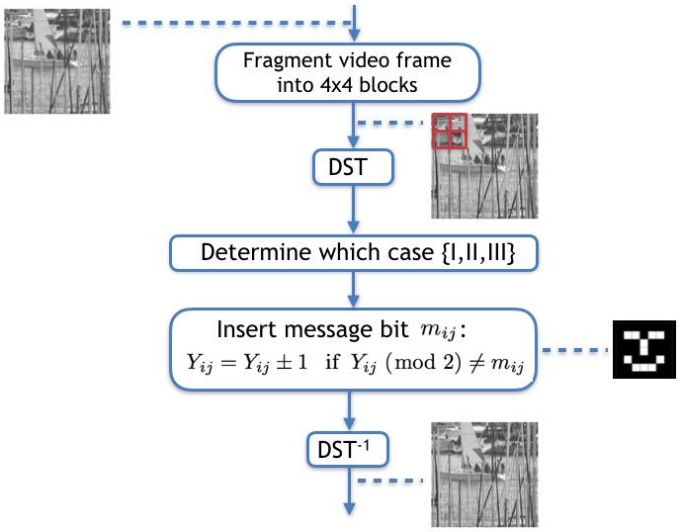



Fig. 2: Embedding steps of the scheme

Bits are embedded dependent on which prediction modes are being used to predict the current block.

- Case (I): if the current block is not dependent on adjacent blocks above the current block
- Case (II): if the current block is not dependent on adjacent blocks to the left of the current block
- Case (III): if the block to be embedded with bits will not influence any adjacent blocks

Let the 4×4 block $Y$ be denoted as follows:

$$Y = \begin{bmatrix} Y_{00} & Y_{01} & Y_{02} & Y_{03} \\ Y_{10} & Y_{11} & Y_{12} & Y_{13} \\ Y_{20} & Y_{21} & Y_{22} & Y_{23} \\ Y_{30} & Y_{31} & Y_{32} & Y_{33} \end{bmatrix} \quad (3)$$

Then embedding is performed based on the above-defined cases:

- Case (I): a bit is embedded into each group of three multi-coefficients: $\{(Y_{00},Y_{02},Y_{03}),(Y_{10},Y_{12},Y_{13}),(Y_{20},Y_{22},Y_{23}),(Y_{30},Y_{32},Y_{33})\}$.
- Case (II): a bit is embedded into each group of three multi-coefficients: $\{(Y_{00},Y_{20},Y_{30}),(Y_{01},Y_{21},Y_{31}),(Y_{02},Y_{22},Y_{32}),(Y_{03},Y_{23},Y_{33})\}$.
- Case (III): 16 bits are embedded into the 4×4 block $Y$ (see elaboration in Equation 3)

Let $(C_1,C_2,C_3)$ represent a multi-coefficient group as per above.

For Case (I) and Case (II) defined above, the embedding proceeds as follows:

- If the embedding bit is 1, then we have:

If $C_1 mod 2 = 0$ and $C_1 \geq 0$,

then $C_1 = C_1 + 1, C_2 = C_2 - 1, C_3 = C_3 + 1$ (4)

If $C_1 mod 2 = 0$ and $C_1 < 0$,

then $C_1 = C_1 - 1, C_2 = C_2 + 1, C_3 = C_3 - 1$ (5)

If $C_1 mod 2 \neq 0$,

then $C_1 = C_1, C_2 = C_2, C_3 = C_3$ (6)

- If the embedding bit is 0, then we have:

If $C_1 mod 2 \neq 0$ and $C_1 \geq 0$,

then $C_1 = C_1 + 1, C_2 = C_2 - 1, C_3 = C_3 + 1$ (7)

If $C_1 mod 2 \neq 0$ and $C_1 < 0$,

then $C_1 = C_1 - 1, C_2 = C_2 + 1, C_3 = C_3 - 1$ (8)

If $C_1 mod 2 = 0$,

then $C_1 = C_1, C_2 = C_2, C_3 = C_3$ )

For Case (III) defined above, the embedding is as follows:

- If the embedding bit is 1, then we have:

If $C_{ij} mod 2 = 0$ and $C_{ij} \geq 0$,

then $C_{ij} = C_{ij} + 1$ (10)

If $C_{ij} mod 2 = 0$ and $C_{ij} < 0$,

then $C_{ij} = C_{ij} - 1$ (11)

If $C_{ij} mod 2 \neq 0$,

then $C_{ij} = C_{ij}$ (12)

- If the embedding bit is 0, then we have:

If $C_{ij} mod 2 \neq 0$ and $C_{ij} \geq 0$,

then $C_{ij} = C_{ij} + 1$ (13)

If $C_{ij} mod 2 \neq 0$ and $C_{ij} < 0$,

then $C_{ij} = C_{ij} - 1$ (14)

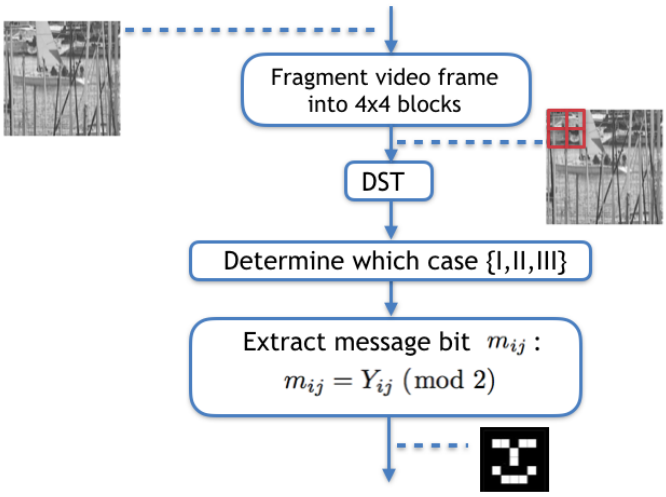
If $C_{ij} mod 2 = 0$,

then $C_{ij} = C_{ij}$. (15)

### 1. B. Data Extraction

The extraction of the embedded data from the video stream is as follows; as illustrated in Fig. 3:

- Case (I): the embedded bit is to be extracted from {( $Y_{00}, Y_{02}, Y_{03}$), ($Y_{10}, Y_{12}, Y_{13}$), ($Y_{20}, Y_{22}, Y_{23}$), ($Y_{30}, Y_{32}, Y_{33}$)} as follows:




- $m_{i0} = Y_{i0} mod 2$ where $i \in \{0, \ldots, 3\}$ (16)

Fig. 3: Extraction steps of the scheme

- Case (II): the embedded bit is to be extracted from {( $Y_{00}, Y_{20}, Y_{30}$), ($Y_{01}, Y_{21}, Y_{31}$), ($Y_{02}, Y_{22}, Y_{32}$), ($Y_{03}, Y_{23}, Y_{33}$)}
- $m_{0i} = Y_{0i} mod 2$ where $i \in \{0, \ldots, 3\}$ (17)

- Case (III): 16 extracted bits can be obtained from the 16 $Y_{ij}$ of the 4×4 *Y* block as follows:
- $m_{ij} = Y_{ij} mod 2$ where $i, j \in \{0, \ldots, 3\}$ (18)

## III. ROBUSTNESS SECURITY OF THE DATA HIDING SCHEME FOR H.265/HEVC VIDEOS

In this section we discuss the security of this scheme, firstly in terms of hiding security, before focussing on robustness security.

### A. Hiding Security

Core security properties that are inherently required in a data hiding scheme are that it is able to remain hidden in terms of two aspects:

- *Secrecy*: embedded message remains hidden
- *Undetectability*: existence of a hidden message cannot be verified [9,10]

By a thorough analysis of the scheme, we can see that from the scheme's extracting steps as in Section II.B that the embedded message can be recovered, irrespective of which of the three cases {I,II,III} is being considered. In more detail, the embedded message bit $\in \{m_{i0}, m_{0i}, m_{ij}\}$ is in fact correlated to whether the DST coefficient $\in \{Y_{i0}, Y_{0i}, Y_{ij}\}$ is even or odd, as per Equations (16~18), and therefore the embedded message bits can be revealed directly from the coefficients. Thus, the *secrecy* property cannot be upheld by the scheme.

In terms of the second property i.e. *undetectability*, an adversary aiming to break this property has to show that it is possible to differentiate between whether a video frame contains an embedded message or does not have any embedded message.

By inspection of the embedding steps as in Section II.A, it can be seen that the message bits are embedded in the least significant bit of the DST coefficients $\{Y_{i0}, Y_{0i}, Y_{ij}\}$. In essence, given a sufficiently large number of coefficients, the distribution of bit values in any bit position is expected to be random, i.e. '0' and '1' have a 50%-50% distribution.

In contrast, if the video frame has been embedded with message bits, which have inherent redundancy [11], therefore their distribution would differ from the 50%-50% random distribution.

Hence, this allows to differentiate between the two cases, thereby breaking the undetectability property.

### B. Robustness Security

Beyond the two hiding security properties discussed in Section III.A, another required security property for data hiding schemes is that of *robustness security* [10], i.e. that the hidden embedded messages are not able to be changed by tampering [12]. Nevertheless, we show an attack on the robustness security of this Liu et al. data hiding scheme. For simplicity of discussion, we use Case III as an example, although the attack equally applies for Case I and Case II.

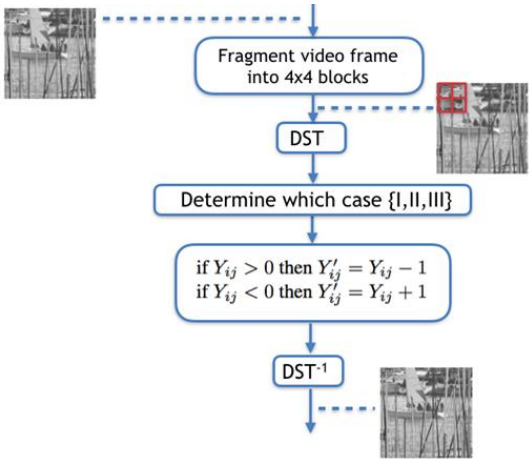



Fig. 4: Robustness attack against the scheme: steps performed by the attacker

The attack steps proceed as follows; see Fig. 4 and Fig. 5 for illustrations:

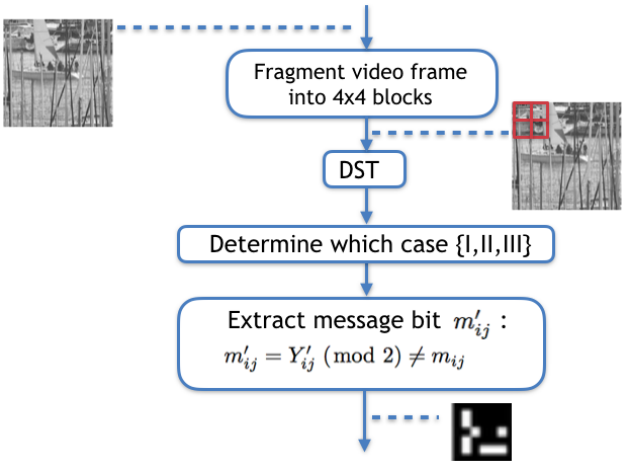



Fig. 5: Robustness attack steps against the scheme: steps performed by the verifying user

2. The video frame *V* is embedded by a legitimate sender with message bits as per the embedding steps of Section II.A, to produce the message-embedded frame *V'*.

3. An attacker aims to illegally modify the embedded message bits in *V'* so that they cannot be detected from the frame *V'* anymore. The attack steps are as follows:

a) Perform the extraction steps of Section II.B on the DST coefficients of *V'* to get the embedded message bits $m_{ij}$.

b) The goal of the attacker is to modify these embedded bits by modifying the message-embedded DST coefficients:

- If $Y_{ij} > 0$,

then $Y'_{ij} = Y_{ij} - 1$ (19)

- If $Y_{ij} < 0$,

then $Y'_{ij} = Y_{ij} + 1$ (20)

c) Perform the data extraction steps of Section II.B on the changed DST coefficients $Y'_{ij}$ to obtain the extracted message bits $m'_{ij}$.

d) Compare the message bits extracted from the modified with the embedded message bits extracted from the original DST coefficients $Y_{ij}$. The robustness attack is successful if $m_{ij} \neq m'_{ij}$.

3. The resultant modified DST coefficients $Y'_{ij}$ are such that the original embedded message bits $m_{ij}$ can no longer be detected because their binary values have been inverted; instead different message bits $m'_{ij}$ would be extracted.

## IV. DISCUSSIONS AND CONCLUDING REMARKS

We analyse the steps of the robustness attacks steps in Section III.B to see why the robustness of the Liu et al. data hiding scheme is broken by the attack, i.e. why $m_{ij} \neq m'_{ij}$.

Recall the embedding steps of Section II.A, namely for Case III, notably the Equations (10~15). If the message bit to be embedded is 1, from Equations (10~11) we see that if the DST coefficient $C_{ij}$ is even, then it is actually changed by the embedding steps to become odd. Otherwise it is left unchanged. This results in effectively embedding a bit value of 1 into $C_{ij}$ because an odd $C_{ij}$ gives $C_{ij} mod 2 = 1$.

Similarly, if the embedding message bit is 0, from Equations (13~14) it can be seen that if the DST coefficient is odd, the embedding steps will change it to be even. Effectively this embeds a bit value of 0 into $C_{ij}$ because an even $C_{ij}$ gives $C_{ij} mod 2 = 0$.

Fig. 6 indicates the corresponding PSNR values computed for some test cover frames of the Xiph.org Video Test Media [13], and can be used to compare the following different cases:

- (i) The attacked marked frame versus the (non-attacked) marked frame. The values for this case are denoted by the small circles in the graph of Fig. 6, and with respect to the left axis. The PSNR is well over 50dB, consistently, therefore this shows the close similarity between the original marked frame and the marked frame after it is attacked to cause changes to the embedded message bits. Therefore, the robustness attack can be performed without compromising frame quality.

- (ii) Marked frame versus the cover frame. For this case, the values are with respect to the right axis. This is the PSNR value that is often used to measure the quality of the cover frame after it has been embedded with message bits.

- (iii) The attacked marked frame versus the cover frame. For this case, the values are with respect to the right axis. This PSNR is used to compare with that of case (ii) to detect if there is any quality difference between the two cases (ii) and (iii).. Fig. 6 shows that these PSNR values for cases (ii) and (iii) are essentially similar, thus showing that there is no significant quality difference between the original marked frame and the attacked marked frame, so the attack

applies without causing any noticeable artefacts.

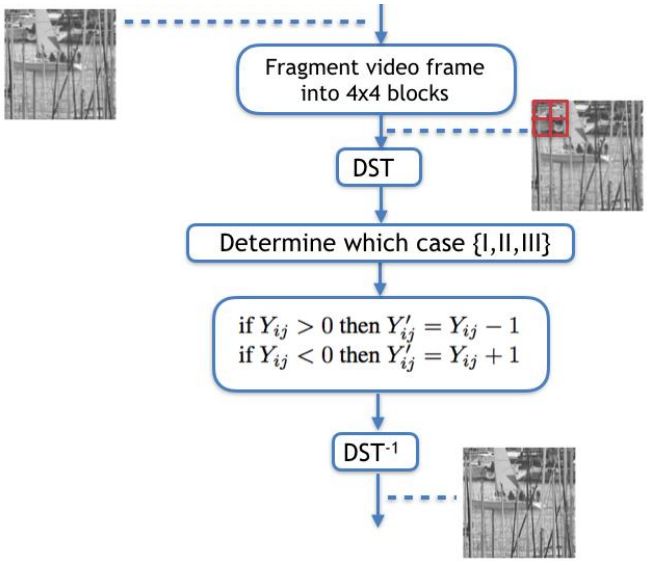



Fig. 7: Comparing the Cover, Marked and Attacked Frames

Fig. 7 illustrates some examples (from the Xiph.org Video Test Media [13] of cover frames before embedding, after embedding, and after the embedded message has been attacked to cause it to be no longer extractable. The frames remain perceptually similar across all these three cases, and the PSNR is consistently over 49dB for all cases. This thereby indicates that the attacks can apply without adversely affecting the visual quality of the frames.

These results show that the Liu et al. scheme is not robust against attacks. The weakness being exploited in the attacks is that the embedded message bits can be recovered (secrecy problem), and one reason for this is because the embedding locations could be determined by the attacker based on case {I,II,III} analysis.

One approach to strengthen against this is to redesign the scheme such that the embedding locations are a selective subset out of all possible block coefficients, and furthermore the order of embedding locations is randomized, initialised by a secret seed that is shared between the embedded and extractor. The robustness attack is then impeded because the attacker since s/he is unaware of the embedding location subset and embedding order, is unable to accurately modify the embedded bits meaningfully.

## REFERENCES

1. G.J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 12, 2012, pp. 1649–1667.
2. T. Luo, L. Zuo, G. Jiang, W. Gao, H. Xu, and Q. Jiang, "Security of MVD-based 3D Video in 3D-HEVC using Data Hiding and Encryption", *Journal of Real-Time Image Processing*, 2018, to be published.
3. A. Boho, G. Van Wallendael, A. Dooms, J. De Cock, G. Braeckman, P. Schelkens, B. Preneel, and R. Van de Walle, "End-to-End Security for Video Distribution: the Combination of Encryption, Watermarking, and Video Adaptation", *IEEE Signal Processing magazine*, Vol. 30, No. 2, 2013, pp. 97–107.
4. Y. Chen, H. Wang, H. Wu, and Y. Liu, "An Adaptive Data Hiding Algorithm with Low Bitrate Growth for H.264/AVC Video Stream", *Multimedia Tools and Applications*, Vol. 77, No. 15, 2018, 20157–20175.
5. M. Long, F. Peng, and H.-y. Li, "Separable Reversible Data Hiding and Encryption for HEVC Video", *Journal of Real-Time Image Processing*, Vol. 14, No. 1, 2018, pp. 171–182.
6. J. Yang, and S. Li, "An Efficient Information Hiding Method based on Motion Vector Space Encoding for HEVC", *Multimedia Tools and Applications*, Vol. 77, No. 10, 2018, 11979–12001.
7. P.-C. Chang, K.-L. Chung, J.-J. Chen, C.-H. Lin, and T.-J. Lin, "A DCT/DST-based Error Propagation-free Data Hiding Algorithm for HEVC Intra-coded Frames", *Journal of Visual Communication and Image Representation*, Vol. 25, No. 2, 2014, 239–253.
8. Y. Liu, S. Liu, H. Zhao, and S. Liu, "A New Data Hiding Method for H.265/HEVC Video Streams without Intra-frame Distortion Drift", *Multimedia Tools and Applications*, 2018, to be published.
9. K. Solanki, N. Jacobsen, U. Madhow, B.S. Manjunath, and S. Chandrasekaran, "Robust Image-Adaptive Data Hiding Using Erasure and Error Correction", *IEEE Transactions on Image Processing*, Vol. 13, No. 12, 2004, 1627–1639.
10. P. Moulin, and R. Koetter, "Data-Hiding Codes", *Proceedings of the IEEE*, Vol. 93, No. 12, 2005, pp. 2083–2126.
11. C.E. Shannon, "Prediction and Entropy of Printed English", *The Bell System Technical Journal*, Vol. 30, No. 1, 1951, pp. 50–64.
12. M. Wu, and B. Liu, "Data Hiding in Binary Image for Authentication and Annotation", *IEEE Transactions on Multimedia*, Vol. 6, No. 4, 2004, 528–538.
13. Xiph.org, "Xiph.org Video Test Media [derf's collection]". Available online at https://media.xiph.org/video/derf, last accessed 17 August 2018.

## AUTHORS PROFILE

**Grace C.-W. Ting** received her B.Eng (Hons) in Electronics and the M.Eng.Sc from Multimedia University (MMU), Malaysia in 2001 and 2007 respectively. She served as a lecturer in the School of Engineering at the Swinburne University of Technology, Sarawak from 2002 to 2006. She is now pursuing her Ph.D with Universiti Tunku Abdul Rahman (UTAR), Malaysia. Her research interests include digital watermarking, multimedia security and digital signal processing.

**Bok-Min Goi** received his B.Eng from University of Malaya (UM) in 1998, and the M.Eng.Sc and Ph.D from Multimedia University (MMU), Malaysia in 2002 and 2006, respectively. He is now the Dean and professor in the Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman (UTAR), Malaysia. He was the General Chair for ProvSec 2010 and CANS 2010, and the PC member for many crypto / security conferences. His research interests include cryptology, security protocols, information security, digital watermarking, computer networking and embedded systems designs.

**Sze-Wei Lee** was born in Malaysia in 1970. He obtained his B.Eng (Hons) in Electronics and Optoelectronics, M.Phil and Ph.D from University of Manchester Institute of Science and Technology, UK in 1995, 1996, and 1998 respectively. He served as a lecturer, senior lecturer and associate professor in the Faculty of Engineering, Multimedia University (MMU) from Jan 1999 to Oct 2008. He joined the Faculty of Engineering and Science, Universiti Tunku Abdul Rahman (UTAR) as a professor in Oct 2008. He is now President of the Tunku Abdul Rahman University College (TARUC). His research interests include digital signal processing, communication systems and protocols.

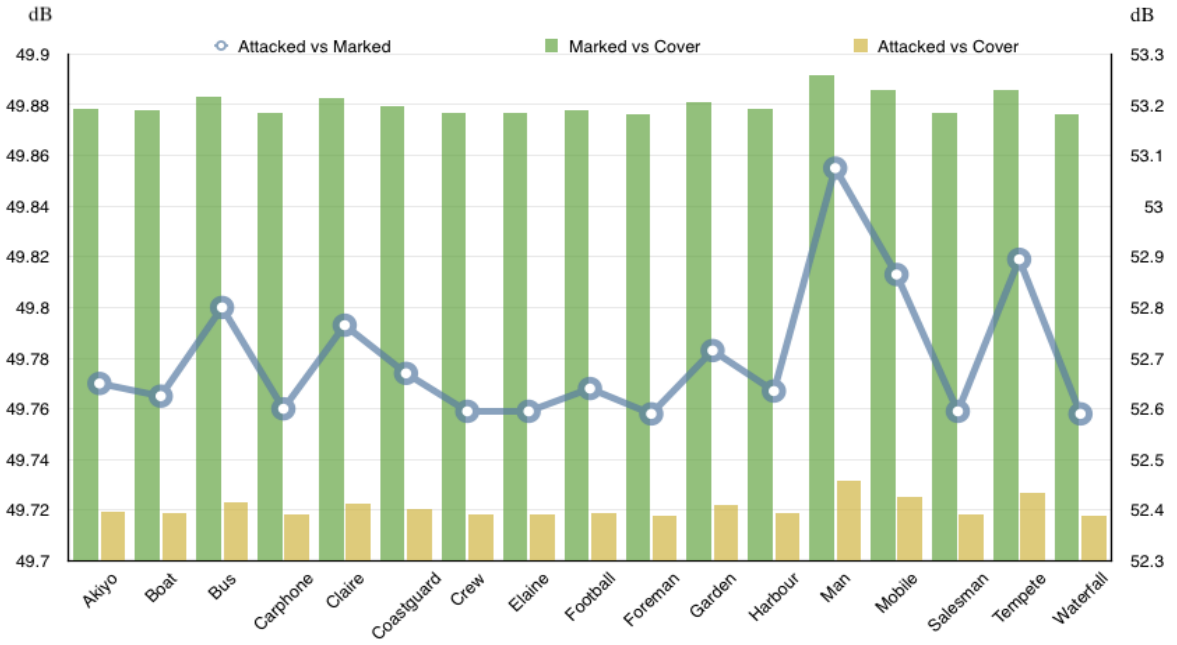



Fig. 6: Comparison of the PSNR for the following cases:
(i) Attacked marked frames vs Marked frames (ii) Marked frames vs Cover frames
(iii) Attacked marked frames vs Cover frames