Multimed Tools Appl (2017) 76:16239–16261
DOI 10.1007/s11042-016-3906-0

Metadata hiding for UAV video based on digital watermarking in DWT transform

Nasr addin Ahmed Salem Al-Maweri[1] · **Aznul Qalid Md Sabri**[1] · **Ali Mohammed Mansoor**[1] · **Unaizah Hanum Obaidellah**[1] · **Erma Rahayu Mohd Faizal**[1] · **Joan Lai P C**[2]

Received: 11 April 2016 / Revised: 21 August 2016 / Accepted: 25 August 2016 /
Published online: 14 September 2016

Abstract As the advent of the Unmanned Aerial Vehicles (UAVs) has been increased, the protection of the information within the transmitted or stored video has become a big challenge. Most known drone systems attach metadata of the recorded video in separate files or in the header of the video. Current techniques make the metadata unsecure and easy to get lost and removed as well as it occupies more storage and bandwidth. In this paper, an efficient method is proposed to hide the metadata of UAVs video using the technology of digital watermarking. Discrete Wavelet Transform (DWT) is used to implement the embedding of the information robustly. The middle frequencies coefficients reside on CH3 sub-band are utilized to hide the watermark bits. In addition, a new scrambling algorithm is proposed to secure the information before hiding. The adaption of the proposed video watermarking algorithm to hide the metadata of the UAV video is achieved. The experimental results prove the high performance of the proposed method. The method had unnoticeable impact on the video quality where the PSNR of 44 dB is attained. The experiments show that the method achieves high robustness under various attacks and provides enough capacity for metadata hiding of UAV video.

Keywords UAV · Drones · Metadata hiding · Video watermarking · Copyright protection · Image scrambling

✉ Nasr addin Ahmed Salem Al-Maweri
senassr_maweri@yahoo.com

[1] Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

[2] UM Centre of Innovation and Commercialization, University of Malaya, 50603 Kuala Lumpur, Malaysia

1 Introduction

Innovations involving Unmanned Aerial Vehicles (UAVs), or commonly known drones have recently progressed rapidly through innovative developments. In recent years [11, 23] unlike the common assumption, the UAVs are used for military missions, but have been incorporated into various civilian applications. The services this device offer can often outperform human capabilities across various places. Nowadays, the UAVs' product specifications have reached sophisticated functionalities. For example, UAVs that can fly up to 50,000 ft can be controlled remotely from thousands of miles from its flying location [13]. Hence, the availability in the market is advantageous. This includes eliminating danger among workers who perform high-risk and life-threatening duties such as those involved in disaster management. The UAVs is increasingly impacting a large contribution in our daily life, organizations and governments. Among the common applications of UAVs are used in the following areas: security and control, military missions, air reconnaissance, traffic monitoring, forest and water search and rescue, firefighting, disaster monitoring and management (i.e., floods, earthquakes), surveys, 3D environment modeling, delivery services, farming and ranching, media reporting and news, border patrol, atmosphere sensing, map guidance, and many more [16, 19, 21, 22].

Although, the advancement in UAVs technology and specifications have shown development of their products rapid progress, issues related to security information to UAV still needs further development and enhancement. By far, developers have added encryption systems and firewalls to achieve trusted transmission. This form of security is considered as the commutation security.

However, once the information is transferred on the drone or transmitted to the Air Traffic Control (ATC), it becomes unsecure. Thus, this situation is considered as the information security issue.

One of the most important tasks of a UAV is to record images or videos during its mission before transferring these data to the ATC. Commonly, UAV will essentially collect the information in the form of videos data that it has captured. This kind of information is called *Metadata*. Metadata contains highly sensitive and valuable information such as the location of the video, the GPS coordinates, the, time and date, the camera angles and what the video is taken for. This type of set of information is notoriously confidential for disclosure especially in areas such as security control, military and maritime.

The current implementation of video systems in UAVs sends the metadata as, either, a separate file or saves the data within the header file of the video. Both techniques do not give concern on the security issues of the metadata. As a result, the data are susceptible to loss or deletion. Recently, a *zombie drone* supported by a system called SkyJack was introduced. The purpose of this type of drone is to hack the systems of other drones enabling it to take advantages of their information and control them [20]. This alarming situation strengthens the importance of addressing the security issues pertaining to the UAVs information.

Various methods and techniques have been released in video watermarking. However, most of the available works still lack high robustness and do not attain optimal trade-off as well as treating low amount of data [6]. These issues made the available algorithms non sufficient for UAVs adoption where more capacity, robustness and security are needed.

Therefore, to protect the metadata from being hijacked, leaked and revealed, we propose a method to hide the metadata using an efficient and robust technique via digital watermarking based on Discrete Wavelet Transform (DWT). In this paper, the proposed work has added a new robust techniques and equations for data hiding in a secure way and improved the

embedding by keeping fine perceptual quality and resisting various malicious attacks. Furthermore it introduced a novel adoption strategy for digital watermarking into UAVs saving the storage as well as increasing the data safety from a potential loss. Moreover, a new scrambling method is developed to encode the metadata before hiding it in the video. The proposed video watermarking algorithm tries to enhance the video watermarking performance in terms of imperceptibility, robustness and capacity compared to the existing video watermarking methods. In addition, we intend to adopt the watermarking technique on the UAVs video metadata. Digital watermarking concerns with embedding some information namely called watermark into another digital signal of data file such as text, video, audio or image. Digital watermarking has always been implemented as three main components watermarks generation, embedding and extraction [5].

The rest of the paper is organized as follows: section 2 reviews the recent video watermarking algorithms. Section 3 discusses the DWT transform. Section 4 explains the proposed algorithm. Section 5 presents the experiments and results of evaluating the proposed video watermarking. Finally, section 6 concludes the paper.

2 Related work

Security challenges and vulnerabilities related to the UAVs development are some issues that air drones vendors worry about since the beginning of the UAVs discovery. For this matter, researchers in the field began to think about solutions in various ways to address these challenges. One research that focused on securing the metadata of the transmitted video was introduced by [14]. In their research, they proposed sending the metadata information within the video by embedding the metadata using digital watermarking. Their method used VLC representation to build the pair trees where the watermark bits are embedded by either changing the specific bits in some blocks to indicate embedding '1' or leaving the block unchanged to indicate embedding '0'. However, the VLC approach in watermarking is now considered obsolete. Researchers in this field have stopped using this method due to the absence of contributions offered to improve this mechanism. That is largely influenced by the complexity in the implementation of such technique. Moreover, [14] have not reported any positive results on the robustness of using VLC trees for the purpose of watermarking. Nevertheless, a previous study by [3] reported some results of using VLC trees for video watermarking. The reported results showed very low performance in term of its robustness compared to the recent works on video watermarking.

In images, digital watermarking has been used firstly by [10] to embed the metadata of taken photos by the digital camera. The associated data such as resolution, date, time, and more were encoded in their algorithm and embedded into the image based on DCT domain. The algorithm showed high PSNR values while the performance was degraded in term of robustness.

Digital video watermarking has evolved rapidly and research improvements in the field are still ongoing. Author in [7] developed an approach for video watermarking using Singular Value Decomposition (SVD) and Discrete Wavelet Transform (DWT). He transformed the video into DWT using two levels of decomposition. The coefficients such as CH2, CV2 and CD2 sub-bands were then transformed into an SVD form before the embedding was performed. The Error Correction Code was integrated to the method to enhance its robustness against any attack. The reported results showed some improvements to the imperceptibility and robustness. However, the Bit Error Rate (BER) was high under some attacks such as

median filtering, frame averaging, JPEG compression and scaling. This indicates that the algorithm is not resilient to potential attacks. Another algorithm was proposed by [15]. In this algorithm, the DWT with three levels of decomposition was used after applying motion detection based on scene change analysis. The watermark bits were embedded using the spread spectrum technique on CH3, CV3 and CD3 sub-bands. This approach showed lower performance compared to [24]. In this result, the Low (PSNR) about 35 with low robustness under attacks such as frame dropping, frame averaging, frame swapping and noise, indicate degraded overall performance.

A different approach was proposed by [24]. In this approach, the embedding was achieved in low frequency coefficients of DCT domain. DCT was applied to only specific regions in the video sequence. These regions were generated based on the matching regions between the frames using the *frame patch matching technique*. The algorithm was concerned to tolerate geometrical attacks. Hence, the KAZE feature matching method was used before extraction is executed to recover the distorted video scene from the original scene. Although the algorithm intends to increase the robustness against geometrical attacks and had survived rotations, it showed low robustness under scaling, frame dropping and frame insertion. Another method for video watermarking that used DCT was proposed by [2]. In their method, DCT was obtained from each frame for the luminance part only after converting the video into the YUV color space. The embedding of the watermark bits was executed diagonally in the coefficients in each block obtained from 8×8 blocks of the DCT transformation. This algorithm showed low PSNR value less than 35. However, it showed some enhancement under certain attacks such as frame swapping. Nevertheless, it showed lower robustness under other attacks such as noising, filtering and geometrical attacks.

Authors in [1] proposed a hybrid video watermarking method which utilized Contourlet Transform (CT), DWT and SVD. In this method, the embedding was performed on DWT coefficients of both middle frequency sub-bands CH and CV. Before embedding, the detection of non-motion frames using histogram difference and applying CT transform for the chosen frames is performed. The watermark image was scrambled using the Arnold transform and sliced into 24 slices using bit plane slicing algorithm. This watermarking approach attained good perception quality with high PSNR values about 60 dB. However, the robustness showed degraded performance under various attacks such as Gaussian noise, salt and pepper noise, frame swapping, and rotation.

Another Algorithm was proposed in [17] to for data hiding in video files which was intended as steganography method. In this algorithm, Kanade-Lucas-Tomasi (KLT) tracking was executed using Hamming codes to track the facial regions and encode the secret message. LSB were used as the main method of embedding the bits on the spatial domain for the chosen regions pixels. Authors of this algorithm tried to increase the visual quality parameter. High PSNR values were attained. But their results showed very low detection accuracy reached 0.760's once the video was attacked which is common with using LSB on spatial domain. Later, the same authors have also proposed a different method in [18] that used transform domain. The method hides the secret message, after encryption, in YUV video planes which is previously transformed to DCT domain. The contribution of the algorithm is only on using BCH error correcting codes for the hidden data before embedding. PSNR attained was between 38.95 and 42.73. However, the robustness of the hidden data was affected under noising attacks as reported.

Recently, a real-time video watermarking algorithm for surveillance networks was proposed by [12]. In this algorithm, the video frames were segmented into candidate-scenes-based after transforming the frames into DCT domain. These selected scenes are obtained by finding the

relationships between the frames. The VLC run level pairs are changed to embed the bits in the macro blocks. The experiments showed an acceptable PSNR value of 42 dB. However, the robustness was still close to the previous mentioned works, if not less, in many situations. Their results were reported using bit correction rate (BCR) in the chosen frames. As reported, the BCR under no attack was negligible between 0.85 and 1 in rare situations.

The lack of works in the field of UAVs security issues and the current threats for the critical data transmitted with UAVs videos as well as the gap in the performance level of the current video watermarking schemes motivates further research and investigation to protect UAVs systems.

3 Discrete wavelet transform

Practically, the techniques of digital video watermarking perform the embedding of the information in the sequence of images that construct the video scene. The embedding is executed either in the spatial domain or in the frequency domain. In spatial domain, the embedding is implemented to modify the pixel values. While, in frequency domain, the media is converted first into another transform and the embedding is executed by modifying the frequency coefficients. However, most of the available works prefer using frequency domain due to providing high robustness, capacity. In addition, this type of domain promotes fewer defects to the visual quality of the watermarked data [4, 9]. The Discrete Wavelet Transform (DWT) is one of the most common frequency forms that can contribute to the digital watermarking technique. It converts the image data from its pixels representation into another coefficients representation using some mathematical equations. For an image I as input for the DWT transform, the transform can be executed by dividing the original image into four equal size matrices, each matrix contains the coefficients of specific features on the images according to frequency level, starting from low frequency to high frequency. Four coefficients representations produce one level of decomposition - Approximations, (CA), Horizontally CH, Vertically (CV) and Diagonally (CD). This level can be further decomposed to multiple levels as shown in Fig. 1. The proposed method is developed based on a DWT which uses the *daubechies* 'db1' filtering family. The CA3 and CV3 coefficients were utilized to perform the embedding in CH3 sub-band as highlighted in Fig. 1.

The principle of utilizing more than one sub-band coefficients for embedding purposes is adopted from [26]. The proposed algorithm is expected to enhance the performance in term of imperceptibility and robustness as well as security by proposing new mathematical embedding equations and mechanism and adapting it to metadata hiding.

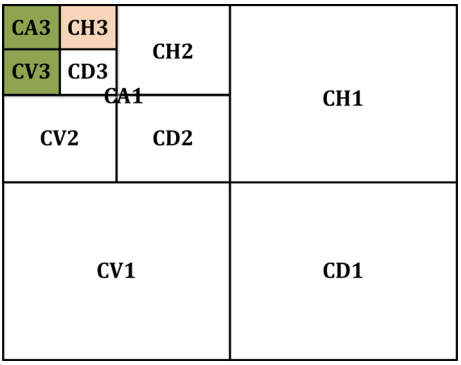


Fig. 1 Chosen DWT sub-bands

4 The proposed watermarking algorithm

The main purpose of developing the proposed digital video watermarking algorithm is to adapt it to the metadata protection and hiding for the videos taken by UAVs systems. This proposed algorithm can be integrated with the videoing system in UAVs either before transmitting the video or once information are received by the data centres, commonly known as the Air Traffic Control (ATC). The proposed video watermarking algorithm is developed in a way that suits hiding the collected data from the UAVs such as the date and time the video is taken, the location where the video is taken, the GPS coordinates of the important scene, altitudes and angels of the camera. These data are sensitive and confidential. Thus, exposure of such data is highly risky. The developed watermarking algorithm will manage scrambling the metadata of the video using two secret keys before generating images of the metadata. This is followed by hiding these images within the recorded video using the DWT domain. Figure 2 illustrates the concept of the proposed metadata hiding application.

4.1 Metadata scrambling and generation

Given that the metadata is confidential, sensitive and susceptible to the recent threats of zombie drones (i.e., UAVs hack systems), it is imperative to encode the metadata before hiding them it in the video. This is achieved in the proposed algorithm by proposing a new scrambling technique. This technique differs from the conventional scrambling mechanisms which have been used in previous watermarking systems. The proposed scrambling technique takes two keys as input and the metadata as image of size (128 × 128) pixels. It scrambles the image of the metadata according to the keys using the proposed mathematical equations. The output from the scrambling process will be in the form of 16 generated and encoded watermarks. These 16 watermarks are generated to adapt the watermarking process for the metadata and video of the UAVs. Figure 3 presents a description of how scrambling of the metadata and generation of the watermarks are performed.

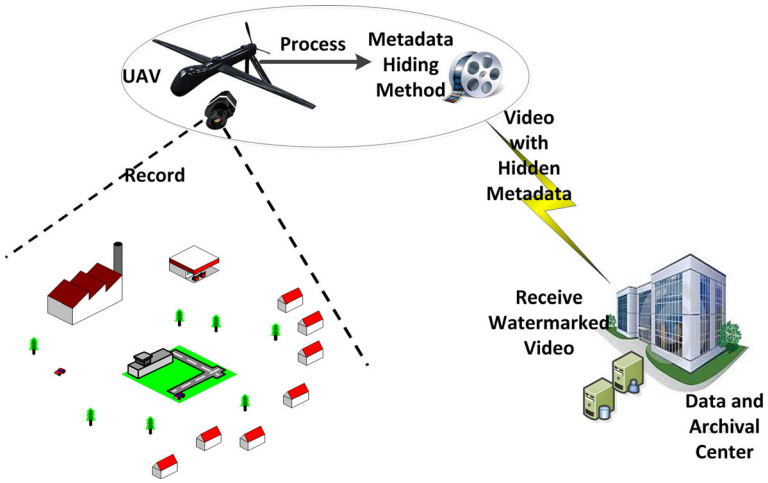


Fig. 2 Metadata hiding in UAVs video

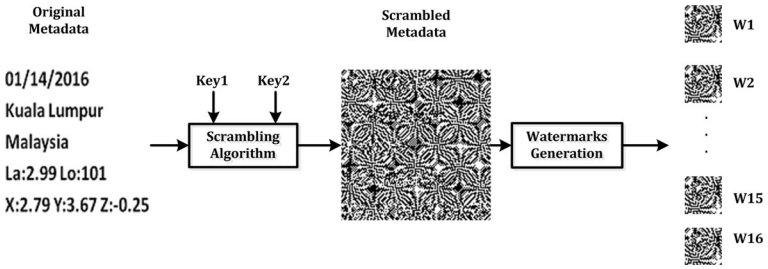


Fig. 3 Metadata scrambling and watermarks generation

The following steps describe the process of scrambling the metadata and generating the watermarks as in Table 1:

4.2 Video watermarks embedding

After generating the scrambled watermarks from the previous process, the embedding of the watermarks in the video taken by the UAV camera is performed. This phase pre-processes the video in which the entire video is split into image sequences or frames. The embedding of the scrambled watermarks is then achieved, and finally, the watermarked frames are reconstructed together to build the watermarked video containing the hidden metadata. Figure 4 shows the proposed video watermarking process.

The embedding of the watermarks in the video frames is executed by embedding 16 watermarks, each in one frame. The process is repeated specific times for another 16 frames, obtained from the original metadata as in Fig. 3, to increase the robustness of the extraction.

Table 1 Metadata scrambling and generation algorithm

Step 1: Input Key1, Key2, OriginalDataImage (128 × 128), where Key 1 and Key 2 are two 5 digit numbers.

Step 2: Create Matrix1 (128 × 128) as:

$$Matrix(i,j) = \frac{round(Key1*(i+j))}{\sqrt{Key1}} \quad (1)$$

Where i, j is the row and column number.

Step 3: Create Matrix2 as:

$$Matrix2(i,j) = \frac{round(Key2+(i*j))}{\sqrt{Key1}} \quad (2)$$

Where i, j is the row and column number.

Step 4: From Matrix1, generate BinaryImage1 as:

$$BinaryImg1(i,j) = \begin{cases} 1, & Matrix1(i,j)\ is\ even \\ 0, & Matrix1(i,j)\ is\ odd \end{cases} \quad (3)$$

Step 5: From Matrix2, generate BinaryImage2 as:

$$BinaryImg2(i,j) = \begin{cases} 1, & Matrix2(i,j)\ is\ even \\ 0, & Matrix2(i,j)\ is\ odd \end{cases} \quad (4)$$

Step 6: Convert OriginalDataImage to Binary.

Step 7: Output = OriginalDataImage XOR BinaryImg1.

Step 8: scrambledImage = Output XOR BinaryImg2.

Step 9: Divide the scrambledImage into 16 equal images of size (32 × 32).

Step 10: Return the scrambled 16 watermarks from Step 9.

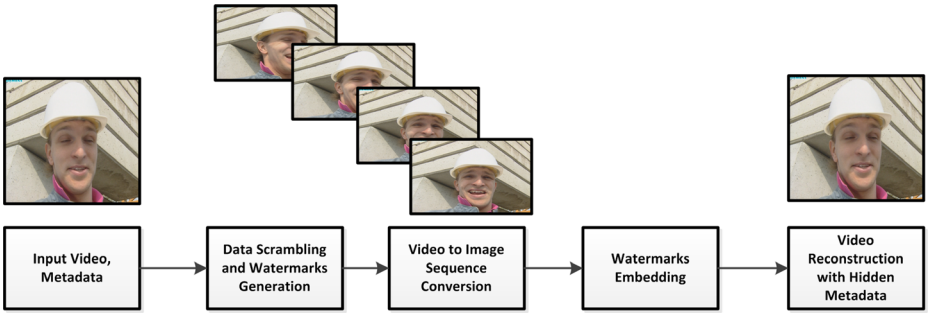


Fig. 4 Video watermarks embedding process

The algorithm is made flexible to repeat the embedding process multiple times according the video length. The following steps explain the embedding procedure in the video as in Table 2:

4.3 Frame watermark embedding

During embedding the watermarks in the video sequence images, each watermark embedding in each frame is performed by calling the frame watermarking procedure as in step 6 in section 4.2. The frame embedding procedure purpose is to insert the bits of the watermark sized as (32 × 32) pixels to the *ith* frame on the image sequences. It takes the frame image as RGB and the scrambled watermark, performs the embedding in YCbCr color space using DWT transform then produces the watermarked frame with the hidden metadata in RGB again. As explained previously, the watermarking is achieved in the middle frequency coefficients, specifically, in CH3, to enhance both visual quality impact and tolerance of the hidden data to survive attacks and video distortions. Figure 5 describes the frame watermarking process.

The frame watermark embedding is executed according to the following steps as in Table 3:

4.4 Video watermarks extraction

Since the proposed watermarking algorithm is developed to have blind extraction scheme, which means that there is no need for the original video to be available upon extraction. The archived

Table 2 Video watermark embedding process

Step 1: Input Video, OriginalDataImage.
Step 2: Scramble and generate 16 watermarks of size (32 × 32) from OriginalDataImage. (Table 1).
Step 3: Read the video and convert it into frames (image sequence)
Step 4: Set $i = 1$. *Stop* = length (video).
For $i = 1$ to Stop
Step 5: Read watermarks from *W1* to *W16*. *Wn* refers to the watermark part.
Step 6: Call *Frame Watermark Embedding* (Table 3) to embed each watermark in each *ith* frame until *W16*.
Step 7: $i = i + 16$.
Step 8: Save the watermarked frames.
Step 9: Loop until Stop, or prior specified *ith* frame.
If i = stop; end
Step 10: Reconstruct the video from the watermarked frames.
Step 11: Return the watermarked video.

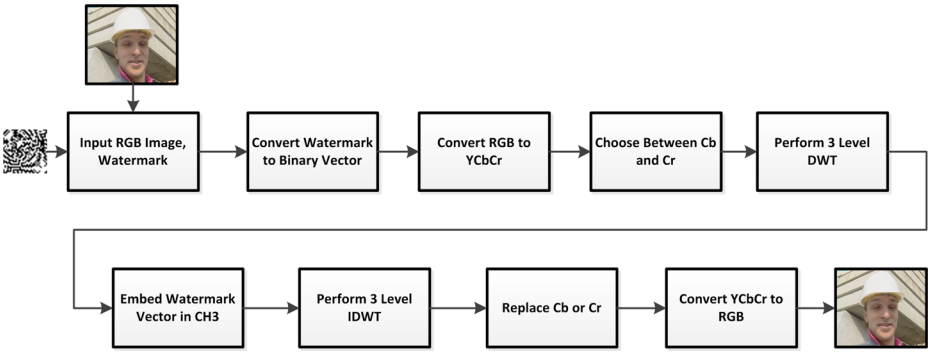


Fig. 5 Frame watermarking process

watermarked video which contains the hidden metadata can be used on the extraction algorithm in a blind way in order to extract the hidden data in the video. The video extraction procedure is performed by splitting the video into image sequence, where the hidden 16 watermarks are extracted. Then the watermarks are reconstructed to form the 128 × 128 pixels image and finally the metadata are descrambled. Figure 6 shows the process of video watermarking extraction.

Table 3 Frame watermark embedding algorithm

Step 1: Input RGB image, Watermark, Scaling factors (α, β).

Step 2: Convert Watermark to Binary Vector (1024 bit).

Step 3: Convert RGB to YCbCr color space.

Step 4: Calculate Average Chrominance Red (Cr) and Chrominance Blue (Cb) where:

$$I = \begin{cases} Cb, & Cr < 3 \\ Cr, & Cr > 3 \end{cases} \quad (5)$$

Step 5: Perform 3 Level DWT for image I using 'db1'. Where db1 is Daubechies wavelet filter.

Step 6: Embed Watermarks Bits one by one in CH3 sub-band as follow:

• $k = 1$; $Stop$ = length (watermark), where k is a loop counter

For $k = 1$ to $Stop$

• Compute Δ_1 and Δ_2

$$\Delta_1 = \frac{|CA3(i,j)| - |CV3(i,j)|}{\beta} \quad (6)$$

$$\Delta_2 = \frac{|CA3(i,j)| - |CH3(i,j)|}{\Delta_1} \quad (7)$$

Where β is scaling factor, CA, CV and CH are the wavelet sub-bands

• Modify CH3 coefficients where:

$$\mathrm{CH3(i, j)} = \begin{cases} \dfrac{\Delta_2 + CH3(i,j)}{a}, & Watermark(k) = 1 \\ CH3(i,j), & Watermark(k) = 0 \end{cases} \quad (8)$$

Where α is scaling factor

• $k = k + 1$.

• Negate back Negative coefficients.

Step 7: Loop in CH3 until all watermarks bits are embedded.

If $k = stop$ end

Step 8: Perform the inverse 3 Level IDWT for modified image I.

Step 9: Convert YCbCr to RGB.

Step 10: Return the watermarked frame.

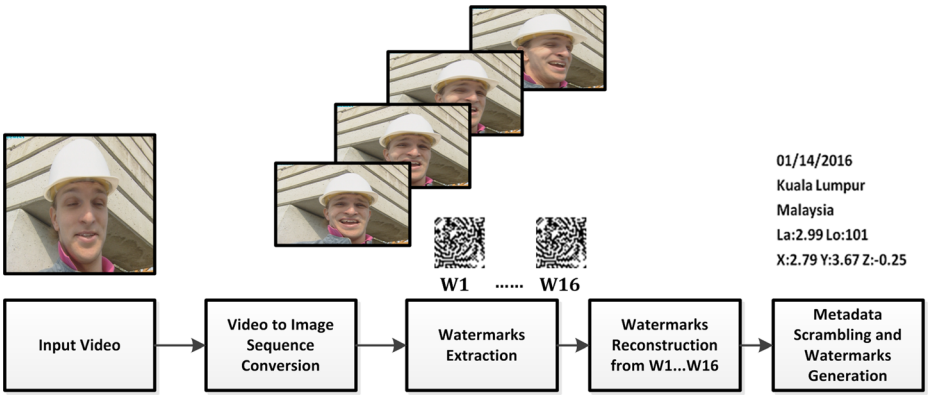


Fig. 6 Video watermarks extraction process

The following steps explain how the extraction procedure from the video is implemented as in Table 4:

4.5 Frame watermarks extraction

To extract each watermark from each watermarked frame, a *Frame Watermark Extraction* is implemented and being called by the main video extraction procedure as in step 4, section 4.4. The watermarked frame in RGB color space and the two scaling factors β and α values used in the embedding process are used again with the same values in the extraction procedure. The RGB image is converted to YCbCr color space and the extraction is done in DWT domain from CH3 sub-band as shown in Fig. 7.

Using the scaling factors in the proposed mathematical embedding and extraction formulas made it possible to enhance the trade-off between imperceptibility and robustness of the watermarking. The steps in Table 5 explain the implementation of the extraction procedure for specific frame:

4.6 Metadata descrambling

By the end of the extraction process from each 16 frames and the reconstruction of the scrambled watermark, the scrambled data image is taken as input to a proposed descrambling technique. This procedure will reverse the encoded data to its original form. The steps in Table 6 clarify the implemented descrambling procedure.

Table 4 Video watermark extraction process

Step 1: Input Video.
Step 2: Read the video and Convert it into frames (image sequence)
Step 3: Set $i = 1$. *Stop* = length (video). Where i is a loop counter.
For $i = 1$ to *Stop*
Step 4: Call *Frame Watermark Extraction* (Table 5) to extract each watermark from each *ith* frame until *W16*.
Step 5: Reconstruct the Extracted Watermark from *W1* ... *W16*, then, Descramble (Table 6) the extracted Watermark (128×128).
Step 6: Save the Extracted Watermark (Hidden Data).
Step 7: $i = i + 16$.
Step 8: Loop specified *ith* frame.
If i = *Stop* end

Table 5 Frame watermarks extraction algorithm

Step 1: Input RGB image, α, β.

Step 2: Convert RGB to YCbCr color space.

Step 3: Calculate Average Chrominance Red (Cr) and Chrominance Blue (Cb) where:

$$\bar{I} = \begin{cases} Cb, & Cr < 3 \\ Cr, & Cr > 3 \end{cases} \quad (9)$$

Step 4: Perform 3 Level DWT for image I using 'db1'.

Step 5: Extract Watermarks Bits one by one from CH3 sub-band as follow:

• $k = 1$; Stop = length (watermark)

For $k = 1$ to *Stop*

• Compute Δ_1, Δ_{1ex} and Δ_{2ex} such that:

$$\overline{\Delta_1} = \frac{|CA3(i,j)| - |CV3(i,j)|}{\beta} \quad (10)$$

$$\Delta_{2ex} = |CH3(i,j)| * a - |CH3(i,j)| \quad (11)$$

$$\Delta_{1ex} = \frac{|CA3(i,j)| - |CH3(i,j)|}{\Delta_{2ex}} \quad (12)$$

• Where Δ_{1ex} and Δ_{2ex} are computed from the watermarked data upon extraction. α, β are the same scaling factors used in embedding.

• Extract Watermark Bit from CH3 coefficients where:

$$Watermark(k) = \begin{cases} 0, & \overline{\Delta_1} - \Delta_{1ex} > x \\ 1, & \overline{\Delta_1} - \Delta_{1ex} < x \end{cases} \quad (13)$$

Where $x = 1, \ldots 10$ according to the attack strength

• $k = k + 1$.

Step 6: Loop in CH3 until all watermarks bits are extracted.

If $k = stop$ end

Step 7: Convert the watermark vector (1024 bit) to 2D binary image (32 × 32).

Step 8: Return the Extracted Watermark.

5 Experiments and results

5.1 Experiments setup

An experiment was performed to evaluate the performance of the proposed video watermarking algorithm. This is done 1) to validate the goal of attaining optimal trade-off between keeping the visual quality of the watermarked (after metadata hiding) video as close as to the quality of the original video, and 2) to evaluate the robustness of the proposed algorithm against malicious attacks. This is also to ensure that both capacity and security metrics are taken into consideration.

Table 6 Metadata descrambling algorithm

Step 1: Input *Key1, Key2, ScrambledDataImage* (128 × 128). Where *Key1, Key2* are the same 5 digit numbers used in embedding.

Step 2: Create $Matrix1_{EX}$ (128 × 128) instead and use equation (1).

Step 3: Create $Matrix2_{EX}$ instead and use equation (2).

Step 4: From $Matrix1_{EX}$, generate $BinImg1_{EX}$ according to equation (3).

Step 5: From Matrix2, generate $BinImg1_{EX}$ according to equation (5)

Step 6: Convert *ScrambledDataImage* to Binary.

Step 7: Output = *ScrambledDataImage* XOR $BinImg2_{EX}$

Step 8: *DescrambledDataImage* = Output XOR $BinImg1_{EX}$

Step 9: Return the *DescrambledDataImage*.

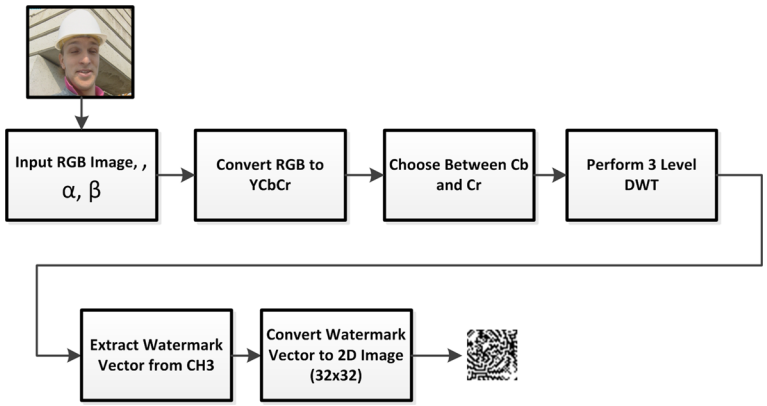


Fig. 7 Frame watermarks extraction process

Four different standard videos known as Foreman, News, Akiyo and Bus with resolution of (352 × 288), 10 s long and 30 frame/s frame rate are used to execute the experiments. Another video was downloaded from WolFang Digital, a video production studio, taken by a UAV of the Kuala Lumpur City, named as UAVVideo is used to repeat the same assessment to validate the measured metrics on actual UAV video.

To measure the imperceptibility and the data invisibility, every video is watermarked with the metadata represented in image of size 128 × 128 pixels as illustrated in Fig. 3. The original video sequence and the watermarked video are used to measure the average Peak Signal to Noise Ration (PSNR) value and similarity percentage [8, 25] according to the following formulas:

$$PSNR = 20\log_{10}\frac{MAX\left(O_{imseq}\right)}{RMSE} \tag{14}$$

Such that:

$$RMSE = \sqrt{\frac{1}{mxn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[O_{imseq}(i,j) - W_{imseq}(i,j)\right]^2} \tag{15}$$

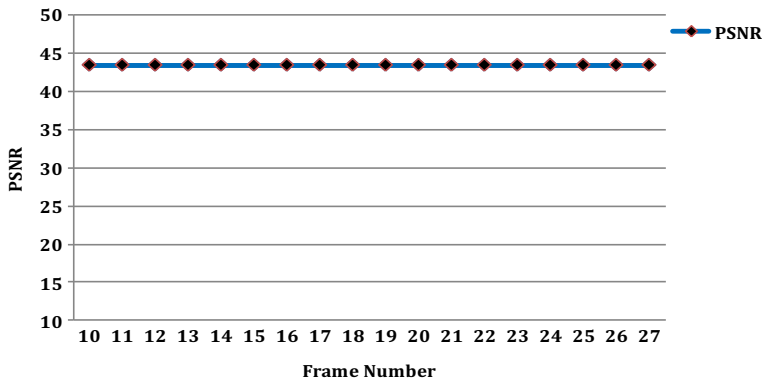


Fig. 8 PSNR for watermarked video frames

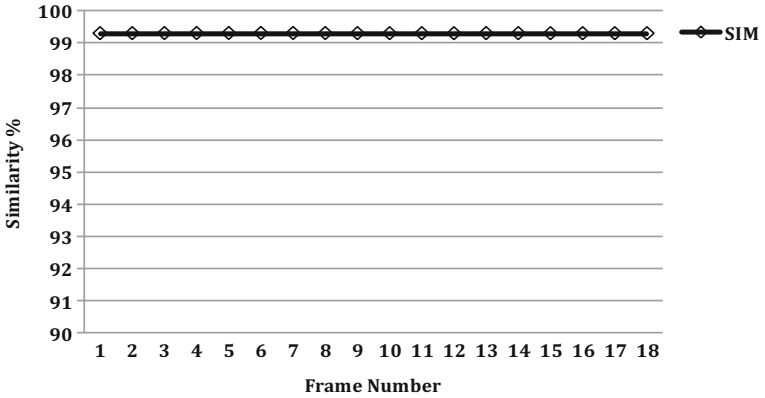


Fig. 9 Similarity measure for watermarked frames

where *mxn* is the size of the image, *(i,j)* is the pixel location, O_{imaseq} is the original image sequence and W_{imseq} is the watermarked image sequence.

$$Similarity = \left(1 - \frac{RMSE}{MAX\left(O_{imseq}\right)} x100\right)\% \tag{16}$$

To measure the extraction accuracy and the robustness of the proposed algorithm, every watermarked video was used to extract the hidden metadata. The accuracy of the extracted watermarks considering various attacks such as salt and pepper noise, Gaussian noise, scaling, frame dropping frame swapping, frame averaging and compression is evaluated by measuring both Bit error Rate (BER) and Normalized correlation (NC) according to the following formula:

$$NC = \left(\frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} O(i,j) - E(i,j)}{\sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} O^2(i,j)} \times \sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} E^2(i,j)}} \right) \tag{17}$$

Where, O is the original watermark and E is the extracted watermark.

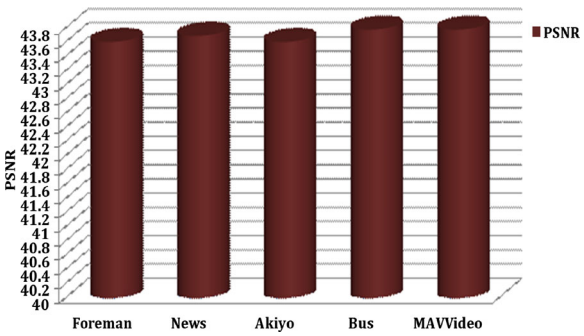


Fig. 10 Average PSNR for various watermarked videos

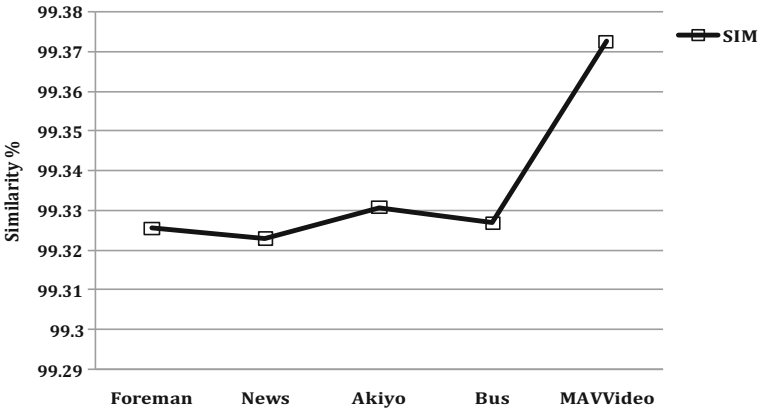


Fig. 11 Average SIM for various watermarked videos

5.2 Results and discussion

5.2.1 Imperceptibility

The impact on the perceptual quality of the watermarked video was tested on all videos mentioned previously. Figures 8 and 9 show the PSNR and Similarity Values, VS the watermarked frames from 10th to 27th frame in Forman video. PSNR showed high value of 44 in all frames and similarity above 99.32 %. This indicates high imperceptibility of the proposed algorithm.

The average PSNR and Similarity Values for different videos are measured and reported in Figs. 10 and 11. As observed, the performance of the proposed algorithm in terms of imperceptibility is stable with different kind of videos. The PSNR value attained is higher compared to the current algorithms in video watermarking as shown in Tables 7, 8 and 9.

5.2.2 Robustness

The goal of the proposed algorithm is to achieve optimal performance by obtaining the trade-off between both PSNR and NC or BER values. The NC values measure the accuracy of the extracted watermarks compared to the original ones that carry the metadata. Figure 12 shows the robustness of the proposed algorithm from the watermarked frames. The NC values were always '1' which indicate perfect accuracy. Table 10 shows the extracted watermarks from

Table 7 PSNR VS NC and BER for Akiyo video

Algorithm	PSNR	NC	BER
T. M. Thanh	37.52	1	0.000
Rakesh	36.00	0.997	0.003
R. J. Mstafa	40.21	1	N/A
Proposed Algorithm	43.60	1	0.000

Table 8 PSNR VS NC and BER for Suzie video

Algorithm	PSNR	NC	BER
T. M. Thanh	37.14	1	0.030
M. Masoumi	35.00	0.994	0.007
Proposed Algorithm	43.69	1	0.000

different videos with their associated values of NC and BER. As observed, all the NC values are 1 and BER value are 0.000.

As mentioned, it was necessary to attain optimal performance while achieving the trade-off between imperceptibility and robustness. Tables 7, 8 and 9 also describe the results of PSNR vs. NC and BER for the proposed algorithm and compares it with other algorithms. As shown, the proposed algorithm performs better than the other reported works, where the PSNR attained from the current work is higher than all of them. At the same time, the NC values reached the highest value 1 and BER was 0.000. Other algorithms attained lower PSNR and NC but higher BER. Although other studies tried to increase the PSNR values, their method affected the NC and BER values. The reasons behind getting better performance in our algorithm can be summarized as follow: Firstly, using the middle frequency coefficients from CH sub-band had fewer defects to the visual quality. Secondly, the proposed modification of the mathematical equations enabled the trade-off and control of the robustness and

Table 9 PSNR VS NC and BER for bus and news video

Algorithm	PSNR	NC	BER
S. Liu (Bus)	40.00	0.97	0.040
S. Liu (news)	40.00	0.98	0.020
R. J. Mstafa	38.95	1	N/A
Proposed Algorithm	43.78	1	0.000

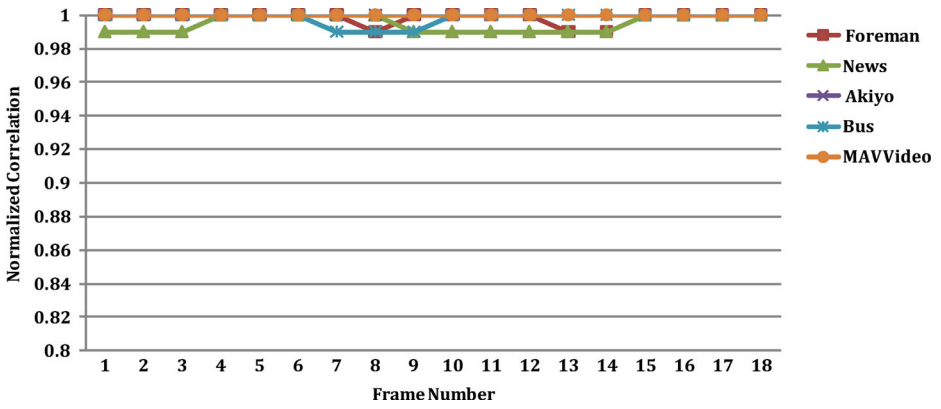
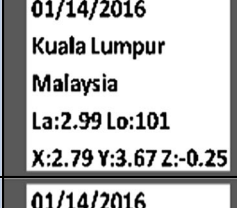
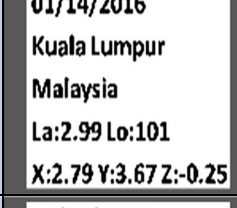


Fig. 12 NC values for extracted watermarks from watermarked frames

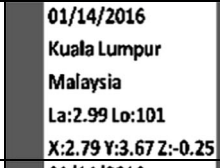
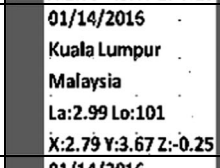
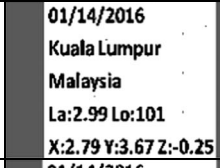
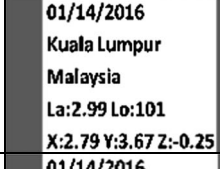
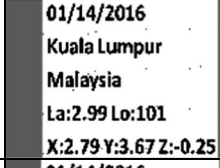
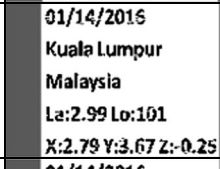
Table 10 Accuracy of the proposed algorithm

Video	Extracted Watermark	NC	BER
Foreman		1	0.000
News		1	0.000
Akiyo		1	0.000
Bus		1	0.000
UAVVideo		1	0.000

imperceptibility. Finally, the strategy of splitting the metadata over 16 frames and repeating the embedding in other frame blocks added high imperceptibility and provided more chance to detect the watermarks in case of malicious attacks.

To validate the resistance of the proposed algorithm against malicious attacks and video distortions, it was tested under certain attacks and the results are reported in Table 11. The extracted watermarks from the attacked videos are shown with each video, the NC and BER values are listed. From the results, it is proven that the algorithm survived most of the attacks with NC value 1 and BER 0.000. The algorithm continued to perform with high NC values around 0.99 on the other attacks such as noising, frame averaging and compression.

Table 11 Robustness against various attacks

Attack Type	Extracted Data Image	NC	BER
No Attack		1	0.000
Salt and Pepper Noise		0.992	0.001
Gaussian Noise		0.996	0.003
Scaling		1	0.000
Frame Dropping		1	0.000
Frame Swapping		1	0.000
Frame Averaging		0.996	0.005
Compression 10%		0.995	0.005
Compression 50%		0.975	0.03

In order to prove the high performance of the proposed algorithm, it was compared with the other algorithms as listed in Table 12. It is obvious from the result that the proposed algorithm worked better than the other algorithms in surviving various attacks. Figure 13 also presents the comparison with the other algorithms (with and without attacks). The neglected difference in NC in some rare situations such as Gaussian noise where other algorithms might be higher is resulted from achieving an optimal trade-off between performance parameters. Hence, the proposed algorithm attains much higher values in other attacks such as scaling.

5.2.3 Capacity and security

As contributed to enhancing imperceptibility and robustness metrics, the capacity and security issues for the proposed video watermarking algorithm is not neglected. The capacity of the algorithm is able to embed a payload size of 128×128 pixels, which means, it is able to hide a total of 16,384 bits using 16 frames. In each frame, 32×32 pixels, which means 1024 bit, can be embedded. Setting up the capacity parameters in such manner helped in attaining high performance on the other side for both imperceptibility and robustness.

The security of the algorithm, as discussed in sections 4.1 and 4.6, is achieved by scrambling the metadata using two secret keys. Without knowing the keys, the hidden data cannot be revealed.

6 Conclusion

In this paper, the problem of disclosing, leaking or revealing the metadata of UAV recorded videos was solved by developing an enhanced video watermarking algorithm to be adapted to UAVs metadata hiding application. The proposed algorithm was implemented in DWT transform and the middle frequency coefficients were utilized. A new scrambling mechanism was implemented to secure the hidden metadata. The imperceptibility and robustness of the algorithm was enhanced to provide an optimal trade-off. The performance of the algorithm has been proven by testing it against various attack and video distortions. Testing and enhancing the algorithm under other geometrical attacks such as rotation is ongoing. Future research and test can be conducted with government agencies like the police force that will be using the drone

Table 12 Comparison between the proposed algorithm and previous works

Attack Type	T. M. Thanh [24]	Rakesh [2]	M. Masoumi [15]	L. Agilandeeswari [1]	Proposed
No Attack	1	0.997	0.994	0.999	1
Salt and Pepper Noise	N/A	0.939	0.956	0.96	0.992
Gaussian Noise	1	0.973	0.956	0.975	0.996
Scaling	0.69	N/A	N/A	N/A	1
Frame Dropping	0.7	N/A	0.949	0.9751	1
Frame Swapping	1	0.983	0.943	0.915	1
Frame Averaging	1	N/A	N/A	N/A	0.996
Compression 50 %	0	N/A	0.9192	N/A	0.975

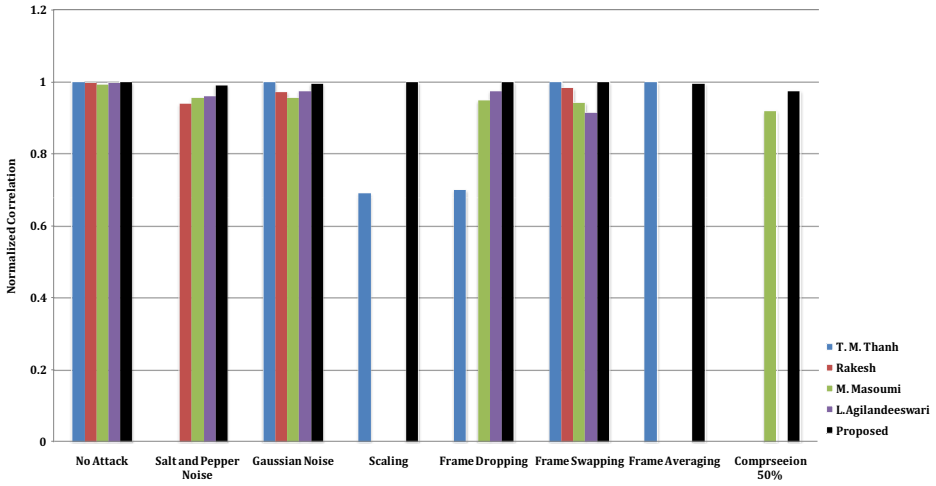


Fig. 13 Comparison of the robustness with various algorithms

or organizations using drone as their service like surveyors etc. The proposed algorithm is tested under normal size videos. Using video watermarking with big data and large files can be considered a challenging. Extending and testing the algorithm to work in real time videos is an open future direction.

Acknowledgments This work is funded under the University of Malaya's Research Grant (UMRG), grant number RP030A-14AET and the Fundamental Research Grant (FRGS), grant number FP061-2014A.

References

1. Agilandeeswari L and Ganesan K (2015) A robust color video watermarking scheme based on hybrid embedding techniques. Multimedia Tools and Applications, Springer
2. Ahuja R and Bedi SS (2015) Copyright protection using blind video watermarking algorithm based on mpeg-2 structure. In: International Conference on Computing, Communication and Automation (ICCCA2015) IEEE, pp. 1048–1053
3. Alattar AM, Lin ET, Celik MU (2003) Digital watermarking of low bit-rate advanced simple profile mpeg-4 compressed video. IEEE Trans Circuits Syst Video Technol 13:787–800
4. Al-Maweri NAAS, Adnan WAW, Ramli AR, Samsudin K, Ahmad SMS (2015) A hybrid digital image watermarking algorithm based on dct-dwt and auto-thresholding. Secur Commun Netw 8(18):4373–4395. doi:10.1002/sec.1371
5. Al-Maweri NAAS, Ali R, Adnan WAW, Ramli AR and Rahman SMSAA (2016) State-of-the-art in techniques of text digital watermarking: challenges and limitations. J Comput Sci
6. Chen Y-H, Huang H-C (2015) Coevolutionary genetic watermarking for owner identification. Neural Comput & Applic 26:291–298
7. Faragallah OS (2013) Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain. AEU Int J Electron Commun 67(3):189–196, [**Online**]. **Available: http://www.sciencedirect.com/science/article/pii/S1434841112001781**
8. Gu K, Zhai G, Yang X & Zhang W (2012) An improved full-reference image quality metric based on structure compensation Signal Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific, 1–6
9. Gupta M, Parmar G, Gupta R, Saraswat M (2015) Discrete wavelet transform-based color image watermarking using uncorrelated color space and artificial bee colony. Int J Comput Intell Syst 8(2):364–380. doi:10.1080/18756891.2015.1001958

10. Huang H-C, Fang W-C (2010) Metadata-based image watermarking for copyright protection. Simul Model Pract Theory 18:436–445
11. Lilien LT, Othmane LB, Pelin A, de Carlo A, Salih RM, Bhargava B (2014) A simulation study of adhoc networking of UAVs with opportunistic resource utilization networks. J Netw Comput Appl 38:3–15
12. Liu S, Chen DB-W, Gong L, Ji W and Seo S (2015) A real-time video watermarking algorithm for authentication of small-business wireless surveillance networks. Int J Distrib Sensor Netw, Hindawi
13. Lowenthal MM (2015) Intelligence: from secrets to policy, 6th ed. USA: CQ Press, SAGE Publication
14. Marcinak MP and Mobasseri BG (2005) Digital video watermarking for metadata embedding in uav video. In: Military Communications Conference, IEEE
15. Masoumia M, Amirib S (2013) A blind scene-based watermarking for video copyright protection. AEU Int J Electron Commun 67:528–535
16. Mohammed F, Idries A, Mohamed N, Al-Jaroodi J and Jawhar I (2014) Uavs for smart cities: Opportunities and challenges. In: Unmanned Aircraft Systems (ICUAS), International Conference on, May 2014, pp. 267–273
17. Mstafa RJ & Elleithy KM (2015) A video steganography algorithm based on Kanade-Lucas-Tomasi tracking algorithm and error correcting codes Multimed Tools Appl
18. Mstafa RJ & Elleithy KM (2016) A DCT-based robust video steganographic method using BCH error correcting codes 2016 I.E. Long Island Systems, Applications and Technology Conference (LISAT), pp 1–6
19. Nex F, Remondino F (2013) Uav for 3d mapping applications: a review. Appl Geomatics 6(1):1–15. doi:10.1007/s12518-013-0120-x
20. Paganini P (2014) Privacy and security issues for the usage of civil drones. Infosec Inistiute
21. Quaritsch M, Kruggl K, Wischounig-Strucl D, Bhattacharya S, Shah M, Rinner B (2010) Networked uavs as aerial sensor network for disaster management applications. E & I Elektrotech Inftech 127:56–63
22. Rango A, Laliberte A, Steele C, Herrick JE, Bestelmeyer B, Schmugge T, Roanhorse A, Jenkins V (2006) Using unmanned aerial vehicles for rangelands: current applications and future potentials. Environ Pract 8:159–168, **[Online]. Available: http://journals.cambridge.org/article_S1466046606060224**
23. Saleem Y, Rehmani MH, Zeadally S (2015) Integration of cognitive radio technology with unmanned aerial vehicles: issues, opportunities, and future research challenges. J Netw Comput Appl 50:15–31, **[Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804514002811**
24. Thanh TM, Hiep PT, Tam TM, Tanaka K (2014) Robust semi-blind video watermarking based on frame-patchmatching. AEU Int J Electron Commun 68:1007–1015
25. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13:600–612
26. Zhao Y and Zhou Z (2012) Multipurpose blind watermarking algorithm for color image based on DWT and DCT. In: 2012 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), pp 1–4

Nasr addin Ahmed Salem Al-Maweri received his Bachelor degree (with first class Honors) in Software Engineering from Albalqa Applied University, Jordan, in 2007. In 2011, he received his Master degree in Computer Systems Engineering from Universiti Putra Malaysia (UPM), Selangor, Malaysia. In 2016, he received

his PhD in Computer and Embedded Systems Engineering from Universiti Putra Malaysia (UPM), Selangor, Malaysia. During 2007 and 2008, he has worked as a software engineer, software systems designer and developer at Batelco Jordan. He also has worked as software engineering trainer and lecturer for 1 year in NCC Education Center Sana'a, Yemen. Currently, He is working as researcher with Faculty of Computer Science and Information Technology, University of Malaya, Malaysia. His research interest covers information security, Software development, image processing, computer vision, computer architecture and operating systems.

Aznul Qalid Md Sabri currently holds the position of Senior Lecturer at the Department of Artificial Intelligence, Faculty of Computer Science and Information Technology (FCSIT), University of Malaya, Malaysia. He is a graduate of the prestigious Erasmus Mundus Master in Vision and Robotics (ViBot), a Master program jointly coordinated by three different universities (University of Burgundy, France, University of Girona, Spain and Heriot-Watt University Edinburgh, United Kingdom). He then completed his Master's degree by performing a research internship program at the Commonwealth Scientific Research Organization (CSIRO) in Brisbane, Australia focusing on Medical Imaging. Next, he pursued his PhD on the topic of "Human Action Recognition" (completed with distinction, très honorable), under a program jointly offered by a well-known research institution in France, Mines de Douai (a research lab) and the reputable University of Picardie Jules Verne, Amiens, France. He is an active researcher in the field of Artificial Intelligence, having published in multiple international conferences as well as international journals. His main research interest is in the field of Computer Vision, Robotics and Machine Learning. He is part of the pioneering members of FCSIT's COVIRO (Cognitive, Vision and Robotics) research group and is currently the principal investigator of multiple research grants.

Ali Mohammed Mansoor received the B.Sc. degree from Amman University, Jordan, and his MSc from University Putra Malaysia, Malaysia, both in computer Science and networks in 2005 and 2008, respectively. He received his PhD degree in Communications and Network Engineering from University Putra Malaysia in 2014. Since December 2009, he is a lecturer at Department of Computer Engineer, Faculty of engineering of Aden University, Yemen. He also worked on broadband networks at Wireless Communications Cluster, MIMOS

Berhad, Malaysia. Currently he is a senior lecturer at the Faculty of Computer Science and Information Technology in University of Malaya, Malaysia. His main research interests are wireless communication and networking, resource management and QoS of Internet of Things (IoT) and emerging wireless technologies standard in 4G and 5G such as WiMAX and 3GPP Long Term Evolution (LTE), LTE-Advance.

Unaizah Hanum Obaidellah completed her MSc degree from the Faculty of Computer Science and Information Technology, University of Malaya, Malaysia in 2008 and PhD degree from the School of Engineering and Informatics, University of Sussex, United Kingdom in 2012. In the 2012, she joined the Department of Artificial Intelligence, University of Malaya, as a Lecturer and became a Senior Lecturer in 2014. Her main specialization is Artificial Intelligence, with an interest in studies related to perception and pattern recognition. More specifically, her work examines how human cognitive function operates with respect to the ability to understand the perceived objects and its environment.

Erma Rahayu Mohd Faizal completed her MSc degree from the School of Engineering, University of Oita, Japan in 2007 and PhD degree from the Faculty of Engineering, University of Technology MARA, Malaysia in 2013. In the 2011, she joined the Department of Artificial Intelligence, University of Malaya, as a Lecturer and became a Senior Lecturer in 2014. Her main specialization is Artificial Intelligence, with an interest in studies related to computer vision and pattern recognition in developing a smart vehicle system.

Joan Lai P C was the Scholarship holder for her doctorates from Ministry of Higher Education and the Chancellor Scholarship. She has received a few fellowships in the Asia Pacific arena including Next Generation Leadership award. She has championed and won a few research and development awards. Her creativity and innovation has won a few Asia Pacific Information Communication Technology (ICT) awards including the prestigious ICT Oscar award APICTA awards as well as voted top 3 favorites @ Next Bank Asia. She also enjoys her works with researchers as well as bringing innovation to the marketplace working closely with industry players and facilitating Master students.