Optimum Parameters for MPEG-4 Data over Wireless Sensor Network

Farizah Yunus[1], Sharifah H. S. Ariffin[1], S. K. Syed-Yusof[1], Nor-Syahidatul N. Ismail[1], A. H. F. Abdul Hamid[1], Norsheila Fisal[1]

[1]UTM-MIMOS Center of Excellence, Universiti Teknologi Malaysia, 81310 Skudai, Malaysia
[1]farizah@fkegraduate.utm.my,

Abstract— **Nowadays, video streaming application is widely used in wired and wireless environment. Extending this application into Wireless sensor network (WSN) applications featuring low data rate transmission, low energy consumption, ease of deployment and low cost has attracted lots of attention in the research community. However, video transmission over such network is more challenging because of the large amount of bandwidth required. To cater this problem, video compression is of utmost importance to decrease the amount of bandwidth required over WSN. MPEG-4 video codec is one of the compression scheme that was identified to be suitable for WSN environment. In this paper, a simulation study for MPEG-4 video encoding scheme based on an experimental model was carried out to determine conformance with IEEE 802.15.4 requirements. The results obtained from this paper would be used as a benchmark for the configuration of the video encoding scheme for WSN applications. There are three parameters that we are concerned with in this experiment, which are quantization scale, group of picture (GOP) and frame rate (fps). The results from this simulation study shows that an optimal selection of the parameters value that enhances the video transmission over WSN.**

Keyword-Wireless Sensor Network, IEEE802.15.4 standard, MPEG-4, Quantization Scale, Group of Picture and Frame per Second.

I. INTRODUCTION

IEEE 802.15.4 is a data communication protocol standard that was designed for wireless personal area networks [1]. The standard deals with applications that require a very low power profile and low complexity at the cost of low data rate. Due to its power profile, it has become the de facto standard for WSN applications where the devices or nodes are mainly powered by a battery and required to operate for a very long time. Additionally, WSN devices are equipped with low to medium end on-board processor as well as sensor for data collection. These specifications enabled WSN devices to be mass manufactured and deployed at a low cost which open up a range of new applications.

The idea of deploying hundreds or thousands of these inexpensive nodes equipped with multimedia sensor has garnered a lot of attention in the research community. The inherent limitations of WSN’s node resources require a new paradigm to be found in order to transmit multimedia data with high traffic volume such as video over WSN. Video transmission over WSN in practical is used to enhance and complement the existing sensor network application. As such, it is crucial to keep the cost of the sensor node and its power profile low by only transmitting a highly compressed video. Therefore, video encoding process is crucial in maintaining the quality of video as well as reducing the traffic volume for transmission. Flood monitoring in remote area is an example of video transmission over WSN that require a low resolution video as well as delay tolerant.

In the literature, video transmission over wireless network mainly focuses on high data rate standards such as WiFi, Bluetooth and other technologies that will be discussed in the next section. In general, these standards are suitable for real-time application because they offer data rate in Mbps with Maximum Transmission Unit (MTU) up to 1500 bytes. However, video transmission over IEEE 802.15.4 standard is more challenging and is given a special interest due to its nature in order to provide a network that have low complexity and low cost but still capable of maintaining a good quality video in term of Peak Signal to Noise Ratio (PSNR) at the receiver side. To realize this, MPEG-4 video encoding scheme was proven to be a suitable video compression technique for video transmission over IEEE 802.15.4 standard in [2][3][4]. Nevertheless, the other researchers do not consider quantization scale, frame rate and the size of the GOP during encoding process that will heavily affect the bandwidth requirement and PSNR measurement for video quality before transmission.

This paper investigates the encoding process and the resulting video quality achieved by MPEG-4 encoding scheme before transmission and the video quality achieved after transmission over WSN for IEEE 802.15.4 standard. Using the results from the investigation, this paper would then identify the optimal parameters required to improve the result. The MPEG-4 video codec was chosen among the available video codecs because of its low bit rate feature capable of achieving as low as 64 kbit/s [5] compared to other video codecs. This simulation-based approach was carried out to determine the optimal encoding quantization scale, frame rate and the size of the GOP by evaluating the video quality under WSN requirements.

The remainder of this paper is organized as follows. Section 2 gives an overview of MPEG-4. Section 3 describes the video encoding process, and section 4 discusses the related works for video transmission over IEEE 802.11 and IEEE 802.15.4 standards. Section 5 provides the simulation results for various parameter settings. Finally, section 5 concludes the paper and the future direction of our research.

II. MPEG-4 ENCODER

The Motion Picture Experts Groups (MPEG-4) are one of the standards for the coding of moving pictures and audio for video streaming. The video samples encoding process play an important role in maintaining the quality of video streaming especially in wireless environments where the channel condition changes rapidly.

Fig. 1 shows the MPEG-4 encoding structure and component that consists of slices, macroblock and block. Each video picture or frames consist of several numbers of slices and every slice consists of macroblock. A macroblock contains a section of the luminance component, and the spatially subsampled chrominance components that carry the shape,motion ad texture information [5]. Sampling format for a macroblock is 4:2:0 where each macroblock contains 4 luminance blocks and 2 chrominance blocks. Meanwhile, each block contains 8x8 pixels and is encoded using Discrete Cosine Transform (DCT) transform [6]. Video frame formats are used to determine the individual video frame in terms of pixels. The Common Intermediate Format (CIF) video frame format has a resolution of 352x288 pixels, and Quarter CIF (QCIF) video frame format has a resolution of 176x144 pixels. The pixels will represent different color spaces. The common color space is red, green and blue (RGB) representation.

The format used for both CIF and QCIF are in the YUV pixel format. The Y component is defined as a luminance (brightness level) while U and V component are defined as hue and intensity. Since human eyes system is more sensitive to luminance component, chrominance component can be a sub-sampled without reducing the video quality. The ratio of luminance to chrominance byte sub-sampling is reduced with a sub-sampling process. The available YUV formats are YUV 4:4:4 (without sub-sampling), YUV 4:2:2, YUV 4:2:0 and YUV 4:1:1. However, YUV 4:2:0 is being used for video conferencing, digital television and modern video coding standard more extensively. Due to this popularity, most of the literature focuses on YUV 4:2:0 [6].

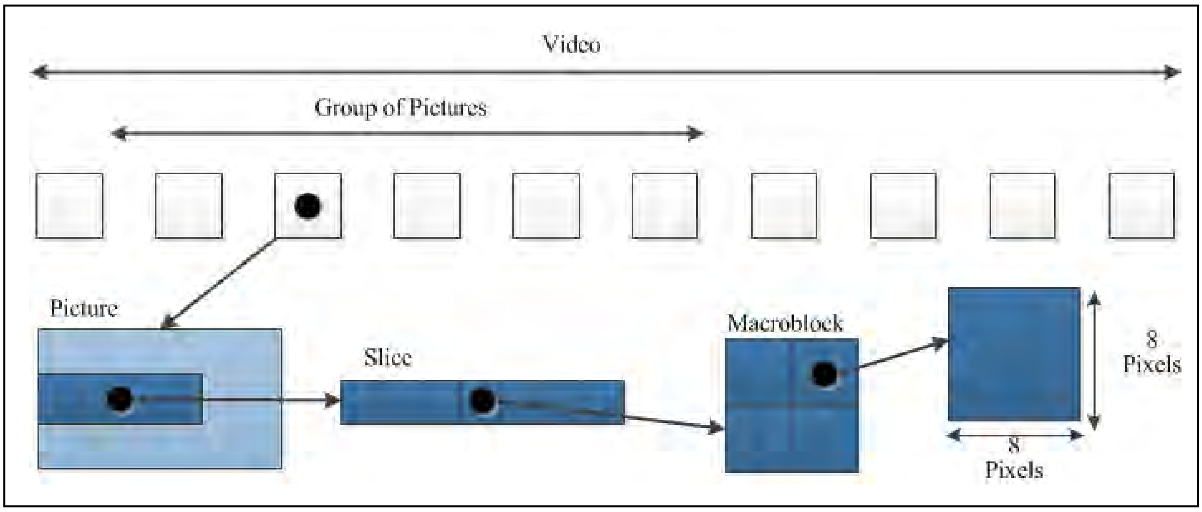

Fig. 1. MPEG-4 encoding structure and component (slices, macroblock and block)

III. VIDEO TRANSMISSION OVER WIRELESS SENSOR NETWORK

The maximum allowable data rate for IEEE 802.15.4 is only 250kbps. Therefore, the data rate for the video must be fulfilled the requirement of the network to ensure the video transmission is successfully arrived at the receiver with the reasonable and acceptable packet loss. This is due to preserve the quality of video in good and acceptable condition.

A. Data Rate for Video Transmission

Data rate is the rate at which information or data being transferred. It can be expressed in term of (amount of information) / (unit of time). For example, the maximum allowable data rate for WSN is 250kbps. Thus, the bandwidth of data or information that will be transferred over the WSN must be less than 250kbps. Since this work is focused on multimedia transmission which is video, it is important to ensure that the bandwidth for video data is sufficient for low data rate device.

To calculate the data rate of the video, we need the compression ratio as defined in equation (1). The compression ratio depends on video codec, the quality of the video and the format of the video.

$$Compression\ ratio = \frac{Uncompressed\ video\ size\,(bits/sec)}{Compressed\ video\ size\,(bits/sec)} \tag{1}$$

$$\begin{pmatrix} Uncompressed \\ video\ size \end{pmatrix} = \begin{pmatrix} resolution \times \\ depth\ of\ color\ format \times \\ frame\ rate \times \\ duration\ of\ video \end{pmatrix} \tag{2}$$

The low compression ratio would result in better video quality which is close to the original video.

Equation (2) is to determine the size of uncompressed video that consists of resolution format, depth of color format, frame rate (frame per second) and duration of the video. The resolution can be either QCIF or CIF format. The depth color format depends on the video color format which is 8 bits/pixel. Thus, we can estimate the data rate of the video before transmission over WSN based on several parameters as mentioned above. Moreover, the bit rate of the video can be calculated using equation (3).

$$Data\ rate = \frac{\begin{matrix} resolution \times depth\ of\ color\ format \times \\ frame\ rate \end{matrix}}{compression\ ratio} \tag{3}$$

B. Related Works

There are several works that have used MPEG-4 as a video codec for high and low rate video transmission. This section explains the related works that utilize MPEG-4 for video transmission over wireless network. Most of the work related to the MPEG-4 codec that have been done on video transmission are intended for the IEEE 802.11 standard that support high data rate transmission. Whereas only several works can be found on MPEG-4 video transmission for the IEEE802.15.4 standard that support low bandwidth.

In [7], the paper introduced frame dropping module. It will drop a frame with low importance such as B and P frame if bandwidth estimation is lower than the transmission rate. The scheme would drop B frame first and then continue with P frame if the condition persist. In the experiments, they used a video with a frame rate of 30 frames per second (fps) with 12 group of picture (GOP) that consists of I, P and B frames. Since IEEE 802.11 has a large bandwidth, this mechanism works well given the condition where the bandwidth estimation is lower than the transmission rate has a low probability of occurring. In case of WSN network where the probability is higher due to the massively deployed nodes and its close proximity to each other, the number of packets dropped would be high and would cause the video quality to be severely degraded.

Z. Ma in [8] proposed a cross layer design between the application layer and data link layer for retransmission of the packet loss for IEEE 802.11. At the data link layer, priority based ARQ (P-ARQ) algorithm is used to decide whether to retransmit the packet or not. The decision is based on class priority and round trip time for the packet loss which is less than its deadline. This project used foreman sample that is encoded in QCIF format with 15 frames per second. However, the delay and the energy consumption induced by P-ARQ would make it unsuitable for low power low rate application utilizing IEEE 802.15.4.

A cross layer design between the application layer and MAC layer in IEEE 802.11 is proposed in [9]. MAC layer will discard frame with low priority if the higher priority frame is lost or met the deadline that have already been assigned at the application layer. The Queuing at the MAC layer would have to be adjusted by dropping all the dependency frames if there are frames that were lost. This approach is good to maintain the frame rate received at an acceptable delay but MAC layer also should monitor maximum delay, and it will become more complex if it's not implemented correctly.

The authors in [2] proved that MPEG-4 can also be transmitted at a low rate and low bandwidth medium such as IEEE 802.15.4 standard. They proposed Contention Free Period (CFP) which is Guaranteed Time Slot (GTS) in IEEE 802.15.4 to transmit the video. However, by using only CFP period, the result reveals that the transmission rate can only go up to 7 frames per second (54kbps) due to the delay. This work then was enhanced in [3] with the introduction of a Cross Layer Multimedia Guaranteed Time Slot (CL-MGTS). Here, they proposed a protocol that controls the optimum MAC gap size, video transmission rate and minimum power consumption. The cross layer decision involves Contention Access Period (CAP) and MGTS gap, the number of CAP and GTS slot, transmission rate and the suitable super frame duration (BO/SO parameters). The hardware device deployment using commercial WPAN hardware for CL-MGTS protocol is described in [4]. The application focuses on agriculture sector with cluster topology.

Even though all the above works used MPEG-4 codec for video transmission over wireless network, they did not explain in details on the parameters that need to be considered, which will affect the bandwidth requirement and PSNR measurement for video quality. Thus, this paper will investigate the several important parameters that need to be taken into consideration for video transmission over WSN in order to ensure the video quality is not compromised during the transmission.

IV. SYSTEM MODEL

A. *Video Encoding Process*

The video samples will be encoded and decoded in offline mode due to the complexity of encoding and decoding process where the results produced would be used for simulating the wireless streaming scenario on the simulator. The flow of the video encoding and testing model is shown in Fig. 2.

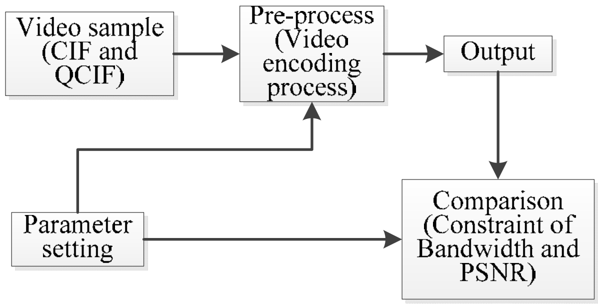

Fig. 2. Flow of video encoding and testing model

Naturally, the encoding process involves video samples. Evalvid tool software is used to encode the video samples into the MPEG-4 format process with a custom parameter setting. As discussed in section 2, the video samples that are used in this simulation can be either Common Intermediate Format (CIF) or Quarter CIF (QCIF) video frame format. The CIF has a resolution of 352x288 pixels, while QCIF has a resolution of 176x144 pixels.

In this work, the parameter settings are divided into three categories, which are quantization scale, frame rate (fps) and group of picture (GOP). All of these three parameters will affect the size of the final video. Since WSN has a stringent constraints and posed a significant challenges compared to others wireless network for video transmission, the determination of video size is important to ensure that the video can be sent over WSN with a good video quality and with the Peak Signal to Noise Ratio (PSNR) in the range of 31-37.

Quantization scale is the parameter responsible for the “lossiness” in the MPEG-4 encoding scheme. It basically determined the output of the DCT in video compression. By having a lower value of quantization scale, the compression ratio would be low and the video quality would remain close to the original video. However, a low value of quantization scale would result in a video size close to the original video. Increasing the quantization scale on the other hand would decrease the compressed video size as well as degrading the video quality. A highly compressed video would produce an artifact because of the missing information during the encoding process. The tradeoff between video quality and compressed video size must be balanced to achieve an acceptable video quality with an acceptable size for video transmission over WSN.

Frame rate (fps) is important because it will show the smoothness of the image transitions in the video and used to determine the quality of a video. The common values of frame rate for MPEG-4 video are between 25 to 30 fps. Frame rate also directly determined the size of the compressed video and thus is important to make sure the data rate required met the WSN specification.

The MPEG-4 video can be divided into several numbers of video sequences that are called group of picture (GOP) [10]. MPEG-4 video generates three different types of frames that consists of Intra-coded frame (I frame), predicted frame (P frame) and bidirectional frame (B frame). I frames are encoded independently and the frame size is larger compared to other types of frames. This is because I frame contains most of the important video information, and the GOP is useless if I frames are lost. While P frames are encoded using predictions from the previous I or P frames. B frames are encoded using predictions from previous and next I and P frames. The choice of GOP structure is important because it will give effect on the frame size and file size. Additionally, it also will give impact to the MPEG video streaming in term of network bit rate and video quality [11].

There are several steps to find the optimal value for three parameters that are taken into consideration to ensure that the video can be transferred over WSN in good quality. Firstly, the video will be encoded without quantization scale with a frame rate of 30 fps and 30 video frame sequences per group of picture (GOP).

After that, the bit rate of the encoded video is checked to ensure that the bandwidth requirement does not exceed 250 kbps in order to conform with WSN specification. If the compressed video data rate exceeded the bandwidth allocation, the quantization scale would be increased with the frame rate and the GOP size remains the same. If the quantization scale reached its maximum value and the compressed video data rate still exceeded the bandwidth allocation, the frame rate would be reduced gradually while maintaining the value of the GOP. This step will be repeated until the compressed video data rate met the bandwidth requirement. When the compressed video data rate met the bandwidth requirement, the PSNR will be checked to ensure the video is in a good quality before it is being transmitted over WSN. However, to increase the PSNR value in order to get better video quality, the GOP value will be decreased gradually while maintaining the bandwidth requirement. This step will be repeated until the optimal value for all three parameters is achieved where the compressed video data rate is within the constraints of bandwidth allocation and PSNR.

B. Simulation Using NS-2

In this simulation study, we used Evalvid [12] simulation software which is a complete framework and toolset for the evaluation of MPEG-4 video quality transmitted over the real or simulated communication network. This toolset used video traffic traces, which provide the relevant video characteristic as an input for video transmission. Fig. 3 shows the Evalvid architecture which is divided into three stages namely pre-process, network simulation and post process.

Pre-process stage will perform the video encoding mechanism to generate video trace file that can be used for network simulation. The original video source (*.yuv file), will be encoded using ffmpeg.exe to produce m4v video. This output then will be fed to MP4Box.exe where it will be converted into mp4 video (*.mp4 file). By employing mp4trace.exe using the mp4 video file as its input, trace file named st_video.txt will be generated. The generated trace file contains the information regarding the mp4 video frames such as its sequence number, type, size in bits, total number of fragmented packets per frame and its encoding time as shown in Figure 4. This information then will be used by NS-2 simulator for the simulation process.

In the simulation stage, the trace file that was generated in the pre-process stage is used as its input. The trace file is read and stored into video.dat file. This file then will be used to generate sd_video.txt file for the sender side and rd_video.txt file for the receiver side.

Post-process is the final stage where the video will be reconstructed based on the result of the video transmission simulation. For this process, the reconstructed video will be compared to the original video to calculate PSNR for evaluation of end-to-end video quality.

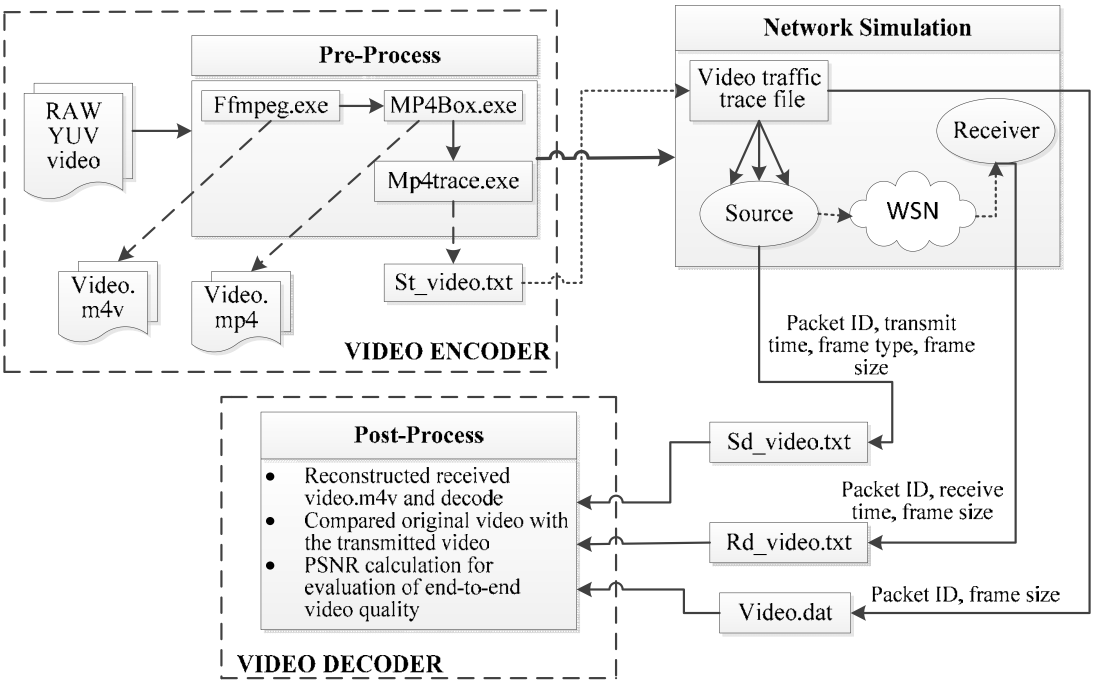

Fig. 3. An overview of NS-2 simulation architecture for video transmission

No.	Frame type	Frame size	No. of segment	Encoding time
1	I	4901	5	0.000
2	P	59	1	0.034
3	P	53	1	0.068
4	P	64	1	0.101
5	P	157	1	0.134
6	I	4892	5	0.168
7	P	60	1	0.201
8	P	162	1	0.234
9	P	176	1	0.268
10	P	169	1	0.301
11	I	4927	5	0.334
12	P	132	1	0.367
13	P	174	1	0.401
14	P	183	1	0.434
15	P	339	1	0.468
16	I	4931	5	0.501
17	P	368	1	0.534
18	P	665	1	0.567
19	P	706	1	0.601
20	P	763	1	0.634
21	I	4907	5	0.667
22	P	233	1	0.701
23	P	336	1	0.734
24	P	334	1	0.767
25	P	432	1	0.801
26	I	4881	5	0.834

Fig. 4. Example generated trace file of st_video.txt

V. RESULTS AND DISCUSSION

A. Configuration of Parameters Encoding

In this section, we study the effects of three parameter configurations on video quality. The parameters are quantization scale, group of picture (GOP) and frame per second (fps). There is only one parameter that is being varied while the other two parameters are fixed at one time.

In this simulation study, three different video sequences with 4:2:0 sub-sampling for both CIF and QCIF video format are used. The video sequences are named Akiyo, Foreman and Mobile as shown in the Fig. 5, Fig. 6 and Fig. 7.

Fig. 5. Akiyo video sequence

Fig. 6. Foreman video sequence

Fig. 7. Mobile video sequence

These video sequences represent low, medium and high motion and scene complexity respectively. In the Akiyo video sequences, female moderator reading news and only her lips and eyes are moving. This will be considered as low video motion. In Foreman video sequences, the fast body movement and quick scene changes which are more complex than low video motion. This will be considered as medium video motion. While in Mobile video sequences, there are toy train moving horizontally and calendar moving vertically in the background and is considered as high video motion.

Quantization is important during encode the video to ensure the acceptable bandwidth requirement to transfer the video over WSN. When the value of the quantization scale increased, the bandwidth requirement to transmit the video will decrease. However, the quality of the video is also decreased by decreasing the value of PSNR. This is due to the quantization processes being a lossy step of texture coding and the information that was discarded during the quantization process cannot be restored.

Each GOP will encode a single I frame and many P and B frames. However, in this simulation, we only used two frames, which are I and P frames. B frames contain bidirectional motion vectors and transform coefficients. In most coding standards, B frames are not being used as a reference to make further predictions to avoid a growing propagation error and thus can be safely dropped without severely degrading the video quality. In this simulation, we need the GOP size to be large in order to reduce the bandwidth requirement during video transmission. This is due to the maximum data rate for WSN is not more than 250 kbps.

Frame rate (measured in fps) is important in determining the smoothness of the video during the playback. Frame rate does not affect the value of PSNR directly. However, the bandwidth requirement will increase when the value of frame rate increases.

The video quality can be measured subjectively and objectively. For an objective measurement, the original video signal will be compared with the reconstructed video signal. The comparison is derived using the metric of mean square error (MSE) and peak signal to noise ratio (PSNR). MSE can be defined as square differences between luminance values of original or reference video (X) and the encoded or impaired video (Y). MSE and PSNR can be computed using equation (4) and (5).

$$MSE = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M=1} [X(i,j) - Y(i,j)]^2 \tag{4}$$

$$PSNR = 10 log_{10} \frac{L^2}{MSE} \tag{5}$$

The sequences of the video frames are composed by NxM pixels. For example, in QCIF format N equal to 176 and M equal to 144. Where the original source frames *(X (i,j))* and reconstructed frame *(Y(i,j))* contains N and M pixels.

Using this information, MSE for each frame will be converted to PSNR in a decibel (dB) by mapping the 8 bits original signal. L reflects the range of values that a pixel can take. Thus, in general cases, L can be computed with the equation (6).

$$L = 2^8 - 1 = 255 \tag{6}$$

When the differences between the original source frame (X(i,j)) and reconstructed frame (Y(i,j)) is large, the value of PSNR will be small. Hence, a higher value of PSNR reflects a better video quality. The MSE and PSNR have been used extensively because of their simplicity [13].

While in a subjective assessment, human ‘grading scale’ is used to evaluate the perceived video quality. Mean Opinion Score (MOS) is one of the techniques for subjective assessment, where a human quality impression is scaled from 1 (bad) to 5 (excellent). Table 1 shows the prediction of video quality based on the Mean Opinion Score (MOS) score. However, subjective assessment is rarely used because of its time consuming, high man-power equipment and the special equipment involved [14].

TABLE I
MOS Conversion from PSNR

Peak Signal to Noise Ratio (PSNR)	**Mean Opinion Score (MOS)**
> 37	5 (Excellent)
31-37	4 (Good)
25-31	3 (Fair)
20-25	2 (Poor)
< 20	1 (Bad)

B. Simulation Results

This section presents the results and analysis for video encoding process before transmission over WSN and results analysis after video transmission over WSN.

Results for video encoding process is important to ensure that the video quality before transmission over WSN meets the bandwidth requirement, and the value of PSNR has remained high. This is due to the video quality being degraded after the transmission over WSN.

1) Results for Optimum Value Parameters: This section presents the results analysis for video encoding process before transmission over WSN.

Result before video transmission over WSN is important to ensure that the video quality after encoding process meets the bandwidth requirement, and the value of PSNR is high. This is due to the video quality being degraded after the transmission over WSN.

The video encoding process is performed using Evalvid tools set as discussed in section IV. The three parameters (Qscale, fps and GOP) were taken into consideration during the video encoding process. Fig. 8 and Fig. 9 shows the effects of quantization scale to bandwidth requirement and peak signal to noise ratio (PSNR) for all video samples respectively.

Fig. 8 shows the effects of quantization scale to the bandwidth requirements for all video samples. From the result, it can be summarized that the relationship between quantization scales and bandwidth requirement is inversely proportional. A higher value quantization scale will result in a smaller size of the encoded video frames and therefore, decreases the bandwidth requirement of video transmission. In other words, a higher value quantization scale is desirable in order to meet the WSN bandwidth limitation. However, we observed that three video samples, foreman_cif, mobile_qcif and mobile_cif require more bandwidth than WSN can provided even when the value of quantization scale is as high as 11. Therefore, the suitable samples that will be considered in this work, which require bandwidth below than 250kbps after quantization scales are akiyo_qcif, akiyo_cif and foreman_qcif.

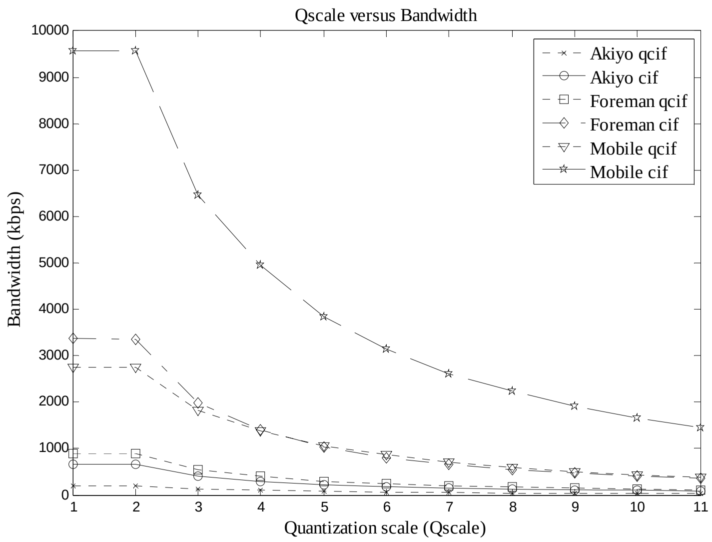

Fig. 8. Quantization scale versus bandwidth requirement for all video samples

Based on the simulation result as illustrated in Fig. 9, it has been proven that the video quality (PSNR) is degraded when the value of quantization scale is increased. This means that the relationship between quantization scale (Qscale) and PSNR is also inversely proportional. This is due to the fact that a high quantization scale causes more information to be lost, resulting in a low video quality (PSNR).

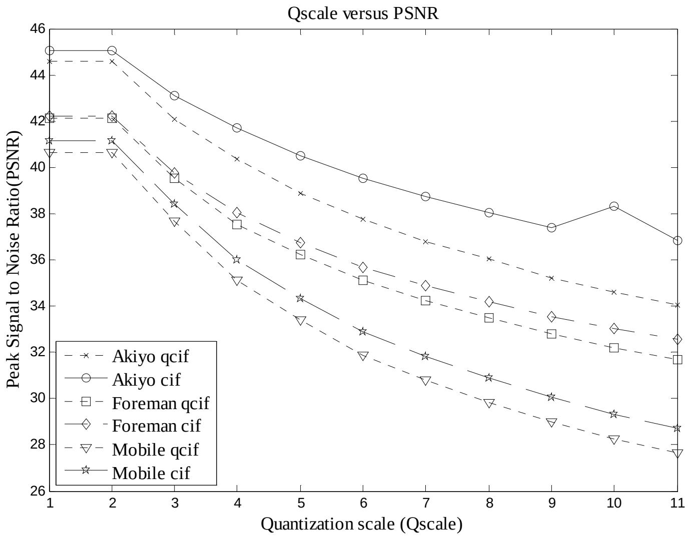

Fig. 9. Quantization scale versus Peak Signal to Noise Ratio (PSNR) for all video samples

TABLE II
Optimum Value Parameters for All Video Samples

Type of motion	**QScale**	**Frame per Second (fps)**	**Group of Picture (GOP)**	**Bandwidth**	**Peak Signal to Noise Ratio (PSNR)**	**MOS**
Akiyo (qcif)	2	30	20	226	44.67	5
Akiyo (cif)	4	25	30	240	41.72	5
Foreman (qcif)	4	15	10	229	37.83	5
Foreman (cif)	4	5	20	242	38.06	5
Mobile (qcif)	4	5	15	240	35.26	4
Mobile (cif)	11	5	30	243	28.70	3

Table II shows the optimum value for three parameters that can be obtained when the encoded video meet the requirement of bandwidth and PSNR value. It is observed that when the value of PSNR is in between 31 and 37 as presented in Table I, the bandwidth requirement is not more than 250 kps, and the video quality in term of PSNR is good.

However, for the bandwidth requirement to be low, the frame rate of the video has to be low and the size of the GOP has to be large. This is because, when the total number of I frame per seconds decreases with a large GOP size, the bandwidth requirement would be reduced due to the fact that I frame is much larger than P frame. Therefore, the value of PSNR will decrease when the quantization scale and the size of the GOP are increased.

2) Results for MPEG-4 Video Transmission over WSN: Network Simulator-2 (NS-2) was used to simulate the video transmission over WSN using Real-Time Load Distribution (RTLD) [15] routing protocol. The simulation parameters are shown in Table III using 802.15.4 MAC and physical layer with default power transmission (1mW).

A typical setup of one traffic configuration which is video traffic between one source node and a sink node in WSN application is used. The video traffic configuration is based on optimum value parameters as shown in Table II. In this simulation works, 4, 9, 16 and 25 nodes are distributed in a region with the grid topology as shown in Fig. 10. The payload used in the traffic is the video data with user datagram protocol (UDP) as the transport protocol. Hence, it is unreliable because there is no retransmission for any packet loss.

Energy consumption, delivery ratio, packet loss and video quality (PSNR) are the metrics used to analyze the performance of video transmission over WSN. Delivery ratio is defined as the ratio of packets received at the sink node to the total packets sent from the source node in the network and the energy consumption is defined as the energy consumed in each sensor node during the simulation task.

TABLE III
Simulation Parameters

Parameter	IEEE 802.15.4
Propagation model	Shadowing
Path loss exponent	2.5
Physical type	Phy/WirelessPhy/802.15.4
MAC type	Mac/802.15.4
CS threshold	1.10765e-11
RX threshold	1.10765e-11
Frequency	2.4Ghz
Initial energy	3.6 Joule
Power transmission	1mW
End-to-end deadline	100s
Traffic	Video CIF format (352x288 resolution) Video QCIF format (176x144 resolution)

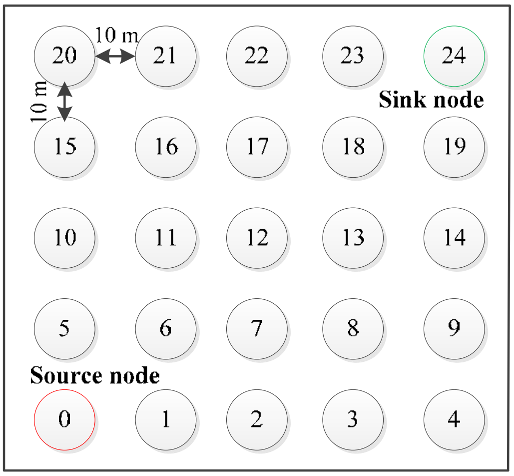

Fig. 10. Network Grid Topology

Figure 11 demonstrates the performance of video transmission in term of energy consumption for all video samples. We observed that the energy consumption increased when the number of hops increased. The same scenario happens when normal data were sent over WSN. But the graph shows that the energy consumption for normal data is higher than the energy consumption for Akiyo_qcif and Akiyo_cif video samples even when the size of the video frame is larger than the normal data where the delivery ratio for normal data is higher than the video frame.

This phenomenon is observed due to the frequent frame lost during video transmission. Since there are no packet retransmissions, the energy consumed is limited to the point where the packet is lost. However, the overall observation reports that the video frame transmission consumed more energy compared to normal data because of the large number of the video frames as shown in Fig. 11. In this simulation, normal data is defined as a constant bit rate (CBR) traffic.

Packets delivery ratio also decreased when the number of hops increased as shown in Fig. 12. The packet delivery ratio decreases because there is no packet retransmission algorithm when the packet is loss. For that reason, the pattern of packet loss is also the same where the packet loss increases as the number of hops increases as shown in Fig. 13.

Even though the packet delivery ratio is low, the value of PSNR is still high and considered acceptable, which is in between 31 and 37 for akiyo_qcif and akiyo_cif as presented in Fig. 14. This is because most of the packets that are lost are P frame not I frame. Thus, the video quality remains good and can be playback at the receiver side. The loss of I frame will result in an invaluable video at the receiver even when the P frame is received successfully. Therefore, more preservation needs to be given to I frame compared to others frame.

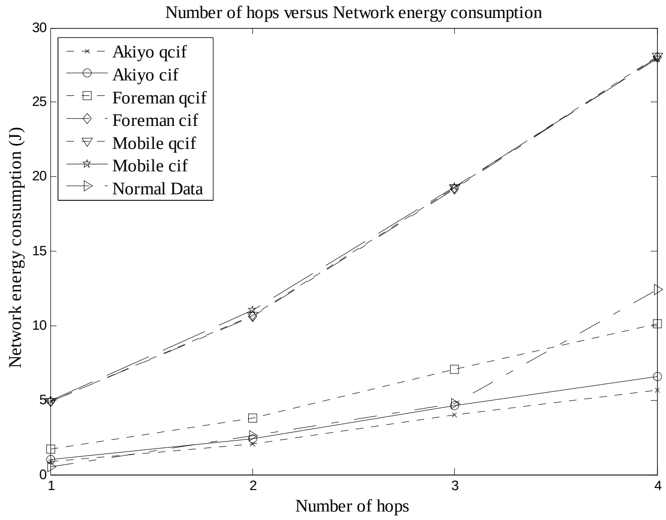

Fig. 11. Number of hops versus network energy consumption for all video samples

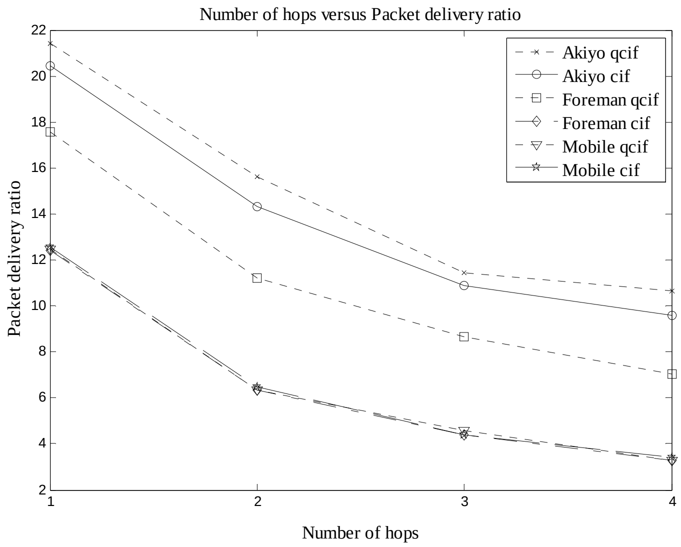

Fig. 12. Number of hops versus packet delivery ratio for all video samples

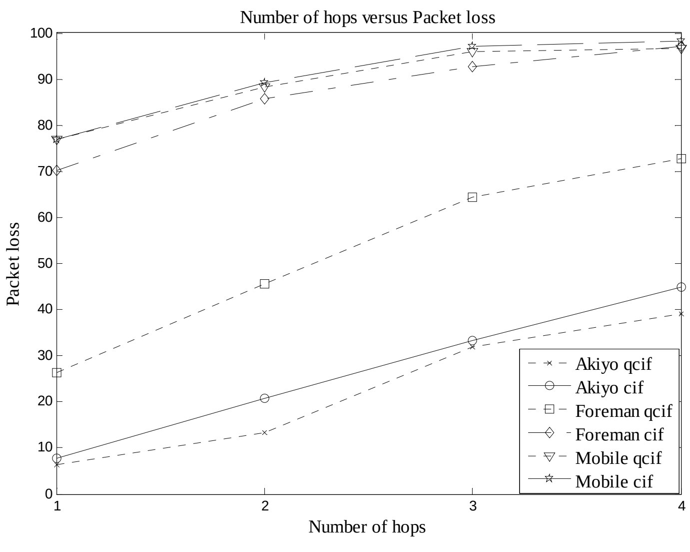

Fig. 13. Number of hops versus packet loss for all video samples

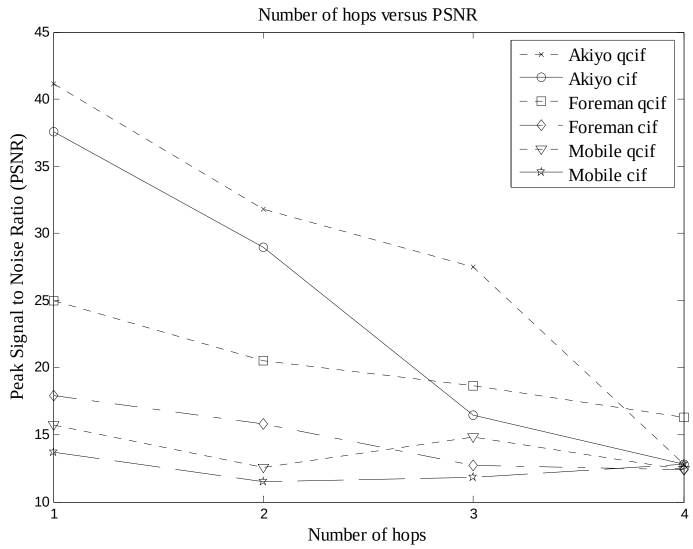

Fig. 14. Number of hops versus peak signal to noise ratio (PSNR) for all video samples

VI. CONCLUSION AND FUTURE WORK

The simulation study for MPEG-4 video encoding and experimental model was carried out to determine the best possible value for three parameters in order to tailor with the environment of wireless sensor network (WSN). The three parameters that we are concerned with in this experiment are quantization scale (Qscale), group of picture (GOP) and frame per second (fps). These parameters have been selected because of their importance in determining video quality (PSNR), and bandwidth requirement of the encoded video. The results revealed that the three parameters were crucial in maintaining the video quality in order to maintain the bandwidth requirement to be less than 250kbps.

It is also observed that when the delivery ratio is low, the video quality (PSNR) is not severely degraded. This is because most of the packets lost during video transmission or P frame, and not I frame. In this work, only low motion for cif and qcif video samples and medium motion of qcif video sample were used in order to ensure the video quality to be within an acceptable and good condition before its being transmitted over WSN. The other three samples which are medium motion of cif video samples and high motion for cif and qcif video samples are not suitable to be used in this work because of its high bandwidth requirement even after being encoded.

For the future work, a more intelligent encoding process is required in order to choose the best possible configuration of the three parameters. Particle swarm optimization (PSO) method will be used instead of trial and error method that was used in this work. The objective function will be computed from the previous data.

Furthermore, the algorithm at the transport layer for reliability mechanism is crucial to ensure a high packet delivery ratio with reasonable packet loss to increase the video quality (PSNR). The proposed transport protocol also needs to take energy consumption into consideration due to WSN low profile requirement. Hence, our proposed transport protocol algorithm also needs to allow maximum network lifetime with multimedia transmission, which is a high power consuming task.

ACKNOWLEDGMENT

The authors would like to thank all those who contributed toward making this research successful. The authors wish to express their gratitude to the Ministry of Higher Education (MOHE), Malaysia Research Management Center (RMC) for the sponsorship, and Telematic Research Group (TRG), Universiti Teknologi Malaysia for the financial support and advice on this project. (Vot number Q.J130000.2623.08J19).

REFERENCES

[1] J. Adams, "An introduction to IEEE STD 802.15. 4," *IEEE Aerospace Conference, 2006*, pp. 1–8, 2006.
[2] A.-J. Garcia-Sanchez, F. Garcia-Sanchez, and J. Garcia-Haro, "Feasibility Study of MPEG-4 Transmission on IEEE 802.15.4 Networks," *2008 IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, pp. 397–403, Oct. 2008.
[3] A.-J. Garcia-Sanchez, F. Garcia-Sanchez, J. Garcia-Haro, and F. Losilla, "A Cross-layer Solution for Enabling Real-time Video Transmission over IEEE 802.15.4 Networks," *Multimedia Tools and Applications*, vol. 51, no. 3, pp. 1069–1104, Feb. 2010.
[4] A.-J. Garcia-Sanchez, F. Garcia-Sanchez, and J. Garcia-Haro, "Wireless sensor network deployment for integrating video-surveillance and data-monitoring in precision agriculture over distributed crops," *Computers and Electronics in Agriculture*, vol. 75, no. 2, pp. 288–303, Feb. 2011.
[5] T. Ebrahimi and C. Horne, "MPEG-4 natural video coding – An overview," *Signal Processing: Image Communication*, vol. 15, no. 4–5, pp. 365–385, Jan. 2000.

[6] Iain E. G. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia*. 2003.

[7] P. Shin and K. Chung, “A Cross-Layer Based Rate Control Scheme for MPEG-4 Video Transmission by Using Efficient Bandwidth Estimation in IEEE 802 . 11e,” *International conference Information Networking ICOIN*.

[8] Z. Ma, J. Zhou, and T. Wu, “A Cross-layer QoS Scheme for MPEG-4 Streams,” in *IEEE International Conference on Wireless Communications, Networking and Information Security (WCNIS), 2010*, 2010, no. 609731, pp. 392–396.

[9] Y. Xiao, X. Du, F. Hu, and J. Zhang, “A Cross-layer Approach for Frame Transmissions of MPEG-4 over the IEEE 802 . 11e Wireless Local Area Networks,” in *IEEE Wireless Communications and Networking Conference, 2008. WCNC 2008.*, 2008, pp. 1728–1733.

[10] S. Moradi and V. W. S. Wong, “Technique to Improve MPEG-4 Traffic Schedulers in IEEE 802.15.3 WPANs,” in *IEEE International Conference on Communications, ICC 2007*, 2007, pp. 3782–3786.

[11] Á. Huszák and S. Imre, “Analysing GOP Structure and Packet Loss Effects on Error Propagation in MPEG–4 Video Streams,” *Proceeding of the 4th International Symposium on Communications, Control and Signal Processing, ISCCSP 2010*, no. March, pp. 3–5, 2010.

[12] J. Klaue., “Evalvid.” [Online]. Available: http://www.tkn.tu-berlin.de/research/evalvid/fw.html.

[13] “Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment,” 2003.

[14] F. Gebali, *Analysis of Computer and Communication Networks*. Springer, 2008, pp. 1–684.

[15] A. Ahmed and N. Fisal, “Secure Real-time Routing Protocol with Load Distribution in Wireless Sensor Networks,” *Security and Communication Networks*, no. May 2010, pp. 839–859, 2011.