Encoding and communication energy consumption trade-off in H.264/AVC based video sensor network

Bambang A.B. Sarif[1], Mahsa T. Pourazad[1,2], Panos Nasiopoulos[1] and Victor C.M. Leung[1]
[1]University of British Columbia, Vancouver, Canada
[2]TELUS Communications Inc., Canada
{bambangs, pourazad, panos, vleung}@ece.ubc.ca

Abstract— **Video sensor networks (VSN) offer an interesting platform for a distributed and flexible surveillance system. In such a system, video compression and wireless transmission are the major operations on each video node. For a battery-powered wireless video sensor, it is essential to maximize the power efficiency of these two operations. Currently, H.264/AVC is the most widely used ITU-T and ISO/IEC video coding standard. Previous works on determining the trade-off between compression and transmission that minimizes energy consumption consider oversimplified coding configurations, thus not taking full advantage of the flexibility and advanced features of H.264/AVC. Choosing the right configuration and setting parameters that lead to optimal encoding performance is of prime importance for video sensor network (VSN) applications, especially since VSN is constrained in terms of bandwidth and energy resources. This paper studies the relationship between the picture quality, the transmission rate, and the complexity of the encoder to expound the energy consumption trade-off between encoding and transmission in VSN. The results of our study can be used as guidelines in optimizing the overall power consumption of a VSN system as it detailed in the paper.**

Keywords-component; Wireless sensor network, video sensor network, H.264/AVC, energy consumption.

I. INTRODUCTION

A wireless sensor network (WSN) is a collection of networked sensor nodes that are distributed spatially to monitor physical attributes of the monitored environment. Considering that visual information can significantly improve the perceived information gathered from the sensed environment, there is a growing interest in incorporating video applications and transmissions over WSN [1–3]. Wireless video sensor network (VSN) has the potential to improve the ability to develop user-centric surveillance applications to monitor and prevent harmful events [4][5]. Unlike the conventional WSNs, VSNs require a large amount of resources for encoding and transmitting the video data. Therefore, maximizing the power efficiency of coding and transmission operations in VSNs is very important.

Currently, H.264/AVC is the most widely used ITU-T and ISO/IEC advanced video coding standard [6][7]. A number of published research works on H.264/AVC's complexity can be found in the literature [8][9]. Unfortunately, the focus of these studies is to obtain the optimal configuration of the encoder without considering the total energy consumed by the encoding/transmission process. The researchers in [10][11] compare the total energy consumption of some video codecs including H.264/AVC. However, in these studies only one or two configurations of H.264/AVC are used, which limits the capabilities and flexibilities of the encoder.

The performance of H.264/AVC in terms of quality, bitrate, and complexity, is determined by a large number of encoding parameters. In a resource constrained environment such as a VSN, it is very important to choose the right configuration and settings parameters that lead to the optimal coding performance. Therefore, in this paper, we investigate the relationship between the transmission rate, and the complexity of the H.264/AVC encoder to expound the energy consumption trade-off between encoding and transmission in VSN. The rest of the paper is organized as follows. Section II describes H.264/AVC encoding tools and the trade-off between its complexity and performance. Analysis of total energy consumption for encoding and transmitting information in VSN is given in Section III. Conclusions are drawn in Section IV.

II. H.264/AVC COMPLEXITY AND PERFORMANCE

While in traditional WSN, most of energy is consumed for data transmission, in VSN, a significant portion of total energy of sensor nodes is consumed by video compression algorithms. It is thus essential to have an optimal encoding performance to minimize the total energy consumption in a VSN. In order to explore the trade-offs between bitrate, quality, and computational complexity, it is necessary to perform evaluations using a wide range of encoding parameters and input sequences.

H.264/AVC is a block-based hybrid video coding that reduces the bit rate generated by the source encoder by exploiting source statistics. For this purpose intra-frame compression techniques are used to reduce redundancy within one frame, while inter-frame compression (contains motion estimation) is used to exploit redundancies among subsequent frames. In terms of complexity, inter-frame encoding is more involved than intra-frame encoding. However inter-frame coded frames have smaller bitrate than the intra-coded ones. Some initial observations of the encoder's performance led us to a low complexity encoding configuration as shown in TABLE I. Among the listed

TABLE I. ENCODING PARAMETERS FOR THE LOW COMPLEXITY CONFIGURATION

Parameters	Values
Number of reference frame	1
Smallest motion compensation block size	8x8
Entropy coding	CAVLC
B Slices	None
Subpel motion compensation	Disabled
Rate Distortion Optimization	Off
Rate Control	Off
Deblocking filter	Disabled
Motion estimation search range	2
GOP (group of pictures) size	1
QP (Quantization Parameter)	34

parameters in TABLE I, in our study we investigate the effect of the following parameters: group of pictures (GOP) size, search range (SR), and quantization parameter (QP). GOP size determines the number successive pictures within a coded video stream. Increasing GOP size will increase the number of inter-frame coded frames. The search range determines the size of searching area in the reference frame to find the best match to be used for inter prediction. Increasing the search range SR may result in better compression performance. However it comes with the price of increasing the complexity of the encoder. The quantization parameter regulates how much spatial detail is saved, i.e., the quality of the video. When QP is very small, almost all of that detail is retained, but at the price of higher complexity and bitrate.

The H.264/AVC software, JM version 18.2 is used in our experiments. Two representative test videos are selected from the data set provided by MPEG for HEVC call of proposal [13]. They are the BQMall (832x480 pixels, frame rate=60) and Traffic (2560x1600 pixels, frame rate = 30). Furthermore, we downsample these sequences to produce 15 frames per second (fps) videos (total number of frames=75) and resize them to obtain the following resolutions: 4CIF (4 times of common intermediate format, 704x576 pixels) and CIF (352x288 pixels).

The complexity of the H.264/AVC video encoder depends on the profile, the encoding options used, and the targeted application. In some existing studies, the encoding complexity is measured based on the encoder's computation time which may not be accurate since it widely depends on the device architecture and the optimization level of the algorithms involved (e.g., if GPU or video processors are used) and whether CPU is involved with other processes than encoding. To have an accurate measure of the encoding complexity we use the instruction level profiler *iprof* [12], which provides us with the total number of basic instruction counts. Fig. 1 and Fig. 2 illustrate the complexity measure using encoding time and the instruction counts (IC) provided by *iprof* respectively. Note that when GOP>1 the inter-prediction module in the encoder uses the value of SR=2 as stated in TABLE I. Fig. 1 and Fig. 2 are obtained using the BQmall 15 fps CIF video as the input sequence.

It can be seen from Fig. 2 that the increase in IC resulted from increasing the GOP sizes is consistent regardless of the QP values used. It is also observed that the IC value for when GOP size=1 is the smallest among them. This is

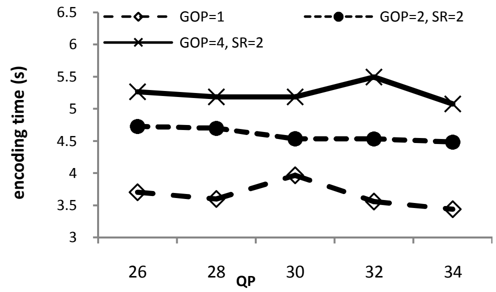

Figure 1. Encoding time for different QP values and GOP sizes.

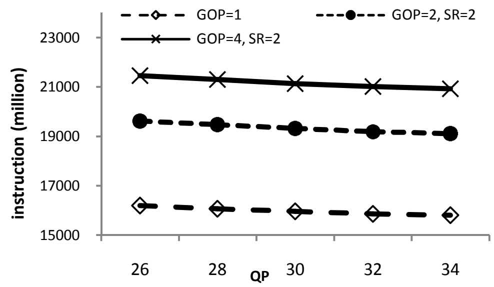

Figure 2. Instruction counts for different QP values and GOP sizes.

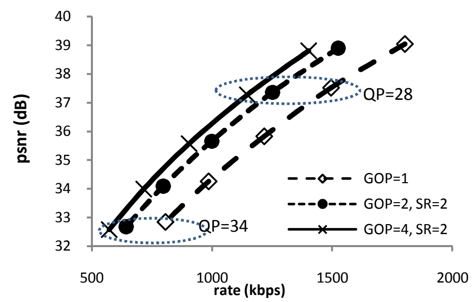

Figure 3. Rate distortion curves for different QP values and GOP sizes

expected since when GOP=1, the motion estimation (ME) algorithm in the encoder is disabled.

Fig. 3 shows the bitrate and the corresponding peak signal to noise ratio (PSNR) of BQmall 15fps CIF video sequence using the same configuration as the one used in Fig. 2. It can be seen here that the bitrate decreases by the increase of GOP size. For example, for QP equal to 34, the bitrate of the bitstream is around 800 kbps (GOP=1) and 570 kbps (GOP=4). The average PSNR of these two videos are 32.8 dB (GOP=1) and 32.6 dB (GOP=4). However, from Fig. 2 it can be observed that the encoder needs around 16 billion of instructions for GOP=1 and 22 billion of instructions for GOP=4 to encode the video. From Fig. 2 and Fig. 3, we can see the trade-off between encoder's complexity and transmission rate. The encoder's complexity can be used to estimate the encoder's power consumption. Thus, the trade-off between encoding and communication energy consumption can be calculated as well.

The JM 18.2 software provides a number of motion estimation algorithms that can be used in the encoder. These algorithms differ in the way of exploring the search space of reference block candidates whose size is determined by the search range. The full search algorithm proceeds by checking all candidates sequentially while some algorithms use heuristics to reduce the search space. The algorithms that use

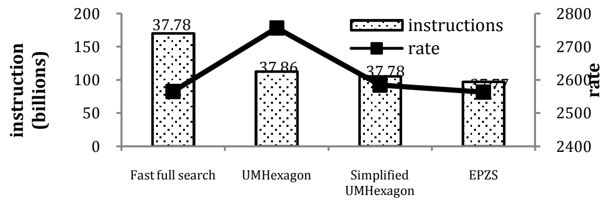

Figure 4. Performance of different ME algorithms

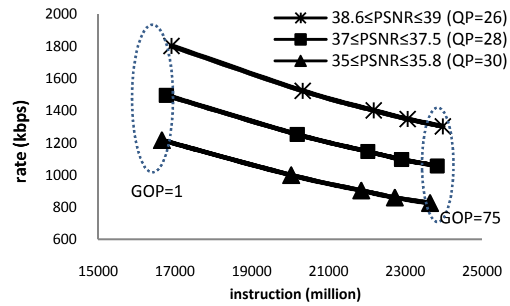

Figure 5. Bitrate and instruction counts trade-off, BQmall 15 fps CIF video, GOP={1,2,4,8,75} and SR=2

heuristics are less complex and faster than the full search, but they may not found the best reference blocks for the inter-prediction. The heuristic based algorithms proceed by selecting the best motion vector predictor from a set of potential predictor, i.e., blocks that are covered in a hexagon (UMHexagon and Simplified UMHexagon) or diamond (EPZS) shape. The algorithms have the option for early termination, if some rules are satisfied, or continue with refinement stage with smaller search space to find the best match iteratively.

Our initial experiments showed that the EPZS algorithm provides the best trade-off in terms of complexity and performance than the other algorithms supported by the JM software. An example of experiment is plotted in Fig. 4. This figure is obtained by encoding BQmall 15 fps 4CIF video sequence with GOP=75 and SR=8. The performance of Simplified UMHexagon and EPZS algorithms are almost identical. However, from averaging the results of a number of experiments, it was observed that EPZS shows better trade-off than the former. Thus, EPZS algorithm is chosen to be used in all our experiments.

Fig. 5 illustrates the rate vs. instruction counts for the BQmall 15 fps CIF video. The GOP is varied from 1 to 75 while SR is set equal to 2. On the other hand, Fig. 6 shows the plot obtained with the same parameter combinations as the one used in Fig. 5 but using a larger resolution version of the video. It can be seen in these figures that the increase in instruction counts is proportional with the increase in resolution. For example, for GOP=2 and QP=30, the number of instructions needed to encode the 4CIF resolution (Fig. 6) is around 80000 million. However, the encoder uses 20000 million of instructions to encode the CIF version of the video with the same encoding parameter. The ratio of bitrate between the two resolutions is, however, smaller than 4. The bitrate is around 1000 kbps for CIF video and 3000 kbps for 4CIF video. Similar trend is observed in the traffic video sequence.

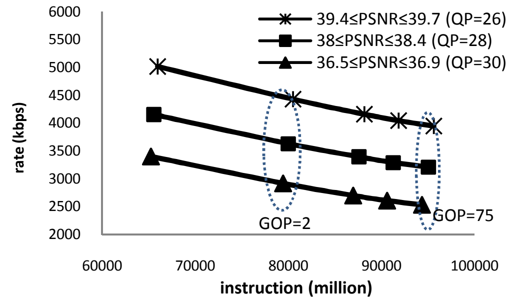

Figure 6. Bitrate and instruction counts trade-off, BQmall 15 fps 4CIF video, GOP={1,2,4,8,75} and SR=2

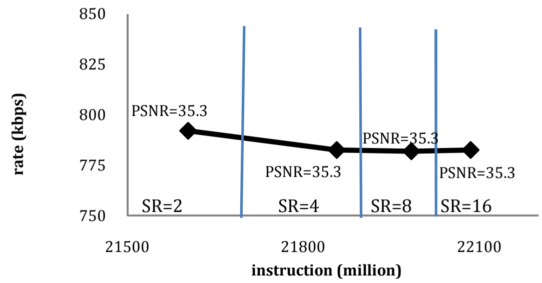

Figure 7. The effect of different SR values on bitrate, complexity, and PSNR (Traffic 15 fps CIF video, GOP=75, QP=28).

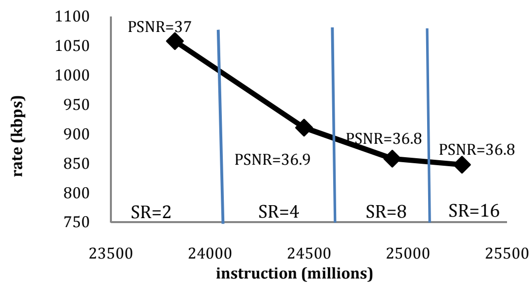

Figure 8. The effect of different SR values to bitrate, complexity, and PSNR (BQmall 15 fps CIF video, GOP=75, QP=28).

As mentioned in the previous paragraph, increasing the search range of motion estimation algorithm may provide us with a better prediction, i.e., better compression. However it comes with the price of increasing the complexity of the encoder. Thus, there is another trade-off between complexity and rate related with the value of SR. Fig. 7 shows the effect of using different values of SR to the bitrate, quality of the encoded bitstream and complexity of the encoder producing the bitstream for the Traffic 15 fps CIF video sequence. The same configurations was applied for the BQMall 15 fps CIF video and shown in Fig. 8.

As we can see from both Fig. 7 and Fig. 8, the quality of the video in terms of PSNR is almost the same regardless of the value of SR used to encode the video. However, the complexity of the encoder increases with the increase of SR. On the other hand, the bitrate of the encoded bitstream decreases with the increase of SR, especially when SR is increased from 2 until 8. The reduction of the bitrate obtained from increasing the SR from 8 to 16 is not really significant. However, when the resolution is increased into 4CIF, we notice that increasing SR up to 16 is still beneficial in terms of bitrate saving. Some interesting observations are

TABLE II. CONFIGURATION ID.

CID	GOP	Search Range
1	1	N/A
2	2	2
3	2	4
4	2	8
5	2	16
6	4	4
7	4	8
8	4	16
9	8	4
10	8	8
11	8	16
12	75	4
13	75	8
14	75	16

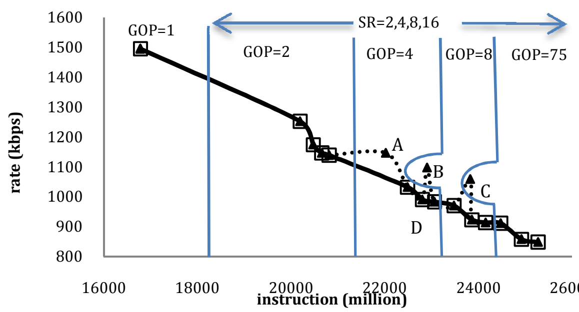

Figure 9. Trade of between complexity and rate for BQmall 15 fps CIF video, GOP={1,2,4,875}, SR={2,4,8,16}, QP=28.

obtained: 1) For the same value of QP, the quality of the encoded video is almost the same, regardless of the size of SR. However, there is a trade-off between encoding complexity and bitrate when different values of SR is used; 2) The value of maximum SR that should be used for a specific video sequence depends on the resolution of the video and its content; and 3) The quality of the reconstructed video depends on the content of the video and QP. The QPs used in Fig. 7 and Fig. 8 are the same. Hence the total number of instructions used for both video sequences are roughly the same. However, we can see that the quality of the encoded Traffic sequence is around 1.5 dB lower than the BQmall sequence.

Therefore, the encoding and communication energy consumption trade-off is content and application dependent. Nevertheless, a scheme to exploit the trade-off can be generalized from the results shown in the previous paragraphs that both GOP and SR sizes affect both the complexity and bitrate of the encoded bitstream. On the other hand, the quality of the video depends largely on the QP and the content of the video. Assuming that the video content of the targeted surveillance application is similar to the BQmall content, we can plot the bitrate vs. instruction count for different values of GOP and SR as shown in Fig. 9.

As we can see in Fig. 9, increasing the GOP size results in higher instruction counts and reduction in rate. Furthermore, for the same GOP value, increasing the value of SR will increase the instruction count but reduce the bitrate. However, looking at the overall picture, there are some configuration settings that are not optimal as shown by the dashed line in the figure. The in-efficient configurations, i.e., point A, B, and C, occurs when GOP>2 and SR=2. It is interesting to note that the configuration settings of point B (GOP=8, SR=2) and point D (GOP=4, SR=4) are using roughly the same number of instructions. However, the bitrate produced by configuration D is smaller than configuration B. Removing points A, B, and C from the plot produces the bold-line plot illustrated in Fig. 9. These configuration settings can be translated into a tabular format as shown in TABLE II. This table shows the configuration ID (CID) and its parameters setting that are shown as the bold-line plot in Fig. 9. It has to be noted that, these parameters setting are used on top of the configuration detailed in TABLE I.

III. POWER CONSUMPTION TRADE-OFF IN VSN

The total energy dissipation at a sensor node consists of the encoding power consumption (P_e), the transmission power consumption (P_t) and the reception power consumption (P_r):

$$P = P_e + P_t + P_r \,. \tag{1}$$

A. Encoding Power Consumption

The encoding energy consumption can be calculated as the number of CPU cycles consumed to execute the encoding task multiplied by the energy depletion per cycle. The value of energy depletion per cycle count can be obtained from the datasheet of the corresponding CPU. For the imote2 sensor platform, the energy depletion per cycle is 1.215 nJ [11].

The number of CPU cycles required to execute a task is usually calculated as the elapsed time of the encoding task multiplied by the CPU frequency. However, there are a couple of issues with this evaluation. First, as mentioned before, the time elapse may not represent the task's actual execution time since other processes could be running in the background. Second, many recent CPUs have the ability to control their clock frequency to save energy. In most cases this process is beyond the user's control. In this regard, *iprof* provides us with the total number of basic instructions needed to execute a certain task.

There are many parameters contributing to the overall energy consumption of the CPU. One of the most important parameter is the average cycle per instruction (*CPI*) [14]. If the total number of instructions used to execute a task is known, the energy depleted to execute that task can be calculated as the multiplication of the total number of cycles to execute that task and the average energy depleted per cycle. The total number of cycles to execute a task can be calculated as the multiplication of the instruction count (*IC*) and *CPI*. Therefore, the average energy consumption required to encode a frame can then be calculated as:

$$E_{ef} = \frac{IC \times CPI \times E_{ec}}{N_f} \,. \tag{2}$$

where, *IC* is the total number of instructions to encode a sequence (provided by *iprof*), *CPI* is the average number of cycles per instruction of the CPU, E_{ec} is the energy depleted per cycle and N_f is the number of frames. The value of *CPI* depends on the processor's architecture and the type of task executed. However, the average *CPI* value is usually provided by the chip manufacturer and it is used to compare

TABLE III. PARAMETERS USED.

Parameters	Description	value
α	Energy cost for transmitting 1 bit	0.5 J/Mb
β	Transmit amplifier coefficient	1.3·10-8 J/Mb/m4
λ	Energy cost for receiving 1 bit	0.5 J/Mb
η	Path loss exponent	4
CPI	XScale average cycle per instruction [16]	1.78
Ec	Energy depleted per cycle for imote2 [11]	1.215 nJ

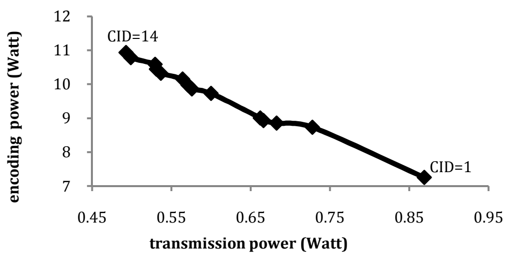

Figure 10. Encoding and transmission power consumption for encoder configuration shown in TABLE II, QP=28, distance to sink=50m.

its performance with other processor. It has been reported that the CPI value for the Intel XScale processor, which is used in imote2 sensor network is 1.78 [15]. With *FPS* denotes the frame rate of the video sequence, the encoding power consumption can then be calculated as follows:

$$P_e = E_{ef} \times FPS\,. \tag{3}$$

B. Communication Power Consumption

In order to calculate the total transmission power consumption of all the nodes in a VSN, the following video sensor node model is used in this paper. Video sensor nodes are assumed to be deployed statically. Each node is powered by battery that has a certain initial amount of energy. It is also assumed that node *i* can communicate with node *j* if a link between those nodes exists. Sensor node *i* can capture and encode video, and then generate data traffic with a source rate R_i.

The general energy consumption model for a wireless communication transmitter and receiver as presented in [16] is used in this paper. The total transmission power consumption of node *i* is the sum of all power consumed to transmit data to other nodes within its transmission range. The transmission power consumption is calculated as follow:

$$P_{ti} = \sum \left(\alpha + \beta \cdot d_i^{\,\eta}\right) \cdot R_i\,. \tag{4}$$

where, P_{ti} is the transmission power consumption of node *i*, α and β are constant coefficients, η is the path loss exponent, and d_i is the distance between node *i* and the relay node/sink. The total reception power consumption of node *j* is the sum of all power consumed to receive data from other nodes, as formulated below, where λ is a constant coefficient:

$$P_{rj} = \sum \lambda \cdot R_j\,. \tag{5}$$

Furthermore, the total communication power consumption dissipated for transmitting the encoded bitstream from node *i* to the sink, separated by *n*-hops can be calculated as follow:

$$P_c = n \cdot P_t + (n-1) \cdot P_r\,. \tag{6}$$

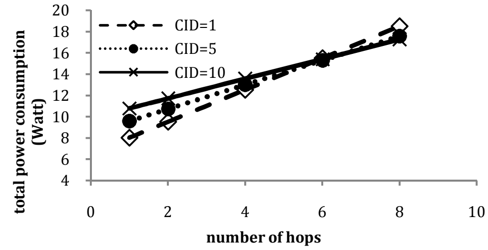

Figure 11. Total power consumption for three select encoder configurations in multi-hop transmission, QP=28, hop distance=20m.

C. Trade-off Between Computation and Communication

The experiments carried out in previous sections shows that by appropriately selecting some combinations of encoding parameters, we can analyze the trade-off between encoder complexity and rate. This can be done while still guaranteeing the quality of the encoded video to be in a specific range of PSNR value by setting the value of QP. In this regard, the configuration parameters shown in TABLE II can be used as a model to analyze the trade-off between encoding and communication power consumption. For the purpose of this analysis, parameters shown in TABLE III are used.

Fig. 10 shows the encoding and communication power consumption for the configuration shown in TABLE II and QP=28. In this figure, it is assumed that the transmission distance is 50m. As can be seen from the figure, for CID=1, the encoding power consumption of the node is equal to roughly 7.2 Watt while its transmission power consumption is around 0.87 Watt. However, for CID=14, the encoding power consumption is roughly 10.9 Watt while its transmission power consumption is around 0.49 Watt. This figure also shows the trade-off between complexity and rate of the encoder and its effect to node's power consumption. Certainly, by performing more encoding process, i.e., to get a lower bitrate, a node can save transmission power consumption. On the other hand, when the node is required to save encoding energy consumption, it can afford to spend more energy on transmission. Therefore, the encoding configuration can be varied according to the configuration ID detailed in TABLE II, resulting into different pairs of complexity and rate.

The knowledge of complexity-rate pair of encoding process is especially helpful in a multi-hop environment such as VSN. This information can be used to minimize the total power consumption of the VSN. For example, as shown in Fig. 11, when the node is 1-hop away from the sink, the optimal configuration for the node is CID=1. However, for nodes that are four hops away from the sink, they need to use configuration CID=10 in order to save energy. It has to be noted that, in Fig. 11, it is assumed that the communication is between a single source-sink pair separated by *n*-hops. When the number of sources increases, relay nodes that are closer to the sink will spend more energy in transmission. Nevertheless, we believe that by knowing the complexity-rate trade-off as shown in this paper, each node can select the best configuration ID to be used to minimize overall power consumption in such cases.

TABLE IV. EXAMPLE ENCODING CONFIGURATION SCHEME IN A SIMPLE VSN, USING CID FROM TABLE II.

Scheme used	Node's configuration (CID)		
	Node 1	*Node 2*	*Node 3*
Scheme 1 (simple encoding)	1	1	1
Scheme 2 (low transmission rate)	14	14	14
Scheme 3 (complexity-rate trade-off)	1	5	10

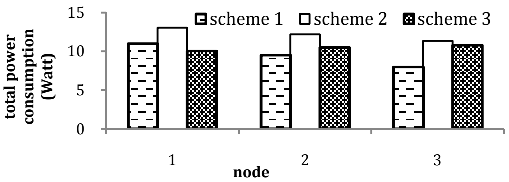

Figure 14. Nodes's total power consumption for the three schemes.

In order to illustrate the usefulness of knowing the encoder's complexity-rate trade-off, consider a simple VSN consisting of 3 nodes and a sink. The three nodes are separated by 20m each. Node 3 can only forward its data to node 2, while node 2 can only relay data to node 1. Node 1 is the one responsible to relay all information to the sink. It is also assumed that all nodes can capture, encode, and transmit the data at the same time. TABLE IV shows three example scenario that one can apply in such a VSN. Scheme 1 assumes that energy consumption can be minimized by applying simple encoding scheme (intra-coding only). On the other hand, in scheme 2, nodes are supposed to encode better to reduce transmission energy consumption. In scheme 3, the complexity-rate trade off is used to determine the nodes' configuration. It has to be noted that, the CIDs used in scheme 3 is not the optimal solution for this simple VSN. It is chosen just to show an example on using the encoding-transmission energy trade-off explained earlier.

Fig. 12 shows the nodes' total power consumption for the three schemes. It can be seen that, in general, node 1's energy consumption is the highest among all nodes, regardless whichever schemes used in the VSN. However, the difference in power consumption for the three nodes in scheme 3 is very small compared to the the other schemes. Furthermore, node 1's energy consumption in scheme 3 is equal to 10 Watt. This is smaller than what is obtained using scheme 1 (11 W) and scheme 2 (13 W). The utility of a VSN system can be measured by its network lifetime. For an important application such as a surveillance system, the minimum node lifetime is used in this paper. Assuming that all nodes has the same initial energy resources of 50 kJ, the lifetime of the VSN using scheme 3 is then 4636s. This is in comparison with 4544s (scheme 1) and 3829s (scheme 2) obtained by the other schemes.

IV. CONCLUDING REMARKS

In this paper, we have shown that by varying some H.264/AVC encoder parameters such as QP, GOP size, and search range for motion estimation, we observe different pairs of encoder power consumption and bitrate produced for a specific range of PSNR values. These pairs of encoder power consumption and bitrate can be used as the operating points for the video node in order to minimize the total power consumption. However, it was seen that the total power consumption of the video nodes depends also to their distance to the sink or the number of hops between the source and the sink. In addition to that when the number of sources increases, relay nodes that are closer to the sink will spend more energy in transmission to relay the information from downstream nodes. Yet, to prolong the utility of VSN applications, the power consumption of all nodes should be roughly equal. In this regard, the complexity-rate trade-off shown in this paper can be used as a guideline for minimizing the overall power consumption.

ACKNOWLEDGMENT

This work was partly supported by Natural Sciences and Engineering Research Council of Canada (NSERC) and the Institute for Computing Information and Cognitive Systems (ICICS) at UBC. This work was also made possible by NPRP grant # NPRP 4-463-2-172 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 51, no. 4, pp. 921–960, Mar. 2007.

[2] X. Ren and Z. Yang, "Research on the key issue in video sensor network," presented at the Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on, Chengdu, 2010, vol. 7, pp. 423 – 426.

[3] A. Seema and M. Reisslein, "Towards efficient wireless video sensor networks: a survey of existing node architectures and proposal for a flexi-WVSNP design," *Communications Surveys & Tutorials, IEEE*, vol. 13, no. 3, pp. 462 – 486, Third Quarter 2011.

[4] J. Chen, Z. Safar, and J. A. Sorensen, "Multimodal Wireless Networks: Communication and Surveillance on the Same Infrastructure," *Information Forensics and Security, IEEE Transactions on*, vol. 2, no. 3, pp. 468–484, Sep. 2007.

[5] T. D. R¨aty, "Survey on Contemporary Remote Surveillance Systems for Public Safety," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 40, no. 5, pp. 493–515, Sep. 2010.

[6] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[7] I. E. Richardson, *The H.264 Advanced Video Compression Standard*, Second ed. John Wiley & Sons, Ltd, 2010.

[8] H. K. Zrida, A. C. Ammari, M. Abid, and A. Jemai, "Complexity/Performance Analysis of a H.264/AVC Video Encoder," in *Recent Advances on Video Coding, InTech*.

[9] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity," *IEEE Circuits and System Magazine*, vol. 4, no. 1, pp. 7–28, 2004.

[10] N. Imran, B.-C. Seet, and Alvis C. M. Fong, "A comparative analysis of video codecs for multihop wireless video sensor networks," *Multimedia Systems*, vol. 18, no. 5, pp. 373–389, 2012.

[11] J. J. Ahmad, H. A. Khan, and S. A. Khayam, "Energy efficient video compression for wireless sensor networks," presented at the Information Sciences and Systems, 2009. CISS 2009. 43rd Annual Conference on, Baltimore, MD, 2009, pp. 629 – 634.

[12] P. M. Kuhn, "A Complexity Analysis Tool: iprof (version 0.41)," *ISO/IEC JTC1/SC29/WG11/M3551*, Jul-1998.

[13] ISO/IEC JTC1/SC29/WG11, "Joint Call for Proposals on Video Com-pression Technology." Jan-2010.

[14] S. Gurun and C. Krintz, "Energy Characterization of the Stargate Sensor Network Gateway," Department of Computer Science, University of California, Santa Barbara, 2006-08, Jun. 2006.

[15] D. Chinnery and K. Keutzer, *Closing the Power Gap between ASIC & Custom: Tools and Techniques for Low Power Design*, 1st ed. Springer, 2007.

[16] T. S. Rappaport, *Wireless communications: principles and practice*, 2nd ed. Prentice Hall, 2001.