PRoFIT—Priority-Based Routing Framework for Image Transmission in Visual Sensor Networks

Adil A. Sheikh, Atif Naseer, and Emad Felemban

Abstract—**The primary goal of VSN deployment is to capture images and send them to power rich sink nodes. As image data is much larger than data sensed by a typical wireless sensor network (WSN), applications of VSN require much bigger amount of data to be transferred to the sink. Due to constraints of WSN such as low energy, limited CPU power and scarce memory, transmission of large amount of data becomes challenging. Furthermore, to take action, some VSN applications require critical image features sooner than the entire image. In this paper, Priority-based Routing Framework for Image Transmission (PRoFIT) is presented along with experimentation on real camera testbed. PRoFIT is designed to deliver critical image features at high priority to the sink node for early processing. PRoFIT does not reduce the total amount of data to be transmitted, however image features critical for some VSN applications are received at sink node much sooner using PRoFIT. Peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) analyses show that PRoFIT improves VSN application response time as compared to priority-less routing.**

Index Terms—**Multimedia transmission, priority routing, visual sensor network, camera test-bed.**

I. INTRODUCTION

Sensor Networks (VSN) are wireless sensor networks (WSN) equipped with image sensors or cameras. Typical WSNs sense and transmit environment factors [1], whereas VSNs sense and transmit images in addition to typical WSN data. VSN applications, such as periphery monitoring, intruder detection, surveillance, etc. require very large amounts of image data to be transmitted from camera nodes to sink. One measurement of a typical WSN sensed data such as light intensity, humidity, pressure, etc. is typically 4 bytes. However, an image captured by VSN is of several kilobytes. For example, a CIF image is of 352 x 240 pixel resolution. Having 32 bits per pixel makes one CIF image of 337; 920 bytes in size. Hence, VSN nodes send very large amount of image data as compared to typical WSN nodes. Although, some VSN architects use image compression techniques such as Discrete Cosine Transforms [2] or Discrete Wavelet Transforms [3], yet these techniques do not reduce data size significantly. Other methodologies such as progressive image

transmission [4] and bit plane coding [5] transmit image in multiple layers. These methodologies partition the image into layers such that the primary layer contains the most significant features of the image, whereas successive layers contain details that when integrated with the primary layer, reconstruct the complete image. Each layer is of a fraction in size of the complete image. If packets of the primary layer are transmitted at a priority through a VSN, they will reach the sink node much sooner than the complete image. Some VSN applications may improve response time by processing the primary layer(s) first instead of waiting for the complete image to be received. This motivation led to the design of Priority-based Routing Framework for Image Transmission (PRoFIT) for VSN. The usefulness of PRoFIT was demonstrated using simulations in our previous work [6]. Our previous work focused on end-to end delays and packet delivery ratios at different deadlines. In this paper, we extend our previous work by implementing and testing PRoFIT on a real VSN. We use Peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) analyses to show that PRoFIT improves VSN application response time as compared to priority-less routing. The remaining of the paper is divided as follows. Section II describes the design of PRoFIT. Section III discusses the details of PRoFIT's implementation. Section IV elaborates the results, whereas the paper is concluded in Section VI.

II. PROFIT DESIGN

This paper discusses only point-to-point communication scenario between VSN node types. (1) The first type is the node with image sensor called camera node. (2) The second type of node is the sink node.

A. Camera Node

Fig. 1 show the processes running on the camera nodes. The camera node captures the image and transmits it in four steps.

Step 1 (Encoding): For a captured grayscale image, 1 byte (8 bits) is used to represent the gray level of one pixel. PRoFIT encodes pairs of pixels of the captured image and stores them into two layers. In the current version of PRoFIT, bit plane encoding [5] is used to encode the image into two layers; primary layer and secondary layer. The primary layer consists of bit planes 8 to 5 of a pixel of the captured image while the secondary layer consists of bit planes 4 to 1 of the same pixel. 56 bytes 1 byte 2-bytes 69 bytes Number Payload This mechanism is shown in Fig. 2. For example, the first and second pixels of the captured image are partitioned into four sets of four bits each. The four most significant bits (MSBs) of the first pixel of captured image are stored at the four MSBs of first pixel of primary layer. The four MSBs of the

Adil A. Sheikh and Atif Naseer are with the Science and Technology Unit, Umm Al-Qura University, Makkah, Saudi Arabia (email: aasheikh@uqu.edu.sa, anahmed@uqu.edu.sa).

Emad Felemban is with the Department of Computer Engineering, Umm Al-Qura University, Makkah, Saudi Arabia (email: eafelemban@uqu.edu.sa).

second pixel of captured image are stored at the four least significant bits (LSBs) of first pixel of primary layer. The four LSBs of the first pixel of the captured image are stored at the four MSBs of the first pixel of the secondary layer. The four LSBs of the second pixel of the captured image are stored at the four LSBs of the secondary layer. Similarly, the third and fourth pixel of the captured image becomes second pixel of each layer, and so on. In this way, each encoded pixel of primary layer contains the MSBs of two pixels of captured image, whereas each encoded pixel of the secondary layer contains LSBs of the same two pixels of captured image. Note that the size of both encoded layers combined is same as the original image.
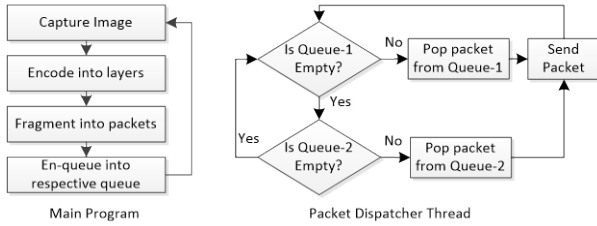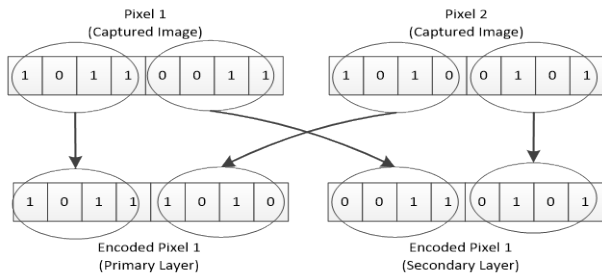

Fig. 1. Processes running on camera node.

Fig. 2. Encoding captured image pixels into two layers.

Step 2 (Fragmentation): Packet size of VSN physical layer may be limited. As image data is much larger, it must be fragmented into packets of size that the physical layer can transit. Maximum packet payload size is 72 bytes in a typical sensor network that uses Zigbee for transmission. Hence, the current version of PRoFIT has been designed with IEEE 802.3 (Ethernet) for transmission with a packet size of 72 bytes. The physical layer can be replaced in the future versions. PRoFIT packet contains three fields; flags, sequence number and payload. Sequence number field is required in order to defragment at the sink. Sequence number field of two bytes can accommodate QCIF (176 x 120) to 4 MP (2240 x 1680) image resolution formats. A QCIF image would be fragmented into 307 fragments, whereas 4MP image will be fragmented into 54,540 fragments. One byte of the packet represents various flags such as priority, new image, etc. The remaining 69 bytes are used as payload to send encoded pixels.

Step 3 (En-queuing Packets): PRoFIT contains two queues; A high-priority packet's queue called Queue-1 and a low priority packet's queue called Queue-2. As the primary and secondary layers of the image are being packetized (Step 2), each packet is inserted at respective queue's tail. The packets of primary layer are inserted into Queue-1 and packets of secondary layer are inserted into Queue-2.

Step 4 (Dispatching Packets): The packet dispatcher runs in a loop as a child thread of the main program. Fig.1 depicts

the mechanism of selecting packets from the queues and transmitting them. If there is a packet in Queue-1, then the packet at the head of Queue-1 is extracted for transmission. If Queue-1 is empty and there is a packet in Queue-2, it will be selected for transmission. If both queues are empty, the dispatcher waits for a packet to arrive in one of the queues. Using the above four steps, high priority packets are always delivered before low priority packets.

B. Sink Node

The sink node uses two steps to reconstruct image.

Step 1 (Defragment Packets): Based on a preset image resolution, the sink pre-allocates memory to store incoming encoded pixels into the primary and secondary layers. The sink node uses the sequence number field to defragment. Packet with sequence numbers 1, 3, 5 and so on are part of primary layer, whereas packets with sequence numbers 2, 4, 6 and so on are part of secondary layer. Each packet contains 69 encoded pixels. As a packet is received, the sequence number is extracted to find the location of the encoded pixels in the respective layer. Encoded pixels from the packet with sequence number 1 are stored in place of the first 69 pixels of the primary layer. Encoded pixels from the packet with sequence number 3 are stored in place of the second 69 pixels of the primary layer, and so on. Encoded pixels from the packet with sequence number 2 are stored in place of the first 69 pixels of the secondary layer. Encoded pixels from the packet with sequence number 4 are stored in place of the second 69 pixels of the primary layer, and so on.

Step 2 (Decode Packets): On a preset periodic deadline, the sink node decodes the primary and secondary layers by splitting each encoded pixel and re-assembling the bit planes. The first pixel of the primary and secondary layer are split into four sets of four pixels each. The 4 MSBs of first pixel of primary layer are concatenated with 4 MSBs of the first pixel of secondary layer to reconstruct first pixel of sent image. The 4 LSBs of first pixel of primary layer are concatenated with 4 LSBs of the first pixel of secondary layer to reconstruct second pixel of the sent image. The 4 MSBs of second pixel of primary layer are concatenated with 4 MSBs of the second pixel of secondary layer to reconstruct third pixel of sent image. The 4 LSBs of second pixel of primary layer are concatenated with 4 LSBs of the second pixel of secondary layer to reconstruct fourth pixel of the sent image, and so on. At any given time, more packets of the primary layer than secondary layer would have been received at sink. Therefore, almost half way through image transmission, critical image features are available at the sink node for processing.

III. IMPLEMENTATION

Fig. 3 depicts the architecture of PRoFIT's implementation. For the camera node, PRoFIT has been implemented on Roseek Cheetah-1 Series RSCT128SK-H000G Smart Cameras. It comes with a Quad-core Intel ATOM CPU and Ubuntu 16.04. Schematics, specifications and data sheets of the camera system are available online[1]. OpenCV[2] version 2.6 has ported on it. A thin application programming

interface layer has been ported over encoding, decoding, fragmentation, defragmentation and enqueuing modules. The image capturing module uses OpenCV to capture an image. Ethernet has been selected as the physical layer, however packet sizes have been kept to emulate VSN physical layer. After the image has been encoded and fragmented, En-queuing modules uses UDP to send the packet over Ethernet.
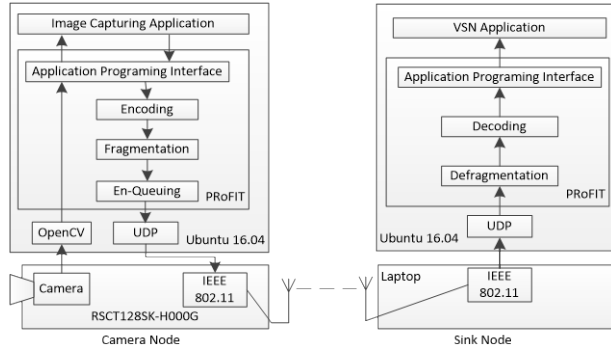

Fig. 3. PRoFIT architecture.

For the sink node, only defragmentation and decoding modules are implemented on an Intel i3 based laptop. When the packet is received on the sink node through UDP, it is defragmented into the primary and secondary layers. When the preset periodic timer expires, the decoded image is constructed. Peak Signal-to-Noise (PSNR) and Structural Similarity Index (SSIM) are calculated to check the quality of image received. Source code of implemented modules is available at the project website[3].

IV. RESULTS

Two standard images of different resolutions were selected to be used to calculate the quality of image received. These were obtained from The Standford Center for Image Systems Engineering (SCIEN)[4]. These are baboon.tif with 256 x 256 resolution and lena.tif with 512 x 512 resolution. Multiple experiments were run to transmit these images to from camera node to sink node. The figures below show the average SSIM and PSNR of respective images resolutions.

Fig. 4 (Top) shows the PSNR and Fig. 4 (Bottom) shows SSIM values as image is received at sink node. It can be seen that with PRoFIT enabled, the PSNR and SSIM are higher right from the start of transmission.

Fig. 5(a) shows the image reconstructed at 20,000 microseconds after first packet was received at the sink with PRoFIT enabled. Below it, Fig. 5(d) shows the image reconstructed at the same time without PRoFIT. Fig. 5(b) shows the image reconstructed at 35,000 microseconds with PRoFIT enabled. Below it, Fig. 5(e) shows the image reconstructed at the same time without PRoFIT. Finally, Fig. 5(c) shows the image reconstructed when the last packet has been received at about 58,000 microseconds with PRoFIT enabled. Below it, Fig. 5(e) shows the image reconstructed at the same time without PRoFIT. It can be concluded that at any given time, the quality of reconstructed image is better with PRoFIT enabled.

[3] www.emadfelemban.com/gvsn

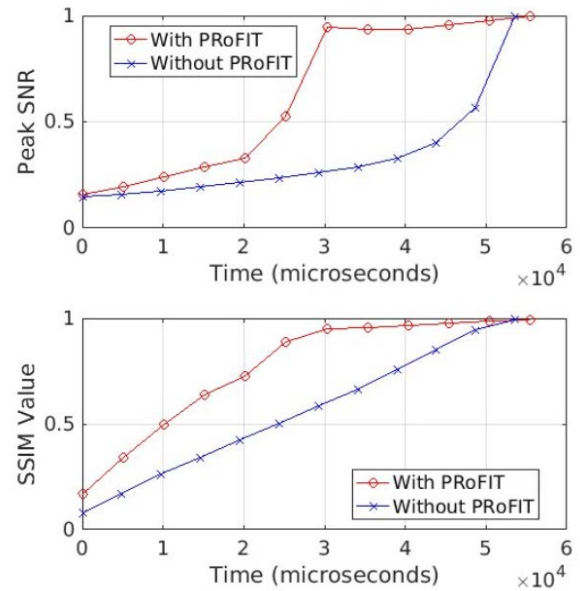[4] scien.stanford.edu/index.php/test-images-and-videos/

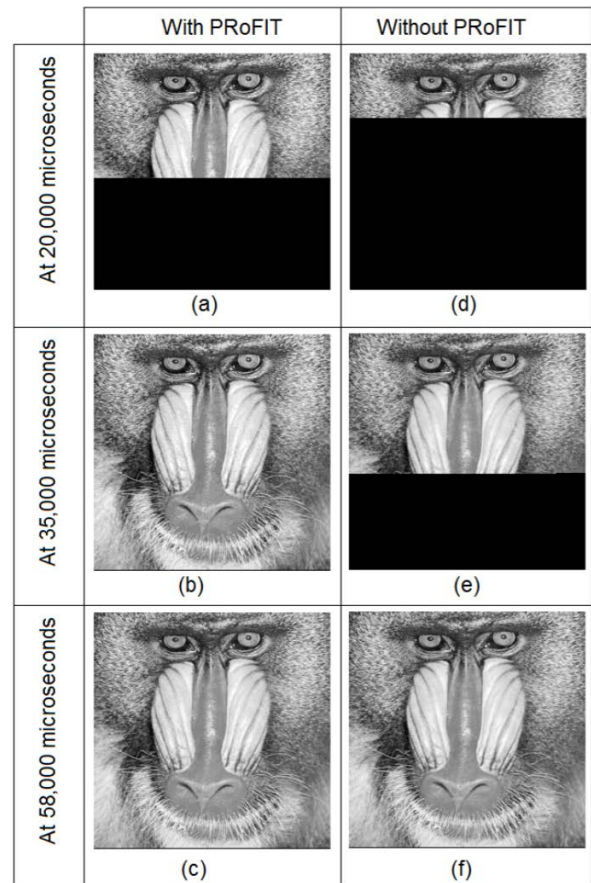Fig. 4. PSNR and SSIM for baboon.tif (Resolution 256 X 256).

Fig. 5. Received images at various deadlines.

At half-time of transmission (35000 microseconds, approximately), values of both PSNR and SSIM (Fig. 4) have reached above 0.9, whereas without PRoFIT, PSNR and SSIM are 2.6 and 6.0, respectively. At this point, the PSNR peaks because the primary layer has been fully received (Fig. 5(b)). On the other hand, without PRoFIT, a little more than half of the image has been received but in high quality (Fig. 5(e)). Without PRoFIT, none of the pixels for the lower part of the image have been received. This impacts the overall SSIM and PSNR of the image. It is also concluded that if primary layer containing half information of the image (bit

planes 8 to 5) is fully received results into better PSNR and SSIM as compared to half of the entire image with all bit planes.
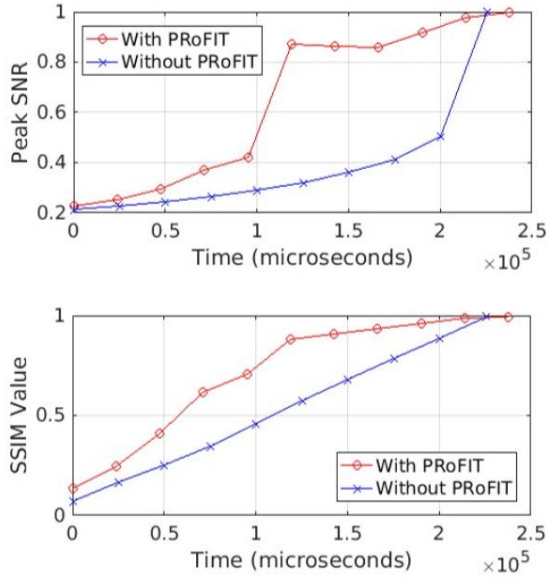

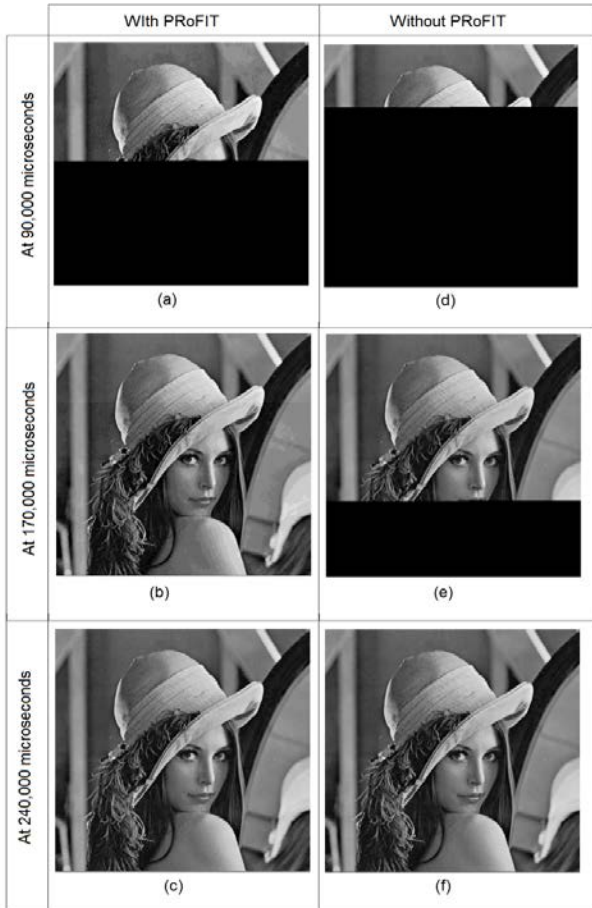Fig. 6. PSNR and SSIM for lena.tif (Resolution 512 X 512).

Fig. 7. Received images at various deadlines.

Similar results are seen in Fig. 6 and Fig. 7. This demonstrates that PRoFIT provide better results irrespective of image resolution. Fig. 8 shows parts of Fig. 7(b) and Fig. 7(e). If Fig. 8(a) and Fig. 8(b) are compared visually at area marked by A then the quality of both areas is same. At this mark, pixels of both primary and secondary layer have been received. The reconstructed area A has high SSIM and PSNR.

On the other hand, only primary layer pixels have been received for area marked by B. It can be seen that the quality of image with PRoFIT is lower. Finally, for area marked by C shows that the primary layer has been received with PRoFIT enabled, however no pixels have been received without PRoFIT. Area B and C with low quality are better than high quality area B and no pixels for area C. This is the trade off of using PRoFIT that the user must take into account.

Fig. 8. Trade off between use of PRoFIT.

V. RELATED WORK

Most of the work done in the field of image transmission has been for traditional power rich camera networks. For WSNs, Nasri *et al.* [3], [7] propose adaptive image transfer using JPEG2000 and Discrete Wavelet Transforms. Through simulations show that their scheme optimizes network lifetime and reduces memory requirements. Our framework follows a similar approach.

Mostly, authors have focused on different aspects of VSN performance, such as network lifetime. Zhou *et al.* [8] propose fast visual background extraction algorithm to detect motion and then use prioritized transmission to balance between delay and energy consumption.

Bhosale *et al.* [9] proposed prioritized scheduling technique at inter-node level to maximize the lifetime of network.

Spachos *et al.* [10], [11] proposed a dynamic routing protocol called Content Relevance Opportunistic Routing (CROR) for improving overall network performance by focusing on Quality of Service parameters such as energy efficiency and packet delay.

Han *et al.* [12] have surveyed routing protocols that balance out the trade-off between network lifetime and QoS requirements. They provide a baseline for "green routing protocols".

Similar to ours, Bondi *et al.* [13] provided a framework that is implemented on real VSN nodes. They advocate on comparing analyze-then-compress (distributed approach) verses compress-then-analyze (centralized approach).

Eriksson *et al.* [14] also demonstrate that distributed optimization can improve system response time for VSNs.

VI. CONCLUSIONS AND FUTURE RESEARCH

This paper discussed the implementation and the results of experiments of PRoFIT, our proposed priority-based image transmission protocol over a real point-to-point test-bed. Both distribution and centralized image processing techniques can be used to transfer image data. PRoFIT used prioritized routing for critical information. Using PSNR and SSIM, we demonstrated that at any given time, the quality of received image at the sink is better than if an image was sent without PRoFIT. In this work, IEEE 802.11 was chosen as the physical layer, however in future we plan to integrate Zigbee with PRoFIT. Moreover, PRoFIT will be extended to run over a multihop VSN.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102-114, Aug. 2002.

[2] L. Makkaoui, V. Lecuire, and J. M. Moureaux, "Fast zonal dct-based image compression for wireless camera sensor networks," in *Proc. of 2010 2nd International Conference on Image Processing Theory, Tools and Applications*, July 2010, pp. 126-129.

[3] M. Nasri, A. Helali, H. Sghaier, and H. Maaref, "Energy-efficient wavelet image compression in wireless sensor network," in *Proc. of 2010 International Conference on Wireless and Ubiquitous Systems*, Oct. 2010, pp. 1-7.

[4] V. Delport, J. Kuhnert, and S. Roessler, "Implementation of an ieee 802.15.4 compliant multi-hop wireless sensor network for energy efficient sensor data and progressive image transmission," in *Proc. of 2014 6th European Embedded Design in Education and Research Conference (EDERC)*, Sep. 2014, pp. 182-186.

[5] H. Kikuchi, K. Funahashi, and S. Muramatsu, "Simple bit-plane coding for lossless image compression and extended functionalities," in *Proc. of 2009 Picture Coding Symposium*, May 2009, pp. 1-4.

[6] E. Felemban, A. A. Sheikh, and M. A. Manzoor, "Improving response time in time critical visual sensor network applications," *Ad Hoc Networks*, vol. 23, pp. 65-79, 2014.

[7] M. Nasri, A. Helali, H. Sghaier, and H. Maaref, "Adaptive image transfer for wireless sensor networks (WSNs)," in *Proc. of 5th International Conference on Design Technology of Integrated Systems in Nanoscale Era*, March 2010, pp. 1-7.

[8] F. Zhou, Z. Lu, J. Liu, Y. Xia, and X. Wei, "Low-cost prioritization of microorganism image data in wireless visual sensor network," in *Proc. of 2016 International Computer Symposium (ICS)*, Dec. 2016, pp. 685-689.

[9] V. D. Bhosale and R. A. Satao, "Lifetime maximization in mobile visual sensor network by priority assignment," in *Proc. of 2015 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, Dec. 2015, pp. 695-698.

[10] P. Spachos, A. K. Marnerides, and D. Hatzinakos, "Content relevance opportunistic routing for wireless multimedia sensor networks," in *Proc. of 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2014, pp. 263-268.

[11] P. Spachos, D. Toumpakaris, and D. Hatzinakos, "Qos and energy-aware dynamic routing in wireless multimedia sensor networks," in *Proc. of 2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 6935-6940.

[12] G. Han, J. Jiang, M. Guizani, and J. J. P. C. Rodrigues, "Green routing protocols for wireless multimedia sensor networks," *IEEE Wireless Communications*, vol. 23, no. 6, pp. 140-146, December 2016.

[13] L. Bondi, L. Baroffio, M. Cesana, A. Redondi, and M. Tagliasacchi, "Ezvsn: An open-source and flexible framework for visual sensor networks," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 767-778, Oct. 2016.

[14] E. Eriksson, G. Dán, and V. Fodor, "Coordinating distributed algorithms for feature extraction offloading in multi-camera visual sensor networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 11, pp. 3288-3299, 2017.

Adil A. Sheikh is a researcher at the Science and Technology Unit in Umm Al-Qura University, Kingdom of Saudi Arabia. He holds a double Masters from RWTH Aachen University and Universit degli Studi di Trento. His research interests include sensor networks, embedded systems, intelligent transportation systems, smart grids and hardware software co-design.

Atif Naseer is an academician and researcher by profession and is currently serving as a researcher at Umm Al Qura University. He has been attached with academia for over seven years and has been part of many research and development projects. His current research interests include internet of things, middleware modeling of wireless sensor networks, semantic interoperability, and big data analysis.

Emad Felemban is an associate professor in Umm Al-Qura University Saudi Arabia. He graduated from Ohio State University in 2009 with PhD degree. His research interest includes wireless sensor networks, wireless networks, underwater wireless networks and intelligent transportation systems. He won many research grants from different research fund agencies in Saudi Arabia.