# Bi-Level Video Codec for Machine Vision Embedded Applications

K. Khursheed[1], M. Imran[1], N. Ahmad[1], M. O'Nils[1]

[1]*Department of Information Technology and Media, Division of Electronics Design,*
*Mid Sweden University,*
*Sweden*
*khursheed.khursheed@miun.se*

*Abstract*—**Wireless Visual Sensor Networks (WVSN) are feasible today due to the advancement in many fields of electronics such as Complementary Metal Oxide Semiconductor (CMOS) cameras, low power electronics, distributed computing and radio transceivers. The energy budget in WVSN is limited due to the small form factor of the Visual Sensor Nodes (VSNs) and the wireless nature of the application. The images captured by VSN contain huge amount of data which leads to high communication energy consumptions. Hence there is a need for designing efficient algorithms which are computationally less complex and provide high compression ratio. The change coding and Region of Interest (ROIs) coding are the options for data reduction of the VSN. But, for higher number of objects in the images, the compression efficiency of both the change coding and ROI coding becomes worse than that of image coding. This paper explores the compression efficiency of the Bi-Level Video Codec (BVC) for several representative machine vision applications. We proposed to implement image coding, change coding and ROI coding at the VSN and to select the smallest bit stream among the three. Results show that the compression performance of the BVC for such applications is always better than that of change coding and ROI coding.**

*Index Terms*—**Wireless sensor networks, low power electronics, embedded computing, image communication.**

## I. INTRODUCTION

The Wireless Visual Sensor Network (WVSN) is composed of many Visual Sensor Nodes (VSNs). Each VSN consist of an image sensor for capturing images of the field of interest, microcontroller or Field Programmable gate Arrays (FPGA) for image processing, a radio transceiver for transmitting the results and energy resource for providing power to all the other components. The VSN is capable of performing complex image processing algorithms using limited energy resources due to the technological advancement in many fields of electronics such as Complementary Metal Oxide Semiconductor (CMOS) image sensors, wireless transceivers, low power computing platforms/embedded systems and distributed computing.

Many authors in the literature e.g. [1]-[3] have focused on different implementation strategies for performing vision processing tasks either locally at the VSN or at the server. Some authors considered to capture the image, compress it and send it to the server for further processing e.g. [1]. In this case the transmission energy consumption is higher because they considered transmitting the raw compressed images directly to the server.

On the other hand, the proposal from other researchers has been on performing all the vision processing tasks at the VSN and transmitting the final object dimensions to the server. In this case, the transmission energy consumption is low but the processing energy consumption is high because the VSN is performing operations for a longer duration.

An example representing all image processing tasks locally at the VSN is SensEye [2] which is a multi-tier network of heterogeneous VSNs and it aims at low power, low latency detection and wakeup. They implemented a surveillance application using the SensEye and performed advance image processing operations such as object detection, recognition and tracking.

Another such example is presented in [3], where the authors implemented a distributed vision processing system for human pose interpretation on a wireless smart camera network. They extracted critical joints of the subject in the scene in real time by performing local processing at the smart camera. The results achieved by various smart cameras are then transmitted to a server through the wireless channel for the reconstruction of the human pose.

The local processing as well as the wireless communication consumes a huge part of the total energy budget of the VSN. Communicating the results from the VSN without local processing reduces the processing time but the consequence of this is the higher transmission energy consumption due to communicating the huge raw data. In contrast, executing the entire image processing tasks locally at the VSN and transmitting the final results, reduces the transmission energy consumption but its drawback is the higher processing energy consumption because of the increased processing time at the VSN.

The studies on Intelligence Partitioning (IP) between the VSN and the server in [4] and [5] have concluded that selecting a suitable IP strategy has a positive impact on the total energy consumption of the VSN. It is concluded in [6] that compressing the binary image after pre-processing and segmentation is a good strategy for achieving a general architecture for the measurement applications of WVSN.

Figure 1 shows the generalized architecture from [6]. Based on this generalized architecture, we explored the compression efficiency of the six bi-level image coding methods and presented the results in [7]. The conclusion drawn in [7] was that the JBIG2 [8], the Group4 [9] and Gzip_pack [10] offer better compression performance.
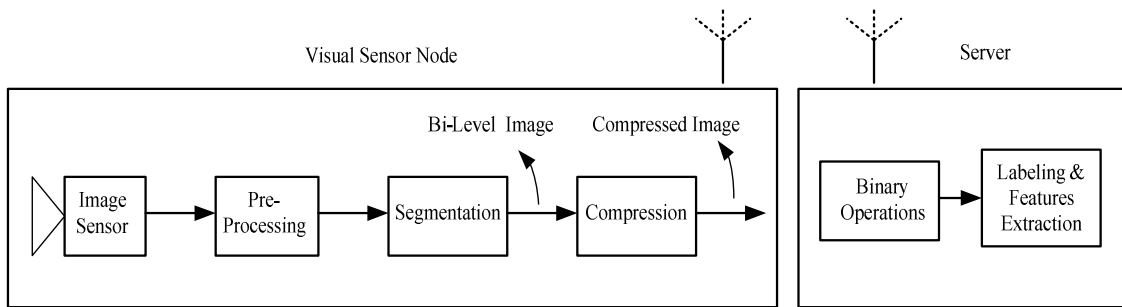


Fig. 1. The suitable intelligence partitioning strategies.

Bi-level image compression standards are helpful in reducing the information amount in the segmented images. But the compression efficiency of the compression standards is restricted to the entropy of the used compression algorithm. So, there is a need for the exploration of other efficient compression techniques which can provide further data reduction.

The adjacent frames of the video in some machine vision applications are almost similar. In other words the differences in the neighbouring frames of the video in these applications are very few. We explored the possibility of further data reduction in [11] based on change coding, where we concluded that change coding is better than image coding (the image coding is investigated in [7]) for up to 95% changes in terms of number of objects in the neighbouring frames. We explored the compression efficiency of Region of Interest (ROI) coding for various geometrical shaped objects in the black images and presented the results in [12]. The results of ROI coding in [12] showed some improvements in the compression performance compared to the results of change coding in [11].

One trouble with both ROI coding and change coding is the degradation of their compression performance for too many objects in the frames. In [13], we proposed a generalized method which provides better compression efficiency compared to both the ROI coding and change coding but we analysed its performance for statistical images only. In current work, our aim is to explore the performance of Bi-Level Video Codec (BVC) for real life machine vision applications. Our aim is to explore and provide a detailed explanation of the compression efficiency based on BVC for three categories of real life machine vision embedded applications.

Our aim is to show that BVC is applicable to all kinds of applications and by doing this, we will be able to show that BVC provides better performance for applications with too few or too many objects in the frames (The compression performance of BVC will never become worse than image coding for frames with too many objects which was the case for ROI coding in [12] and change coding in [11]).

The remainder of the paper is organized as follows. Section II provides the related work while section III presents an overview of the BVC. Evaluated applications are presented in section IV. The results are presented in section V. Finally, section VI concludes the paper.

## II. RELATED WORK

Many methods have been proposed for the compression of scanned textual and medical images [14]–[16]. In machine vision applications, the images are quite different from scanned textual and medical images. The compression efficiency of the well known compression methods is investigated in [7] for machine vision applications. It is concluded in [7] that JBIG2, CCITT Group 4 and Gzip_pack provide better compression efficiency compared to other Bi-level image compression methods.

We investigated the possibility of further data reduction based on, change coding and ROI coding in [11], [12] respectively. We concluded in [11] that change coding provides better compression efficiency than image coding. We concluded in [12] that the compression efficiency of ROI coding is better than that of change coding. But both change coding and ROI coding provides better compression efficiency only for applications attributing too few changes in adjacent frames. A generalized method which is effective for all applications is needed. We think that BVC is such a method and will provide better compression efficiency for all applications i.e. applications with too few or too many objects in the change frame.

## III. BI-LEVEL VIDEO COMPRESSION

A number of lossless Bi-level image compression algorithms exist and some of them are the Ziv-Lempel algorithms [15], Efficient partitioning into rectangular regions [16], Arithmetic coding [17], CCITT Group 4 [9], JBIG2 [8] etc. Any of these best compression schemes can be used for the analysis of BVC in this paper where we have selected the well known CCITT Group 4 because of its relatively high compression efficiency and low computational complexity based on the analysis in [7].

The architecture for BVC is shown in Fig. 2. Every new frame needs to be saved (overwritten) in memory in the form of alternating black and white runs. We have applied run length encoding for frame storage for efficiency purposes. The change frame is determined by performing the exclusive-OR operation on the pixels (segmented pixels) of the current frame from the camera and the pixels of previous frame from memory. The change frame is given as input to the ROI block for finding the ROIs in it.

The Group 4 is used to compress the original frame, the

change frame and the detected ROIs. Smallest of the three compressed bit stream is selected as the output of the BVC.
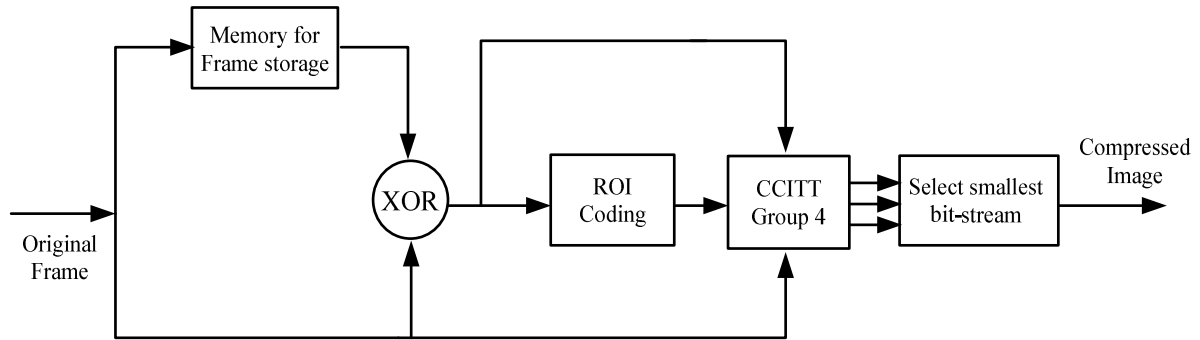


Fig. 2. The architecture for Bi-Level Video Codec.

## IV. EVALUATED APPLICATIONS

We have selected three categories of applications for evaluating the performance of our proposed method. These applications represent three categories regarding the transition from black to white pixels in the frames and vice versa. One category represents applications attributing too many objects in the frames such as in Meter Reading applications. The second category represents application having too few objects in the frames e.g. LED light detection (for tracking/localization of robots/objects) and Magnetic particle detection in hydraulic systems etc. The third category is the one where there are too many transitions from black to white pixels and white to black pixels in the change frames due to the illumination noise or movements of the objects such as in Human Detection applications.

### A. Applications attributing too many objects

The authors in [1] designed and implemented a VSN for Meter Reading application, where they considered transmitting the raw compressed images. In their implementation, they evaluated various image compression standards. In contrast, we proposed to detect the ROIs i.e. the digits in the images and then transmit those ROIs along with the run length codes to the server. Compare to the results in [1], we achieved much better compression performance, in addition to much reduced processing complexity. The interested readers are referred to details of Meter Reading application in [1] and [18].

### B. Applications attributing too few objects

LED light detection has been used for object localization. In both magnetic particle detection and LED light detection, the frames usually contain very few objects. We applied BVC to these applications to analyse its performance for applications attributing too few objects in the frames. The interested readers are referred to the details of node localization using LED Light detection in [18] and [19].

### C. Applications with too much transitions in the pixels

We analysed the application which have too many transitions from black to white and white to black pixels in the images. The selected application is the Human Detection, where due to illumination noise and the movement of the object, there are too many transitions in the change frames.

## V. EXPERIMENTAL RESULTS

This section is comprised of the results based on applying our proposed method to the three categories of real life machine vision applications. The Libtiff library in [17] is compiled and the execution file is generated for performing Group 4 compression scheme. The change frames and the ROIs in the change frames are determined using Matlab. The Group 4 is used to compress the original frames, the change frames and the detected ROIs and the results are shown in tables, where we discussed the remarkable points.

The compressed file size for the image coding, change coding, ROI coding and BVC is shown in Table I for Human Detection. We analysed two data sets (measurements), one where the human is in vertical position and the other in which the human is lying down on the earth. It must be observed in Table I that the compression efficiency of image coding is better than that of change coding and ROI coding. The reason for this is that there are too many transitions in the change frames of this application. So, in such applications, the output of the image coding is selected for BVC.

TABLE I. EFFICIENCY OF BVC FOR HUMAN DETECTION.

| Observations | Image Coding | Change Coding | ROI Coding | BVC |
|---|---|---|---|---|
| Measurement 1 | 474 | 705 | 585 | 474 |
| Measurement 2 | 472 | 577 | 484 | 472 |

The Table II shows the results of BVC for Meter Reading application. In Meter Reading, there are no movements involved, so the change frames have very few changes due to which the compression efficiency of change coding is better than image coding. Also in Meter Reading, there are too many objects (digits) in the frames (around 12–15) due to which the compressed file size of ROI coding is larger than that of change coding which must be observed in Table II. So, ROI coding is not a better option for such applications.

TABLE II: EFFICIENCY OF BVC FOR METER READING.

| Observations | Image Coding | Change Coding | ROI Coding | BVC |
|---|---|---|---|---|
| Measurement 1 | 508 | 287 | 487 | 287 |
| Measurement 2 | 524 | 289 | 507 | 289 |
| Measurement 3 | 502 | 283 | 487 | 283 |
| Measurement 4 | 496 | 284 | 475 | 284 |
| Measurement 5 | 498 | 284 | 479 | 284 |

TABLE III. EFFICIENCY OF BVC FOR LED LIGHT DETECTION.

| Observations | Image Coding | Change Coding | ROI Coding | BVC |
|---|---|---|---|---|
| Measurement 1 | 258 | 271 | 221 | 221 |
| Measurement 2 | 268 | 271 | 229 | 229 |
| Measurement 3 | 296 | 330 | 259 | 259 |
| Measurement 4 | 344 | 370 | 311 | 311 |
| Measurement 5 | 320 | 341 | 285 | 285 |

TABLE IV. EFFICIENCY OF BVC FOR MAGNETIC PARTICLE DETECTION.

| Observations | Image Coding | Change Coding | ROI Coding | BVC |
|---|---|---|---|---|
| Measurement 1 | 266 | 264 | 231 | 231 |
| Measurement 2 | 294 | 289 | 257 | 257 |
| Measurement 3 | 298 | 272 | 259 | 259 |
| Measurement 4 | 316 | 284 | 277 | 277 |
| Measurement 5 | 348 | 282 | 307 | 282 |

The compressed file size for the image coding, change coding, ROI coding and BVC for LED Light Detection and Magnetic Particle Detection is shown in Table III and Table IV respectively. It must be observed in Table III and Table IV that the compression efficiency of ROI coding is better than that of both image coding and change coding. The reason is that there are very few objects in the change frames in such applications. Hence for such applications, the output of ROI coding is selected as the output of the BVC.

## VI. CONCLUSIONS

The processing complexity and the compression efficiency of the compression methods affect the total energy consumption of the Visual Sensor Node (VSN). Optimization in both computation and communication energy consumption of the VSN is required for the energy constrained outdoor embedded applications of Wireless Visual Sensor Networks (WVSN). The compression efficiency of both change coding and Region of Interest (ROI) coding is better than that of image coding for applications involving too few objects in the change frames but it becomes worse than that of image coding if the change frame contains too many objects. In this paper, we analyzed the compression efficiency of the Bi-Level Video Codec (BVC) for various real life machine vision applications. In BVC, we proposed to implement all the three information reduction techniques i.e. image coding, change coding and ROI coding at the VSN and to select the smallest of the three bit stream as the final output. We observed that the compressed file size of the BVC is always smaller or equal to that of change coding, ROI coding and image coding. In contrast, the compression efficiency of both change coding and ROI coding becomes worse than that of image coding for applications attributing frames with too many objects. Thus we conclude that BVC is better than or at least equal to that of change coding, ROI coding and image coding and can be used for all of applications i.e. applications with too many or too few objects in the frames.

## REFERENCES

[1] L. Ferrigno, S. Marano, V. Paciello, A. Pietrosanto, "Balancing computational and transmission power consumption in wireless image sensor networks", in *Proc. of the IEEE Int'l Conf. on Virtual Environments, Human-Computer Interfaces, and Measurement Systems*, Italy, July 2005.

[2] P. Kulkarni, D. Ganesan, P. Shenoy, Q. Lu, "SensEye: A multi-tier camera sensor network", in *Proc. of the 13th Annual ACM int'l conference on Multimedia*, 2005, pp. 229–238. [Online]. Available: http://dx.doi.org/10.1145/1101149.1101191

[3] W. Chen, A. Hamid, K. Richard, "Real-Time Human Posture Reconstruction in Wireless Smart Camera Networks", in *Proc. of the 7th int'l conf. on Information processing in sensor networks (IPSN 08)*, IEEE Computer Society Washington, DC, USA, 2008, pp. 321–331.

[4] K. Khursheed, M. Imran, M. O'Nils and N. Lawal, "Exploration of Local and Central Processing for a Wireless Camera Based Sensor Node", in *Proc. of the IEEE Int'l. Conf. Signal Elec. Syst.*, Gliwice, Poland, Sept. 2010.

[5] K. Khursheed, M. Imran, A. W. Malik, M. O'Nils, N. Lawal, T. Benny, "Exploration of Tasks Partitioning Between Hardware Software and Locality for a Wireless Camera Based Vision Sensor Node", in *Proc. of the Int. Symposium on Parallel Computing in Electrical Engineering (PARELEC 2011)*, 2011. [Online]. Available: http://dx.doi.org/10.1109/PARELEC.2011.21

[6] M. Imran, K. Khursheed, M. O'Nils, N. Lawal, "Exploration of Target Architecture for a Wireless Camera Based Sensor Node" in *Proc. of the 28th Norchip Conference*, Tampere, Finland.

[7] K. Khursheed, M. Imran, N. Ahmad, M. O'Nils, "Selection of bi-level image compression method for reduction of communication energy in wireless visual sensor networks", *in Proc. of the SPIE 8437*, 2012. [Online]. Available: http://dx.doi.org/10.1117/12.923716

[8] F. Ono, W. Rucklidge, R. Arps, C. Constantinescu, "JBIG2 - The ultimate bi-level image coding standard". *in Proc. of the IEEE Intl. Conf. on Image Processing (ICIP)*, Vancouver, Canada, Sept. 2000.

[9] *The link for T.6 i.e. CCITT Group 4.* [Online]. Available: http://www.itu.int/rec/T-REC-T.6-198811-I

[10] *Gzip pack.* [Online]. Available: http://www.gzip.org/

[11] K. Khursheed, M. Imran, N. Ahmad, M. O'Nils, "Analysis of Change Coding for Data Reduction in Wireless Visual Sensor Network", *in IEEE Transactions on Industrial Informatics,* submitted for publication.

[12] K. Khursheed, N. Ahmad, M. Imran, M. O'Nils, "Detecting and Coding Region of Interests in Bi-Level Images for Data Reduction in Wireless Visual Sensor Network", in *Proc. of the 1st International Workshop on Wireless Multimedia Sensor Networks*, submitted for publication.

[13] K. Khursheed, N. Ahmad, M. Imran, M. O'Nils, "Binary Video Codec for Data Reduction in Wireless Visual Sensor Networks", *in Proc. of the SPIE Conf. Real Time Image and Video Processing*, submitted for publication.

[14] R. B. Arps, T. K. Truong, "Comparison of international standards for lossless still image compression", in *Proc. of the IEEE*, vol. 82, no. 6, pp. 889–899, June 1994. [Online]. Available: http://dx.doi.org /10.1109/5.286193

[15] J. Ziv, A. Lempel, "A universal algorithm for sequential data compression", *IEEE Trans. on Information Theory,* vol. IT-23, no. 3, May 1977.

[16] S. Mohamed, A. Fahmy, "Binary Image Compression Using Efficient Partitioning into Rectangular Regions", *IEEE Trans. on Comm.*, vol. 43, no. 5, May 1995.

[17] *The Libtiff library.* [Online]. Available: http://www.libtiff.org/libtiff.html

[18] S. Hengstler, H. Aghajan, "Application Development in Vision-Enabled Wireless Sensor Networks", in *Proc. of the Int. Conf. Systems and Networks Communications*, 2006.

[19] H. Lee, H. Dong, H. Aghajan, "Robot-assisted localization techniques for wireless image sensor networks", in *Proc. 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*, 2006, vol. 1, pp. 383–392.