Distance Based Heuristic for Power and Rate Allocation of Video Sensor Networks

Bambang A.B. Sarif, Victor C.M. Leung, Panos Nasiopoulos
Department of Electrical and Computer Engineering
The University of British Columbia
Vancouver, BC, Canada V6T 1Z4
{bambangs, vleung, panos}@ece.ubc.ca

***Abstract*—Video sensor networks (VSNs) offer an alternative to several existing surveillance technologies. However, unlike in conventional sensor network, video processing and transmission requires large amount of resources both in signal processing, i.e., encoding, and transmission of the encoded data. For such networks, an optimal encoding power and rate allocation method based on a power-rate-distortion (PRD) analysis has previously been proposed, where the power consumption of video encoding can be controlled by managing some encoding parameters. However, these parameters are currently obtained offline by examining the stored video, an approach which may not be suitable for surveillance applications. In this paper, a distance based heuristic for encoding power and rate allocation of VSNs is proposed. The proposed technique is a practical solution since the video coding parameter is controlled by the node's location in the network. Although the proposed technique offers a sub-optimal solution, in some scenarios it achieves performance up to 94% of the optimal solution in terms of network lifetime.**

Keywords—Wireless sensor network; video sensor network; power and rate allocation; energy efficiency; network lifetime

I. INTRODUCTION

Recently, there is a growing interest in incorporating video applications and transmissions over wireless sensor networks (WSNs) [1-3]. While visual sensing is not generally well-supported by WSN architectures, visual information may significantly improve the perceived information gathered from the sensed environment. This is especially true for low-cost monitoring and surveillance applications [4-6].

Wireless video sensor networks (VSN) have the potential to improve the ability to develop user-centric surveillance applications to monitor and prevent harmful events. The availability of inexpensive low-power sensors, radios, and embedded processors enables the deployment of distributed sensor networks to offer information to users in distinct environments and to provide them control over undesirable situations. Networked sensors can collaborate to process and make deductions from the collected data and provide the user with access to continuous or selective observations of the environment [7, 8]. In addition, video based sensor networks offer an alternative to several existing technologies. The wiring costs restrict complicated environment controls and the reconfigurability of these systems. In many cases, the savings in the wiring costs alone justify the use of a VSN.

Sensor nodes are often battery powered, and battery replacement may be infrequent, not preferable, or even impossible in many applications. Tremendous amount of research effort has focused on energy conservation. In conventional WSNs, data processing performed by sensor nodes is assumed to be very simple. Thus, the energy consumption utilized for data collection and processing is often negligible. On the other hand, video processing and transmission requires a large amount of resources both in signal processing, i.e., encoding, and transmission of the encoded data. Thus it is very challenging to prolong the network lifetime of a VSN. A number of algorithms to maximize a sensor network's lifetime have been proposed in the literature [12, 14, 15]. The technique proposed in [12] claims to provide the optimal allocation of encoding power and source rate for a VSN which is based on the power-rate-distortion (PRD) model proposed in [9, 10]. It is assumed that each node can control the power consumption of the encoding process. This can be done by optimizing three video coding parameters: number of frames, number of skipped blocks, and number of examined blocks for motion estimation [9]. However, to the best of our knowledge, these parameters are currently obtained offline by examining the stored video. Thus, it is not practical for real-time monitoring or surveillance systems.

In some VSN applications such as structural monitoring or surveillance system, it is often required to transmit the compressed video data through a large area. In this regard, experimental results suggest that for the same required video distortion, nodes that are closer to the sink should spend less power encoding than the farther ones [12].

In some video coding approaches, i.e., the intra-coding mode, the average power dissipation to encode one frame is generally constant. Thus, we suppose that to achieve a given target distortion, encoding power consumption can be controlled by managing the video frame rate on each node based on their distance from the sink. Hence, a distance based heuristic for encoding power and rate allocation of a video sensor network is proposed in this paper. The rest of the paper is organized as follows. The network lifetime maximization problem is explained in Section II. Section III describes the proposed heuristic. Experimental results are presented in Section IV. Section V concludes the paper.

This work is supported by a BCFRST Natural Resources and Applied Sciences (NRAS) Endowment award and the Institute for Computing, Information and Cognitive Systems (ICICS) at UBC.

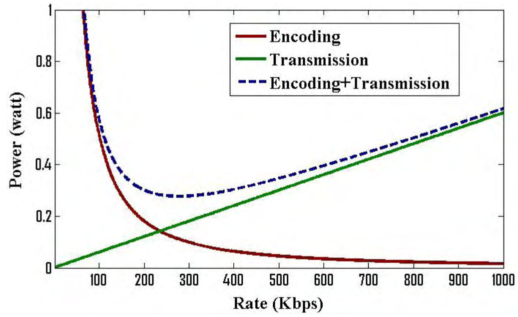

Figure 1 Encoding and transmission power consumption

II. Network Lifetime Maximization of Video Sensor Networks

The distortion of a compressed video depends on the source rate R and encoding power consumption P. According to the PRD model in [9], the encoding distortion D is computed by:

$$D = \sigma^2 e^{-\gamma \cdot R \cdot P^{2/3}} \tag{1}$$

where σ^2 is the average input variance and γ is the encoding efficiency coefficient. To achieve the given encoding distortion requirement, one can either increase the encoding power or source rate. Obviously, there is a trade-off where optimal power consumption of a source node can be achieved, as shown in Fig. 1. This figure is generated using parameters mentioned in [12] and video distortion of 100.

The following video sensor node model is used in this paper. Battery powered sensor nodes are statically deployed in a square area characterized by its width. It is assumed that a standard medium access control (MAC) protocol is applied to resolve the link interference problem. The network is modeled as a undirected graph $G(N,L)$ where N is the set of nodes and L is the set of links. Node i can communicate with node j if a link between those nodes ($L_{ij} \in L$) exists. Sensor node i can capture and encode video, and then generate data traffic with a source rate R_i. Furthermore, each node can also relay the traffic from upstream nodes. The flow conservation law at each node is:

$$\sum r_{ij} - \sum r_{ki} = R_i \tag{2}$$

Here r_{ij} denotes the outgoing rate at L_{ij} while r_{ki} denotes incoming rates at L_{ki}, and L_{ij} , $L_{ki} \in L$.

The total power dissipation at node i consists of the transmission power consumption, the reception power consumption and the encoding power consumption:

$$P_i = P_{ti} + P_{ri} + P_{ei} \tag{3}$$

A general energy consumption model for a wireless communication transmitter and receiver as presented in [13] 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.

$$P_{ti} = \sum \left(\alpha + \beta \cdot d_{ij}{}^{\eta}\right) \cdot r_{ij} \tag{4}$$

Here, α and β are constant coefficients, η is the path loss exponent, and $L_{ij} \in L$. Also, the total reception power consumption of node i is the sum of all power consumed to receive data from other nodes, as formulated below, where λ is a constant coefficient and $L_{ki} \in L$:

$$P_{ri} = \sum \lambda \cdot r_{ki} \tag{5}$$

The encoding power consumption of node i can be determined using the PRD model [9]. For a given distortion requirement, the relation between encoding power consumption and source rate can be stated as follows:

$$P_{ei}^{2/3} \cdot R_i = -\log\left(D_i / \sigma^2\right) / \gamma \tag{6}$$

In this equation, D_i is the given distortion requirement for node i, σ^2 is the average input variance, and γ is the encoding efficiency coefficient.

It is assumed that the node is powered by a battery with initial amount of energy B_i. The lifetime of node i is equal to:

$$T_i = B_i / P_i \tag{7}$$

Minimum node lifetime (T_{net} = min T_i, $i \in N$), i.e., the network fails even if only one of the nodes runs out of energy, is used as the measure of network lifetime in this paper. Since maximizing T_{net} = min T_i is equivalent to maximizing T_{net} subject to $T_{net} \le T_i \quad \forall i$, the objective function can be reformulated as a linear program by introducing variables $q_{net} = 1/T_{net}$, $q_i = 1/T_i$, such that $P_i = B_i \cdot q_i$, and $q_i \le q_{net}$.

The lifetime maximization of the VSN is formulated as a constrained optimization problem as shown in Fig. 2. This algorithm is similar to the centralized power-rate optimization proposed in [12] and for this reason we call it the Power Rate Optimal – Video Sensor Network (PRO-VSN) algorithm.

minimize q_{net}

subject to:

$$q_i \le q_{net},\ \forall i \in N$$
$$\sum r_{ij} - \sum r_{ki} = R_i, \forall i \in N\ \forall j \in N\ \forall k \in N$$
$$-\log\left(D/\sigma^2\right)/(\gamma \cdot P_{ei}^{2/3}) = R_i, \forall i \in N$$
$$\sum\left(\alpha + \beta \cdot d_{ij}{}^{\eta}\right) \cdot r_{ij} + \sum \lambda \cdot r_{ki} + P_{ei} = B_i \cdot q_i, \forall i \in N$$
$$q_i > 0, \forall i \in N$$
$$R_i > 0, \forall i \in N$$
$$P_{ei} > 0, \forall i \in N$$

Figure 2. Optimization in PRO-VSN

III. Proposed Distance Based Heuristic

Results from the implementation of several video encoding algorithms on the Stargate platform show that the average energy consumption of H.264 intra-frame coding for QCIF (Quarter Common Intermediate Format) resolution video is approximately 60mJ/frame, while the baseline H.264 with motion compensation consumes approximately 760mJ/frame. The energy consumption for transmission is approximately 0.9mJ/frame for H.264 intra and 0.2mJ/frame for H.264 with motion estimation coded video data [11].

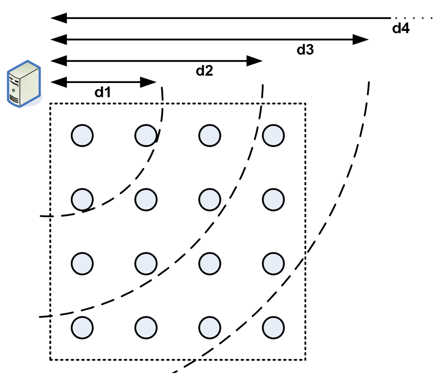

Figure 3. Sensor network's zones

Thus, according to this result and unless we have a very large system in which packets need to be transmitted on a large number of hops, intra-frame coding algorithm is the most suitable option for low energy sensor motes. In the intra-coding mode, the average power dissipation for encoding one frame is generally constant. Since a node's encoding power can be estimated based on its distance from the sink, the network is divided into different zones (Z) as shown in Fig. 3. All nodes in the same zone will dissipate the same encoding power and transmit with the same source rate to achieve the given distortion requirement.

A zone is defined as a section of the network covering an area whose distance to the sink is greater than r_{m-1} and smaller or equal to r_m, $1 \le m \le z_{opt}$. The number of zones (z_{opt}), $1 \le z_{opt} \le z_{max}$, should be carefully chosen as to maximize network lifetime. On the other hand, the value of z_{max} depends on the value of the given distortion requirement and the hardware's ability, as will be explained in the following paragraphs. In Fig. 3, a sensor node i that is located at a distance of d_i from the sink is a member of Z_m if $r_{m-1} < d_i \le r_m$. Each node has to become a member of exactly one zone. However, a zone may have zero member.

As stated earlier, for the same required video distortion, encoding power of nodes that are closer to the sink should be lower than the ones farther. Consequently, the source rate of nodes closer to the sink will be higher than the ones farther. Thus, $P_1^z < P_2^z < \ldots P_m^z \ldots < P_{z_{\max}}^z$ and $R_1^z > R_2^z > \ldots R_m^z \ldots > R_{z_{\max}}^z$, where P_m^z and R_m^z are the encoding power and source rate of any node in zone m. Using the analytical PRD model of [9], we can also derive the following:

$$R_m^z \cdot \left(P_m^z\right)^{2/3} = R_1^z \cdot \left(P_1^z\right)^{2/3} \tag{8}$$

Since the average power dissipation for encoding one frame is assumed to be constant, the encoding power consumption for any node in the first zone can be calculated as follows:

$$P_1^z = P_{1frame} \cdot f_1^z \tag{9}$$

Here P_{1frame} is the average power dissipation for encoding one frame while f_1^z is the frame rate of any node in the first zone. Using (8) and (9), we can calculate the encoding power dissipation and source rate for zone m:

$$P_m^z = P_{1frame} \cdot (f_1^z + m - 1) \tag{10}$$

$$R_m^z = R_1^z \cdot \left(f_1^z / \left(f_1^z + m - 1\right)\right)^{2/3} \tag{11}$$

The value of f_1^z depends on the target distortion requirement and an estimated value of source rate in zone one. Since $R_1^z \ge R_m^z$, the source rate in zone one can be estimated as follows:

$$\hat{R}_1^z = (1+\delta) \cdot W / V \tag{12}$$

Here W is the bandwidth of the network, δ is a small number and V is the number of video nodes. With D_1^z as the value of target distortion for zone one, the values of f_1^z and R_1^z can then be obtained as follows:

$$f_1^z = \left\lceil \left(\frac{-V \cdot \log\left(D/\sigma^2\right)}{\gamma \cdot (1+\delta) \cdot W} \right)^{1.5} / P_{1frame} \right\rceil \tag{13}$$

$$R_z^1 = -\log\left(D_1^z / \sigma^2\right) / \left(\gamma \cdot \left(f_z^1 \cdot P_{1frame}\right)^{2/3}\right) \tag{14}$$

Note that, unless it is mentioned specifically, we assume that the target distortion for all nodes is the same, thus $D_1^z = D_i = D$.

The value of z_{max} is obtained by considering the capability of the central processing unit (CPU) to process the captured picture. The experimental results on encoding power consumption reported in [11] were obtained by calculating the number of cycles required to process a video source. Based on these results, the number of cycles required to process one frame is approximately 20 million cycles. Considering that the maximum CPU frequency of the Stargate platform is 400 MHz, the maximum number of frames that can be processed in one second is approximately 16. Thus, the value of z_{max} can be calculated as: $z_{\max} = 17 - f_1^z$.

Applying the distance based heuristic, maximization of the network lifetime of a VSN is then formulated as the constrained optimization problem shown in Fig. 4. We call this approach the Power-Rate-hEuristic Video Sensor Network (PRE-VSN) algorithm.

minimize q_{net}

subject to:

$$q_i \le q_{net}, \ \forall i \in N$$

$$\sum r_{ij} - \sum r_{ki} = R_i, \forall i \in N \ \forall j \in N \ \forall k \in N$$

$$R_1^z \cdot \left(f_1^z / \left(f_1^z + m - 1\right)\right)^{2/3} = R_i, \ i \in Z_m, 1 \le m \le z_{opt}$$

$$\left\lceil \left(\frac{-V \cdot \log\left(D/\sigma^2\right)}{\gamma \cdot (1+\delta) \cdot W} \right)^{1.5} / P_{1frame} \right\rceil = f_1^z$$

$$-\log\left(D_1^z / \sigma^2\right) / \left(\gamma \cdot \left(f_z^1 \cdot P_{1frame}\right)^{2/3}\right) = R_z^1$$

$$P_{ei} / (f_1^z + m - 1) = P_{1frame}, \ i \in Z_m, \ 1 \le m \le z_{opt}$$

$$\sum \left(\alpha + \beta \cdot d_{ij}^{\ \eta}\right) \cdot r_{ij} + \sum \lambda \cdot r_{ki} + P_{ei} = B_i \cdot q_i, \forall i \in N$$

$$q_i > 0, \forall i \in N$$

Figure 4. Optimization in PRE-VSN

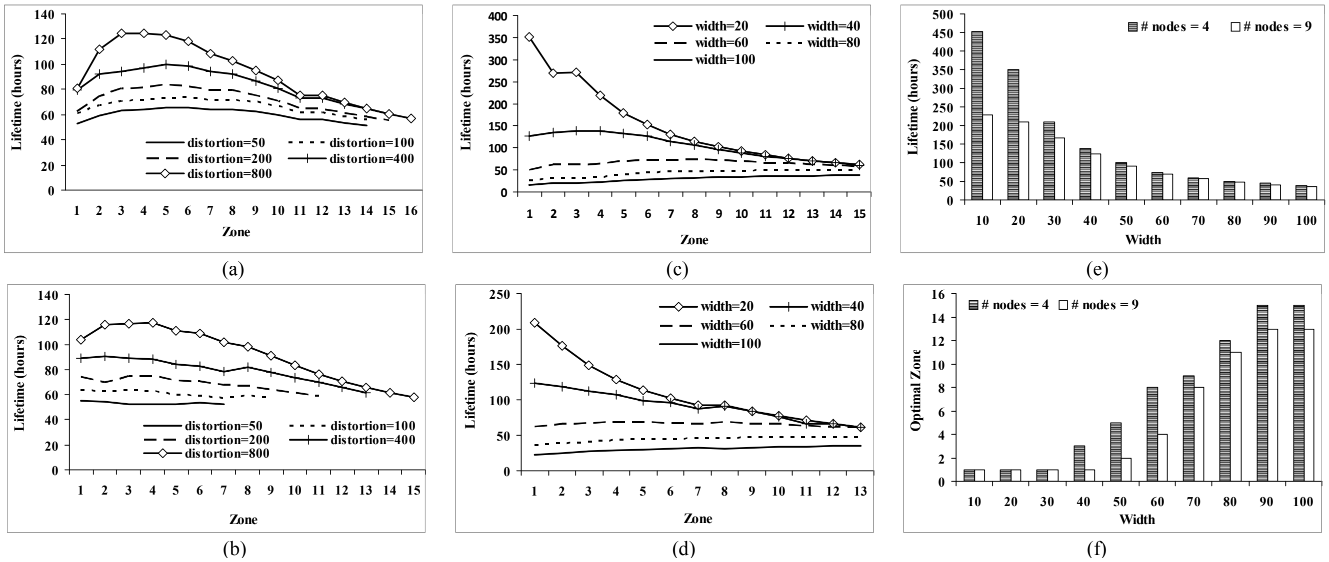

Figure 5. Performance of the proposed heuristic. (a) varying distortion, node=4, width=50m; (b) varying distortion, node=9, width=50m; (c) varying width, node=4, distortion=400; (d) varying width, node=9, distortion=400; (e) maximum network lifetime; (f) the number of zone producing maximum lifetime (z_{opt}).

IV. Experiments and Results

The proposed algorithm is implemented in MATLAB. It should be noted that the optimization of PRE-VSN includes an integer variable in the form of z_{opt}. Solving a mixed integer linear problem is known to be intractable. However, since the value is z_{max} is relatively small, we implement the optimization by performing a loop for all possible z_{opt} values. We hope that a trend or relation between the number of zones and network width can be deduced from our experiments so that we can perform a much better procedure in the future.

Table I shows the parameters used in the experiments. It is assumed that the maximum transmission range is 100m and the network bandwidth is 1 Mbps. It is clear from (4) that the transmission power consumption depends on the transmission distance. Accordingly, we assume that each node is able to adjust its transmission power to achieve the required performance at the receiver. The term network width denotes the width of the square area in which sensor nodes are statically deployed. In some of the experiments, the network width used is smaller than nodes' transmission range. However, since the proposed algorithm tries to maximize minimum lifetime using the distance-based heuristic and rate optimization, each node is assumed to be able to relay some of its packets through intermediate nodes. The rate adjustment coefficient is set according to the number of zones in the corresponding experiment. δ is set equal to 0 when the number of zones is 1 and δ_{max} when the number of zones is equal to z_{max} respectively.

TABLE I. PARAMETERS USED

Parameters	Description	value
B	Initial energy (battery)	100 kJ
α	Energy cost for transmitting 1 bit	0.5 J/Mb
β	Transmit amplifier coefficient	$1.3{\cdot}10^{-8}$ J/Mb/m4
λ	Energy cost for receiving 1 bit	0.5 J/Mb
η	Path loss exponent	4
σ^2	Average variance of video (MSE)	3500
γ	Encoding efficiency coefficient	55.54 $W^{3/2}$ /Mbps
δ_{max}	Max. rate adjustment coefficient	0.1
P_{1frame}	Power consumed to encode 1 frame	60 mJ/frame

Performance of the proposed heuristic in terms of network lifetime and optimal zone is shown in Fig. 5. For these results, the sensor nodes are deployed uniformly in a square-grid layout where each node is deployed at the center of each grid. In Fig. 5(a), the number of nodes is 4 and the network width is 50m. Distortion requirement is varied from 50 to 800. This graph shows some interesting insight. Firstly, the value of z_{max} is smaller when the distortion requirement is small, i.e., z_{max} = 14 when distortion is 50, which means that f_1^z=3 fps. Secondly, the network lifetime graph for low distortion requirement is rather flat. This means that most of the power is consumed for encoding the captured video to achieve the target distortion. A similar result is shown in Fig. 5(b) where the number of nodes is increased to 9.

In Fig. 5(c), the number of nodes is 4 and the target distortion is 400 while the network width is varied from 10m to 100m. Similar settings are used in Fig. 5(d) where the number of nodes is increased to 9. It can be seen that the value of the optimal number of zones is proportional to the network width. When the network is large, the source rate should be small to reduce communication cost. To compensate for the reduction in source rate, sensor nodes need to dissipate higher encoding power to achieve the given target distortion. Fig. 5(e) shows the maximum network lifetime for these experiments. Fig. 5(f) shows the optimal number of zones to achieve those results.

In the techniques proposed in [14, 15], all nodes assume the same encoding power consumption regardless of their location. These approaches are actually a special case of the proposed distance based heuristic where the number of zones is equal to 1. We call this algorithm Power-Rate-Allocated Video Sensor Network (PRA-VSN).

To compare the proposed approach with other existing techniques, sensor nodes are randomly deployed within an area whose width is varied between 10m to 100m. The experiments are repeated 10 times and the distortion requirement is set equal to 100. Two performance merits, normalized lifetime and lifetime offset, are used. Normalized lifetime is defined as the ratio of network lifetime obtained from the examined algorithms with the optimal one produced by PRO-VSN. Lifetime offset is the difference between the optimal network lifetime produced by PRO-VSN and the one obtained from the tested algorithms. The results are shown in Table II and Fig. 6.

It can be seen that the proposed PRE-VSN performs better than PRA-VSN when the network width is greater or equal to 60m. Furthermore, the lifetime offset of PRE-VSN can be as little as 1.78 hours less than that achieved by PRO-VSN when the network width is greater or equal to 40m. These results show that the proposed distance based heuristic can achieve a comparable performance to that obtained by the optimal solution. However, when the network width is smaller than or equal to 30m, PRE-VSN's lifetime is at least 4 hours shorter than that of PRO-VSN. PRO-VSN performs much better in this case because the encoding power is finely optimized. However, since PRO-VSN does not use the proposed zone concept, the allocated encoding power of PRO-VSN cannot be translated easily into video coding parameters which may not be suitable for a real time system. Fig. 7 shows average power consumption for each node when the network width is 60m and target distortion is 100. It is shown that each algorithm is able to control power dissipation so that all nodes spend similar total power consumption. It can also be seen that in PRA-VSN all nodes dissipate similar encoding power. However, it is clear that PRE-VSN performs better than PRA-VSN, while PRO-VSN performs the best in term of minimum total power consumption.

TABLE II. NETWORK LIFETIME OFFSET

Width (m)	Lifetime Offset (hours)		Width (m)	Lifetime Offset (hours)	
	PRA-VSN	*PRE-VSN*		*PRA-VSN*	*PRE-VSN*
10	5.81	5.81	60	2.82	2.09
20	6.97	6.97	70	3.49	2.33
30	4.13	4.13	80	4.08	2.21
40	2.25	2.25	90	4.13	1.78
50	2.13	2.13	100	4.27	2.04

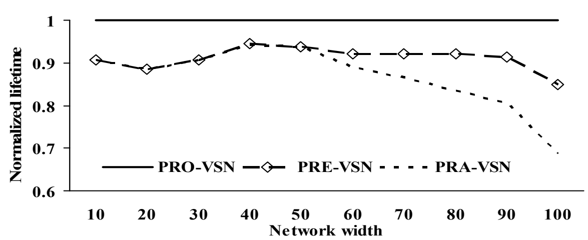

Figure 6. Average network lifetime, normalized to PRO-VSN

V. CONCLUDING REMARKS

A distance based heuristic for encoding power and rate allocation of VSNs is proposed. It is based on the fact that in order to achieve a given target distortion, encoding power consumption can be controlled by managing the video frame rate at each node, based on the node's distance from the sink. It can be deduced from the experimental results that the value of the optimal number of zones is proportional to the network width. Performance evaluations also show that the proposed heuristic can achieve quite a comparable performance to that obtained by the optimal solution when the network width is greater than or equal to 40m and the distortion requirement is 100. The proposed technique may offer some practical benefit in a sense that the frame rate can be estimated according to the node's physical location in the network, without the need to examine the captured video. We are currently working on improving the performance of the heuristic by fine-tuning other parameters.

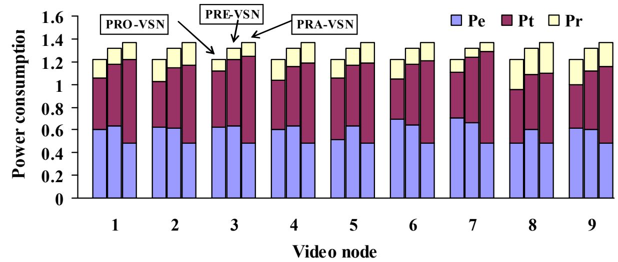

Figure 7. Average power consumption profile

REFERENCES

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

[2] I. Akyildiz, T. Melodia, and K. Chowdhury, "A survey on wireless multimedia sensor networks," Computer Networks, vol. 51(4), 2007, pp. 921—960.

[3] A. Kandhalu, A. Rowe, and R. Rajkumar, "DSPcam: a camera sensor system for surveillance networks," IEEE International Conference on Distributed Smart Cameras (ICDSC), 2009.

[4] W.T. Chen, P.Y. Chen, W.S. Lee, and C.F. Huang, "Design and implementation of a real time video surveillance system with wireless sensor networks," IEEE Vehicular Tech. Conf., 2008, pp. 218 – 222.

[5] J. C. Yang and C. L. Lai, "Vision based fire/flood alarm surveillance system via robust detection strategy," IEEE Int. Instrumentation and Measurement Technology Conference, I2MTC 2008.

[6] N. Xu, S. Rangwala, K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A wireless sensor network for structural monitoring," Proc. of the Second International Conference on Embedded Network Sensor Systems, 2004, pp. 13–24.

[7] T.D. Raty, "Survey on contemporary remote surveillance systems for public safety," IEEE Trans. on System, Man and Cybernetics – Part C: Applications and Reviews, vol. 40, no. 5, 493—515, 2010.

[8] J. Chen, Z. Safar, and J. A. Sorensen, "Multimodal wireless networks: communication and surveillance on the same infrastructure," IEEE Trans. Inf. Forensics Secur., vol. 2, no. 3, Sep. 2007, pp. 468–484.

[9] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, "Power-rate-distortion analysis for wireless video communication under energy constraints", IEEE Trans. on Circuits and Sys. for Video Tech., vol. 15, no. 5, pp. 645 - 658, 2005.

[10] Zhihai He, Dapeng Wu, "Resource allocation and performance analysis of wireless video sensors", IEEE Transactions on Circuits and Systems for Video Technology, vol.16, no.5, pp.590 -- 599, 2006.

[11] J. J. Ahmad, H. A. Khan, and S. A. Khayam, "Energy efficient video compression for wireless sensor networks," 43rd International Conference on Information Sciences and Systems (CISS), 2009.

[12] Y. He, I. Lee, and L. Guan, "Distributed algorithms for network lifetime maximization in wireless visual sensor networks", IEEE Trans. on Circuits and Sys. for Video Tech., Vol.19, Iss.5, pp.704 --718, 2009.

[13] T.S. Rappaport, Wireless communications: principles and practice, Prentice Hall PTR Upper Saddle River, NJ, USA, 2001.

[14] J.H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," IEEE/ACM Trans. Netw., vol. 12, no. 4, pp. 609–619, August 2004.

[15] S. Soro and W. Heinzelman, "On the coverage problem in video-based wireless sensor networks," in Proc. Int. Conf. Broadband Netw., vol. 2, Boston, MA, Oct. 2005, pp. 932—939.