Aggregation In Wireless Multimedia Sensor Network

By

Archana Kurde

11MCEC20

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AHMEDABAD-382481

May 2013

Aggregation In Wireless Multimedia Sensor Network

Major Project(Part-II)

Submitted in partial fulfillment of the requirements

For the degree of

Master of Technology in Computer Science and Engineering

By

Archana Kurde

(11MCEC20)

Guided By

Prof. Vijay Ukani

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AHMEDABAD-382481

May 2013

Certificate

This is to certify that the Major Project Report entitled **"Aggregation In Wireless Multimedia Sensor Network"** submitted by **Ms. ARCHANA KURDE (Roll No. 11MCEC20)**, towards the partial fulfillment of the requirements for the award of Degree of Master of Technology in Computer Science & Engineering of Nirma University of Science and Technology is the record of work carried out by him under my supervision and guidance. The work submitted has in my opinion reached a level required for being accepted for examination. The results embodied in this major project work to the best of my knowledge have not been submitted to any other University or Institution for award of any degree or diploma.

Prof. Vijay Ukani
Guide
Institute of Technology,
Department of Computer Science & Engineering,
Nirma University,
Ahmedabad,

Dr. Sanjay Garg
Head and Professor,
Department of Computer Science & Engineering,
Institute of Technology,
Nirma University, Ahmedabad

Dr K Kotecha
Director,
Institute of Technology,
Nirma University,
Ahmedabad

Abstract

Image transmission in WMSN is expensive due to its large image size and significant energy consumption. One possible way to reduce the image size is to aggregate the received images on the sensor nodes before transmission. By eliminating the redundant image, image aggregation in WMSN could transmit the overlapped region of the images only once to reduce the total energy consumption. Image aggregation requires image processing at the aggregator node that incurs additional node processing energy. At the same time image aggregation scheme helps in reducing the communication energy by reducing the amount of data transmitted. As compared to the non-aggregation scheme, image aggregation scheme consumes only 46 percent of the energy consumption.

An energy efficient communication protocol is proposed which works by leveraging the correlation of visual information of camera sensors. A spatial correlation coefficient is derived to describe the degree of correlation of visual information observed by cameras with overlapped field of views.

We make use of correlation between captured images in the network scheduling process and propose a correlation based scheduling algorithm to maximize network lifetime of WMSN.

Acknowledgments

I owe a huge debt of gratitude to my guide, **Prof. Vijay Ukani**, PG Co-ordinator, Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad for the valuable inputs at every step of this work. The appreciation and continual support he has imparted has been a great motivation to me in reaching a higher goal. His in-depth knowledge about the subject has provided me a strong base and direction towards this work. His guidance has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come.

It gives me an immense pleasure to thank **Dr. Sanjay Garg**, Hon'ble Head of Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his kind support and providing basic infrastructure and healthy research environment.

A special thank you is expressed wholeheartedly to **Dr. Ketan Kotecha**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout course of this work.

I would also thank the Institution, all the faculty members in Department of Computer Science and my colleagues without whom this project would have been a distant reality.

I would like to thank my colleagues Samir Kariya, and Sameer Kapadia for being with me in any condition. And I am thanking them to keep me calm and cool every time.

Last, but not the least, no words are enough to acknowledge constant support and sacrifices of my family members because of whom I am able to complete the first part of my dissertation work successfully.

- Archana Kurde

11MCEC20

Contents

List of Figures

Chapter 1

Introduction

A wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to a main location. Wireless sensor networks (WSNs) composed of tiny low-power sensors that cooperatively operate in an ad hoc manner have fostered a series of innovative monitoring and control applications.

1.1 Introduction to WSN

Nodes in WSNs are disposable electronic devices commonly equipped with a transceiver, a limited energy supply, a sensing unity and memory and processing resources, although additional modules can be found, such as a Global Positioning System. Compared with other traditional network, sensor network has the following features:

1. Resource-constrained: Because of limited cost, size and power consumption, the computing power, program space and memory space of nodes are much weaker and smaller than common computer.

2. Large-scale use: In order to finish regional monitoring missions, usually there are thousands of sensor nodes are densely deployed in the target area, so as to make use of the high degree of connectivity among nodes to assure the fault tolerance and invulnerability of system.

3. Self-organization structure. Network deployment and launch don't depend on any pre positioned network facility.

4. Multi-hop communications: In wireless network, communication distance between nodes is limited. Each node can only communicate with its neighbor. If the communication with nodes outside its radio frequency range is expected, information must be transferred through intermediate nodes by means of routing.

5. Strong dynamic: In network, nodes may quit from net operation because battery power runs out or other faults. Some nodes may be moved or new nodes can be added to the net because of mission requirements. All these can lead to changes of network topology.

6. Safe and reliable: Sensor network is applied to monitoring missions in harsh environment or security sensitive areas. Thus sensor network is required to have the ability to prevent monitoring data from being stolen and identify counterfeit monitoring information.

The challenges in the hierarchy of: detecting the relevant quantities, monitoring and collecting the data, assessing and evaluating the information, formulating meaningful user displays, and performing decision-making and alarm functions are enormous.[1]

1.2 Introduction to WMSN

The integration of low-power wireless networking technologies with inexpensive hardware such as complementary metal-oxide semiconductor (CMOS) cameras and microphones is now enabling the development of distributed, networked systems that we refer to as wireless multimedia sensor networks (WMSNs).

Harvesting scalar information from deployed sensors in a monitored field such as humidity, pressure, temperature, luminosity, seismic variations, among others, these applications have raised many challenges that have been extensively addressed.

The WMSN is totally different from traditional WSN in three ways. In WMSNs, interconnected sensor devices collaborate to retrieve multimedia information such as video and audio streams, still images, and scalar data from the environment. Apart from the primary goal of retrieving multimedia data, WMSNs will also be able to store, process in real time, correlate and fuse data originated from heterogeneous sources. These networks can be an integral part of systems such as security surveillance, traffic enforcement, health care delivery, environmental monitoring, and industrial process control.

1. New definition of data source nodes: FOV (field of vision) is to be considered so that the camera captures images of the event from a certain direction.

2. System resource consumption: Transmitting image in WMSN is challenging as the bandwidth and frequency spectrum are limited resource in WMSN.

3. Energy-efficient design: Besides the power for image transmission, the power for image processing should be addressed carefully to get power-efficient design in WMSN.

1.3 The key issue in video sensor network

1. **Emergency Saving Control Strategy :** The environment and condition of video sensor application doesn't allow a battery change, and its energy consumption is also larger than traditional sensor, we need to save the limited battery energy on each node and try to extend the life of the whole network.[2]

2. **Deployment and Coverage of Video Senor Nodes :** Traditional deployment and coverage optimization of sensor network is not applicable to video sensor nodes which have directional perception. So we have to employ a video sensor network coverage optimization that consider both energy efficiency and reliable monitoring, so as to extend the work life of network and provide sufficiently high monitoring coverage.

3. **Network Architecture of Video Sensor Nodes :** Existed simple sensor network structure based on cluster cannot perfectly process large data media services including images and video information. It is an important issue to make use of isomorphic or heterogeneous video sensor nodes to build network system structure which is energy efficient, expansible, flexible and fault tolerant.

4. **Real- Time Transmission :** Video information has high requirements for the delay and synchronization of transmission. We should design real-time, reliable transport protocol and strategy that are applicable to video sensor network, so as to save network energy and extend work life of network.

5. **Collaborative Processing Technology of Video Sensor Nodes :** A single node cannot separately process large-scale scene monitoring tasks. In order to give full play to sensor network's monitoring ability, we must coordinate lots of sensor nodes in the net with limited resources so as to efficiently solve the problem of complex media data's collecting, processing, transmission and exhibition.

1.4 Applications of WMSN

WMSN is mainly applied to environment monitoring. The main featured applications are the following[3]

- Battlefield monitoring: Video sensor network can realize many functions such as the monitoring of enemy troops and equipment, real-time monitoring of the battlefield, target locating, battle field assessment, etc .

- Traffic monitoring: Video sensor network can monitor the traffic of traffic hub, ring road and highway, so as to count vehicles passing by, whether there is illegal target parking and whether there are fault vehicles. It can also provide the latest situation about road congestion, recommend the best route, and remind the drivers to avoid accidents, etc.

- Security-sensitive area monitoring: Monitor security-sensitive working environments such as mines, power stations, coal mines, watch towers, etc.

- Intelligent home and target tracking: Building intelligent kindergarten to monitor children's early education environment and track children's activities can allow parents and teachers to make a comprehensive study of students' study and life process. And it can also be used to monitor people's behaviors.

- Public safety monitoring: Video sensor network can also be widely applied to security monitoring of public places like airports, railway stations, customs, stadiums, parking lots, business places, etc. It is also used for real-time monitoring and forecast of emergent environments.

1.5 Introduction to Data Aggregation

The smart nodes of WMSN are more powerful in function and are of smaller size. It can acquire and process not only simple information like temperature and humidity, but also complex information like video and image. As a result, in addition to basic characters of WSN, WMSN has two other important characteristics. First, the amount and sort of information acquired from environment is more plentiful than before. Second, the cost of one WMSN is more expensive than conventional WSN. In order to avoid waste of data resource and have a high return on investment, WMSN should be more customizable and reusable than conventional WSN. Therefore, WMSN should be data-centric network.[3]

The amount of data in WMSN is much huger than traditional WSN, so collaborative in-network processing is necessary to avoid transmitting large amount of raw streams to the sink. When redundant data is coped by collaborative in-network processing, Source coding technologies provide another way to decrease the amount of data which should be transmitted. This process of reduction of data that is to be transmitted is termed as data aggregation.

In WMSNs, multiple camera sensors are deployed to provide multiple views, multiple resolutions, and enhanced observations of the environment. There exists correlation among the visual information observed by cameras with overlapped field of views.

There are various types of In-network processing techniques. Few of them are discussed in chapter[2].

Chapter 2

Literature Survey

In WMSN multiple camera nodes capture images which can generate redundant images. To reduce this redundancy various techniques such as image processing before transmission, transmission of low resolution images, etc. can be adapted.

2.1 In network Multimedia data processing techniques

In WMSNs, multiple camera sensors are deployed to provide multiple views, multiple resolutions, and enhanced observations of the environment. There exists correlation among the visual information observed by cameras with overlapped field of views. This causes substantial redundancy in the network traffic.Multimedia source coding is a common approach to remove the redundancy of visual information.

However, the resource constraints of the sensor nodes bring new challenges when applying source coding globally in the entire network. Also it requires extensive computation at the encoder, which places heavy burden on the resource-constrained sensor nodes.

In contrast, distributed source coding only requires low-complexity encoding and leaves the intensive computations at the decoder. However, this coding strategy requires each sensor node to have the knowledge of global correlation structure, which would incur significant additional costs. For these reasons, multimedia source coding is infeasible to be applied globally in a large scale network, despite its outstanding compression gains.

Here we discuss few Multimedia data reduction techniques that reduce the overburden on network from heavy flow of data resulting due to multimedia data transmission.

2.2 Correlation Based Video Processing technique

We consider a 2-D model where the sensing area of a video sensor s is a sector denoted by 4-tuple (P, R, V, α). Here P is the location of the sensor node, R is the sensing, radius, V is the center line of sight of the camera's field of view which will be termed sensing direction and α is the offset angle of the field of view on both sides of V.[4]

A point P1 is said to be covered by sensor s if and only if the following conditions are met:

1. $d(P, P_1) < R$, where d(P, P_1) is the Euclidean distance between the location P of sensor s and point P_1
2. The angel between $\overrightarrow{PP_1}$ is within $[-\alpha, \alpha]$

Given the sensing areas of video sensors s1 and s2 as F(s1) and F(s2) respectively, the correlation degree between S1 and S2, denoted by c(s1, s2) is defined as

$$c(s_1, s_2) = \frac{F(s_1) \bigcup F(s_2)}{\alpha R^2}$$

To simplify the analysis, instead of treating the overlapping region as a continuous area, we adopt the view of the sensing region as a set of discrete points.

Assume sensor s has sensing direction $\overrightarrow{V} = (V_x, V_y)$ and located at P = (x_0, y_0) . Let P_1 and P_2 denote their respective locations and $\overrightarrow{V_1} = (V_{1,x}, V_{1,y})$ and $\overrightarrow{V_2} = (V_{2,x}, V_{2,y})$ denote their respective sensing directions. Assume that segment A_1A_2 is the scan line in the image plane and A_m is the middle point of A_1A_2.

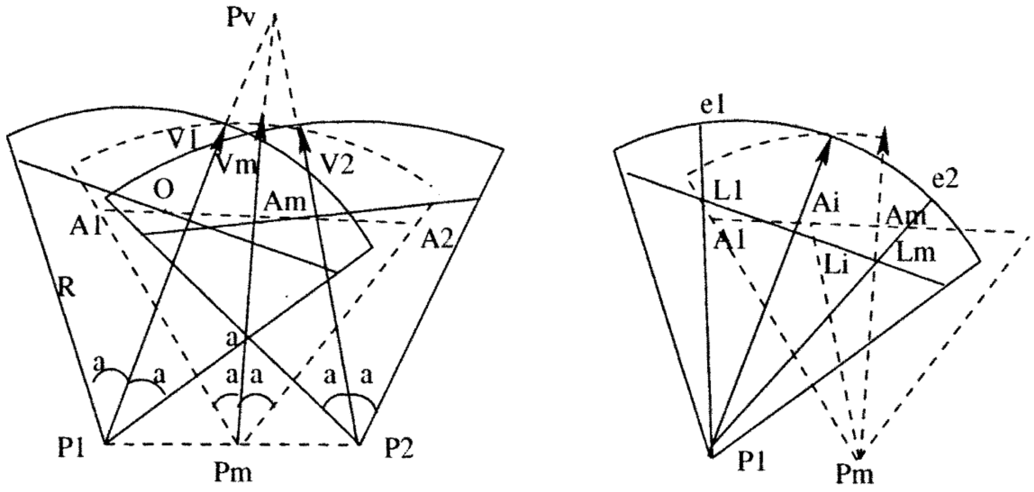


Figure 2.1: Video processing

If the images delivered by s_1 and s_2 can be utilized to restore A_1A_m and A_mA_2 respectively, they can then be fused together to construct the virtual image. For this purpose, sensor s needs to cover the sector defined by point P_1, e_1, and e_2 that corresponds to L1Lm in its image plane. The distance between the location and the image plane for all sensors are the same and denoted by f. The scan line that spans A_1A_m can be represented as

$$x = x_m + fV_{m,x} - tV_{m,y}$$
$$y = y_m + fV_{m,y} + tV_{m,x}$$

Notice that distance between left end point A_1 and middle point $A_m(x_m, y_m)$ is f tanα. Therefore the coordinate of A_1 is

$$x_{A1} = x_m + fV_{m,x} - ftan\alpha V_{m,x}$$
$$y_{A1} = y_m + fV_{m,x} + ftan\alpha V_{m,x}$$

For sensor s_1 located at $P_1 = (x_1, y_1))$ and with sensing direction $\overrightarrow{V_1} = (V_{1,x}, V_{1,y})$. The middle point of its scan line l_1 is $(x_1 + fV_{1,x}, y_1 + fV_{1,y})$ and hence L_1 can be represented as

$$x = x_1 + fV_{1,x} - tV_{1,y}$$
$$y = y_1 + fV_{1,y} + tV_{1,x}$$

In the virtual camera located at P_m and with sensing direction, assume that number of pixels shall cover segment $[A_1, A_m]$. According to Equations 1 and 3, points A_1 and A_m in the image plane will be mapped to $\left(M - \frac{t_1}{p}\right)$ and $\left(M - \frac{t_m}{p}\right)$ respectively in $s_1's$ image plane, where p = $\frac{ftan\alpha}{M}$ and t_1 and t_m can be obtained by replacing (x_i, y_i) in above equation with the coordinates of A_1 and A_m

$$t = \frac{(x_1 + fV_{1,x} - x_i)y' - y_1 + fV_{1,y} - y_i x'}{V_{1,x}x' + V_{1,y}y'}$$

where $y' = (y_m - y_i)$ and $x' = (x_m - x_i)$.

Once the images are delivered to the sink, they will be fused together to construct the composite image from the two sensors. The fusion process is depicted in fig(2.2). Images from sensor s1 and s2 will be utilized to generate half of the fused image G shown in Fig.2.2

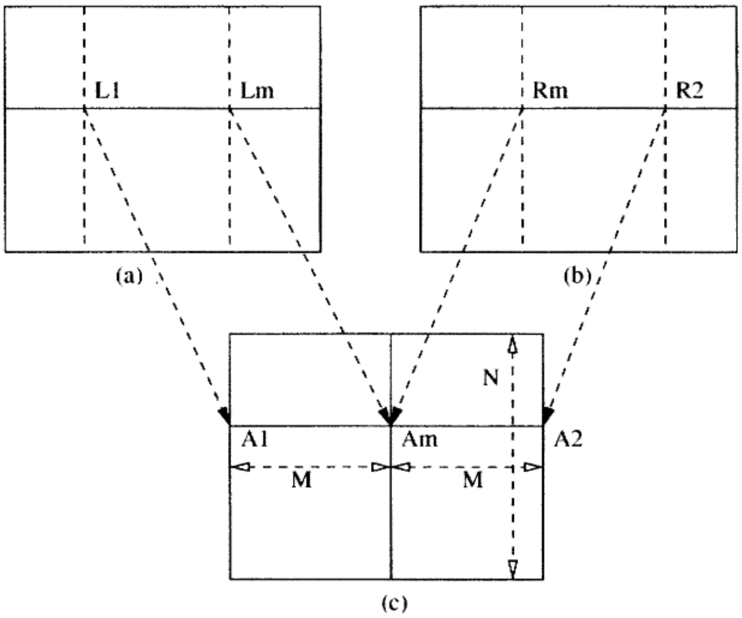


Figure 2.2: Video fusion

For each pixel of the resulting image G, one corresponding pixel needs to be found in images from either s_1 or s_2, depending on which half it residing in. Pixels [1... A] of one horizonal scan line of G correspond to pixels $\left[M - \frac{t_1}{p}, M - \frac{t_m}{p}\right]$ on the scan line of the image captured by s_1; pixels [M 2M] of one horizonal scan line of G correspond to pixels $\left[M - \frac{t_m}{p}, M - \frac{t_2}{p}\right]$ on the scan line of the image captured by s_2.

2.2.1 Drawbacks of Correlation based technique :

- This technique is only valid when the sensing directions of the two sensors do not differ very much.
- This processing method is limited between two sensors. How to deal with the cooperative processing of more than two sensors is a problem that has not been well investigated.

2.3 Image Processing Based technique

Image processing techniques are used to obtain correlation and design collaborative processing. Images from correlated views are roughly registered using correspondence analysis. In spatial correlation is obtained by an image shape matching algorithm, while temporal correlation is calculated via background subtraction.

2.3.1 Correspondence analysis and super-resolution based technique :

Image a scenario in which a number (N) of sensors share different looks at a common scene, with one particular camera having the perspective of most interest (call it the primary sensor) each of the N - 1 non-primary cameras have some overlap between their fields of view and that of the primary camera. The intersection of these N - 1 overlapping regions describes a common region of overlap, where each sensor sees the same image block in part of its field of view.

We can compute pair-wise registrations between the image from the primary camera and those from each of the N - 1 non primary cameras. These registrations allow us to transform the non-primary images into the frame of reference of the primary camera.[6]

2.3.2 Registration Via Correspondence Analysis :

To realize a low-bandwidth means of communicating image information among sensors we employ the shape context algorithm. This technique allows for the correspondence of similar images in terms only of their critical feature points via a shape descriptor known as the shape context. First both images are sampled to extract two sets of points which capture the critical features of the images. Then. in both sets, for each point a shape context is computed.

The shape context is a histogram which describes the spatial relationship of all other points in the set to the sample point for which the context is computed. This yields two sets of sample points with associated shape contexts. Bi-partite graph matching is then employed to find the best one-to-one matching.

Two views of a scene (say, one of the N - 1 non-preferred images I, in Figure(a) and the preferred image I,, in Figure(b) are captured by a neighboring pair of sensors. Following the shape context registration process, there exists a warping transform which can take I, to the frame of reference,of Ip, with the result being the image in Figure(c). To evaluate the effectiveness of this transformation, the difference image between the warped I , and the original I, is shown in Figure(d).

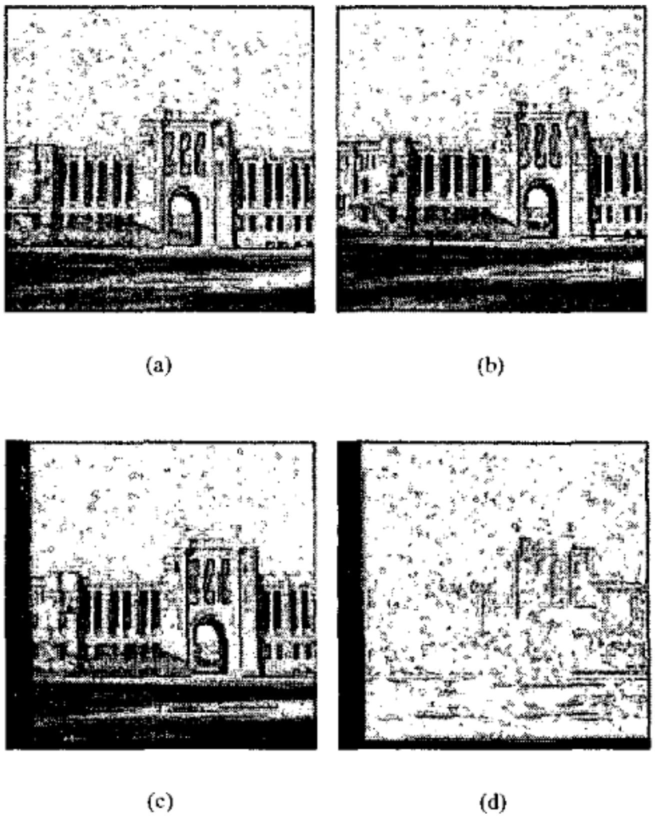


Figure 2.3: Sample super-resolved images (a) original high-resolution image (b) sample low-resolution image (c) super-resolved image (d) zerotree-coded image using equivalent bitrate

2.3.3 Decoding Via Super Resolution :

The images in this transformed set can then be intersected with each other to identify the largest common pixel block in the set (i.e. the largest block in the area of overlap).

Once this common block has been identified in each transformed image, we have exactly the data required for super resolution. We distribute the coding of this block among the group of sensors by requiring each to transmit a down sampled version of the post-registration block to the receiver. Thus, the coding cost for this block in the sensor of interest is spread among the entire neighborhood of sensors, reducing the transmission load on the preferred sensor and thus increasing its operational lifetime.

2.4 Shape Matching Based Aggregation :

We treat an object as set of points and assume that the shape of an object is captured by a finite subset of its points. These can be obtained as locations of edge pixels as found by an edge detector, giving us a set $\rho = \{p_1,...,p_n\}$.[7]
Figure 2.4 shows sample points for two shapes.

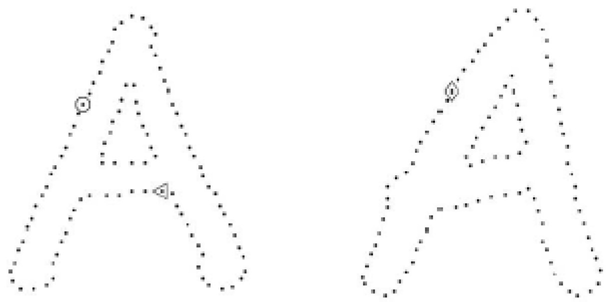


Figure 2.4: Edge Pixels

For each point p_i on first shape, we want the best matching point q_j on second shape. Matching is easy if rich descriptor is used. That is why we use shape context. To find shape context of a point p_i on the shape, we compute a coarse histogram h_i of the relative coordinates of the remaining n-1 points

$$h_i(\text{k}) = \{\text{q} \neq p_i : (\text{q}-p_i) \in \text{bin(k)}\}$$

This histogram is defined to be the shape context of p_i .

Consider a point p_i on the first shape and a point q_j on the second shape. Let C_{ij} = C(p_i,q_j) denote the cost of matching these two points.

$$C_{ij} \equiv \mathrm{C}(p_i,q_j) = \frac{1}{2}\sum_{k=1}^{k} \frac{[h_i(k)-h_j(k)]^2}{h_i(k)+h_j(k)}$$

where $h_i(k)$ and $h_j(k)$ denote the k bin normalized histogram at p_i and q_j respectively. The cost of matching the two points should be minimized to reduce the overhead of matching on nodes reduces. This takes time of O(N^3) which increases when the number of nodes in the network go on increasing.

2.4.1 Drawbacks of Image Matching Based Technique:

The task of shape matching is very compute intensive, thus using such highly complex logics for comparison of images consumes additional energy of nodes.

2.5 Drawbacks of Image Processing Based Technique:

- The performance of image processing algorithms are application dependent: different types of images will require different processing schemes
- Image processing techniques are complicated and computation extensive, which will bring about extra computation costs for sensor nodes.

2.6 Entropy based Framework

Since uncompressed raw video streams require excessive bandwidth that is impossible to be supported by wireless multihop networks, multimedia source coding must be employed to achieve high compression efficiency. But video coding achieves high compression performance at the expense of extensive computation at the encoder.

Collaborative multimedia in-network processing is suggested as an effective way to avoid the transmission of redundant information. According to the requirements of specific applications, each sensor node can filter out uninteresting events locally, or coordinate with each other to aggregate correlated data.

2.6.1 Advantages of Entropy based Framework :

- This Framework is independent of the specific coding algorithms and images types, thus providing a generic architecture that allows users to freely customize the WMSN applications based on them.

Chapter 3

Entropy Based Framework

The spatial correlation of camera sensors causes substantial redundancy in the observed visual information in WMSNs, which can be removed to improve energy efficiency and network performance. We propose to remove the redundancy of visual information through joint compression/ coding among multiple correlated cameras. To maximize the compression gain of the whole network, we partition the network into a set of clusters where camera sensors with high joint compression gains are grouped together. Since entropy serves as the lower bound of coding rates, we can estimate the joint entropy of multiple cameras as a prediction of the joint coding efficiency.

3.1 Spatial Correlation

For Camera i and Camera j in the group S = S1, S2, ..., SN, we will derive a correlation coefficient ρ_{ij} to describe the degree of correlation between image Xi and image Xj . For the group of camera sensors, the correlation among the images observed by these cameras (X1,X2, ...,XN) will be represented as a correlation matrix C, denoted as $C = (\rho_{ij})_{N*N}$, where ρ_{ij} is the correlation coefficient of image Xi and image Xj .

Correlation-based Camera Selection : If only M cameras($M < N$) are allowed to transmit their observed images to the sink, how to select M cameras out of the N cameras so that the sink can gain the maximum amount of information.[5]

3.1.1 Spatial Correlation Model for Visual Information :

Sensing Model :

A camera also has limited sensing range. It can only observe the objects within its field of view (FoV). A cameras field of view is determined by four parameters (P,R,$\vec{V}$,α), where P is the location of the camera, R is the sensing radius, $\vec{V}$ is the sensing direction, and α is the offset angle.

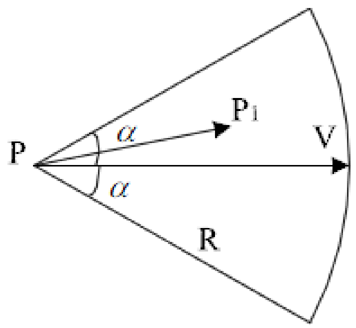
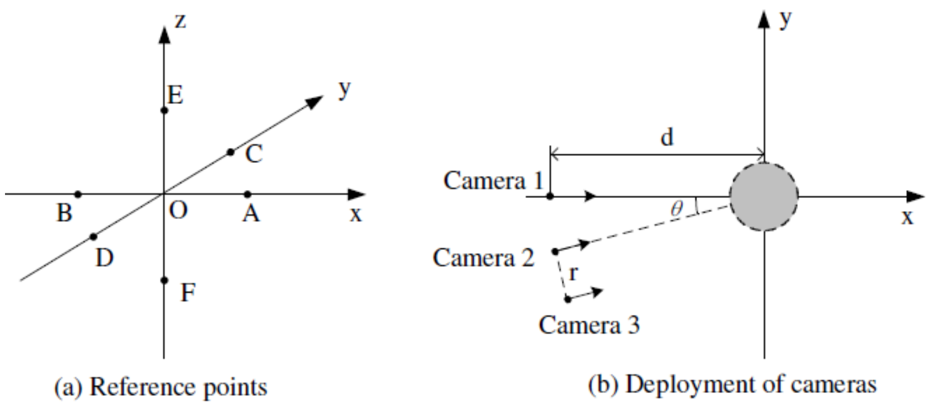


Figure 3.1: Field of view

Seven reference points, which can also be regarded as feature points or key points in a scene, are chosen as: O(0, 0, 0) ,A(1, 0, 0), B(-1, 0, 0), C(0, 1, 0), D(0,-1, 0), E(0, 0, 1), F(0, 0,-1). These reference points form six unit reference vectors along the orthogonal directions in the 3-D world: $\vec{OA}$, $\vec{OB}$, $\vec{OC}$, $\vec{OD}$, $\vec{OE}$, $\vec{OF}$.



(a) Reference points (b) Deployment of cameras

If the projection of OA is $o_i a_i$ =(u_i, v_i) on Camera i, and $o_j a_j$ = (u_j, v_j) on Camera j, the distance is calculated as $d_{OA} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$

The disparity between the images at Camera i and Camera j, denoted by δ, is defined as the average distance of the four vectors:

$$\delta = \frac{1}{4}(d_{OA} + d_{OB} + d_{OC} + d_{OD})$$

Generally, for Camera i and Camera j with position parameters (d_i,r_i,θ_i) and (d_j,r_j,θ_j), the disparity between the images at the two cameras is given by

$$\delta_{ij} = \frac{1}{4}(\left|\frac{-d_i sin\theta_i - r_i cos\theta_i}{d_i + cos\theta_i} - \frac{-d_j sin\theta_j - r_j cos\theta_j}{d_j + cos\theta_j}\right| + \left|\frac{d_i sin\theta_i + r_i cos\theta_i}{d_i - cos\theta_i} - \frac{d_j sin\theta_j + r_j cos\theta_j}{d_j - cos\theta_j}\right| + \left|\frac{d_i cos\theta_i - r_i sin\theta_i}{d_i + sin\theta_i} - \frac{d_j cos\theta_j - r_j sin\theta_j}{d_j + sin\theta_j}\right| + \left|\frac{-d_i cos\theta_i + r_i sin\theta_i}{d_i - sin\theta_i} - \frac{-d_j cos\theta_j + r_j sin\theta_j}{d_j - sin\theta_j}\right|)$$

The disparity value increases as the sensing direction difference increases. The larger the disparity value, the more differences exist between the two images, i.e. the images are less correlated. we bound the disparity value from 0 to 1 as follows:

$$\delta = min(\delta, 1)$$

Thus the correlation coefficient that is complementary to the disparity function:

$$\rho = 1 - \delta$$

When the correlation coefficient is 0, it means that the two images are independent of each other. If it equals to 1, the two images are highly correlated. The larger the correlation coefficient, the more correlated are the two images.

3.1.2 Joint Effect of Multiple Correlated Cameras :

We device a method to measure the amount of visual information from multiple correlated cameras.

Entropy-based Approach : The concept of entropy is used to measure the amount of information of a random source. If an image is interpreted as a sample of a "gray-level source", the source's symbol probabilities can be modeled by the gray-level histogram of the observed image. An estimate of the source's entropy can be generated as

$$\widetilde{H} = -\sum_{k=1}^{L} p(r_k logp(r_k))$$

where L is the number of all possible gray-levels, and p(r_k) is the probability of the kth gray-level. It denotes the average amount of information per pixel in the image.

If a camera Si transmits its observed image Xi to the sink, the amount of information gained at the sink is H(Xi). If the group of camera sensors, S = S1, S2, ..., SN, transmit

their observed images X1,X2, ...,XN to the sink, the amount of information gained at the sink will be the joint entropy H(X1,X2, ...,XN).
Thus the joint entropy of A and B is

$$H(A, B) \approx (1 - \frac{1}{2}\rho)$$

Therefore, the amount of information that can be gained from image A and image B together depends on the correlation degree between A and B. The more A and B are correlated, the less joint entropy can be gained from A and B together. That is to say, if two camera sensors transmit their images to the sink, the amount of information gained at the sink will be larger if the two sensors are less correlated.

3.2 A Collaborative Image Compression Framework Using Clustered Source Coding

The clustering strategy has been proved to be an effective way to improve network scalability and energy efficiency for sensor networks. This strategy uses the hierarchical concept where the entire network is divided into regions.

This framework consists of two components: (i) compression efficiency prediction, and (ii) coding hierarchy construction. The compression efficiency prediction aims to estimate the compression gain from joint encoding of multiple cameras before the actual images are captured.

To achieve this, an entropy-based divergence measure (EDM) is proposed, which only takes the camera settings as inputs without requiring the statistics of real images.

In the EDM, the overlapping pattern of the FoVs of multiple cameras is first identified. Then, the correlation degree among the observations from cameras with overlapped FoVs is obtained through a spatial correlation model. Based on the correlation characteristics, a dependency graph based algorithm is designed to estimate the joint entropy of multiple cameras. This joint entropy effectively predicts the compression performance for joint encoding of multiple cameras.

3.3 Multi-camera Entropy Estimation Problem

Joint entropy serves as a lower bound of the overall coding rate of multiple sources for both centralized and distributed source coding.

$$\sum_{i=1}^{N} R_i \geq H(X1, X2, ..., X_N)$$

Our objective is to estimate the joint entropy of multiple cameras in WMSNs through low computation and communication costs. Given a cluster of cameras with observations X1,X2,...,XN, the joint entropy H(X1,X2,...,XN) will be described as a function of the individual entropy (H(Xi)) and field of view (Ai) of each camera, and the correlation coefficients between any two cameras (ρ_{jk}) in the cluster.

3.3.1 Entropy based divergence Measure (EDM) :

The EDM scheme consists of the following two components.

1. Area division for FoVs. Given a group of cameras, their FoVs are divided into several regions, such that each region is covered by the same set of cameras.

2. Joint entropy estimation for regions. For each region, a dependency graph is constructed based on the correlation among the cameras. The joint entropy of the region is then estimated by traversing the dependency graph. Finally, the total joint entropy for the group of cameras is the sum of the entropies of all the regions.

3.3.2 Area Division for Overlapped Field of Views :

An arbitrary point O1 is in the FoV of the camera if it satisfies

$$\begin{cases} \left|O\overrightarrow{O_1}\right| & \leq R \\ \theta \leq \alpha \end{cases}$$

where θ is the angle between $\overrightarrow{OO_1}$ and $\overrightarrow{V}$.

We consider the case when N cameras (C1,C2,...,CN) are deployed on the ground plane and all the cameras are homogeneous, i.e., they have the sensing radiuses (R), and offset angles (α). Denote the FoV of an individual camera Ci by Ai(Oi,Ri,$\overrightarrow{Vi}$,α_i), and the overall FoV for these cameras by A (A = A1,...,AN).

3.3.3 Dependency Graph Based Joint Entropy Estimation :

We study a two camerasf case first. Suppose there are only two cameras (C1 and C2) in a region Pi. We can depict their relationship using a dependency graph: C2 → C1.

The joint entropy of the observations from C1 and C2 can be calculated by traversing the dependency graph. The source node C2 contributes the entropy of its observations, H(X2(Pi)), and the node C1 contributes the conditional entropy with respect to C2, H(X1(Pi)| X2(Pi)), so the joint entropy is calculated by adding these two terms: $H(X_1(P_i), X_2(P_i)) = H(X_2(P_i)) + H(X_1(P_i) \mid X_2(P_i))$. The dependency graph can also be constructed as C1 →C2, from which we can get the same result of joint entropy.

The two cameras case can be extended to estimate the joint entropy of more than two cameras. Denote the dependency graph by G(V,E), where V is a collection of cameras, and E is a collection of directed edges that stand for dependencies. Joint entropy of the region is calculated by traversing all the nodes in the graph along the directed edges. if camera Cj is most correlated with camera Ck, we say that Cj is dependent on Ck, and we can construct a directed edge starting from Ck and ending at Cj : Ck →Cj . Cj is said to be a direct successor of Ck, and Ck is a direct predecessor of Cj.

The dependency graph is designed to be a directed acyclic graph with the following constraints: a camera is either a source node (i.e., a node that has no predecessors), or a direct successor of one of the other cameras; a dependency graph may have several source nodes, but each node can have at most one direct predecessor; and there should be no loops in the graph, e.g., Ck →Cj and Cj →Ck cannot exit in the same graph.

These properties could be guaranteed through the procedure of constructing the dependency graph. For each node Ck, if another node Cj is most correlated with it, i.e., neighbor(Cj) = k, the algorithm adds Ck →Cj into the graph only when two conditions are met: i) Cj has no predecessors, and ii) Cj is not a predecessor of Ck . The first condition guarantees that each node can have at most one direct predecessor, and the second one guarantees that there are no loops in the graph.

Given a dependency graph with the above features, the joint entropy is estimated by traversing all the nodes in the graph and adding the entropies of the nodes together. A source node contributes its individual entropy to the joint entropy, while a non-source node contributes its conditional entropy with respect to its direct predecessor to the joint entropy.

Since the FoVs for a group of cameras are divided into several independent regions, the total joint entropy is the sum of the entropies of all the regions. For a group of cameras with observations (X1,...,XN), with their FoVs divided into regions (P1, . . . , PM), the total joint entropy is given by H(X1,...,XN) = H(P1) +...+ H(PM).

3.3.4 Optimal Coding Clustering Problem :

An optimal coding clustering (OCC) problem, which we define as: find a set of coding clusters with the minimum total entropy, such that each camera node is covered by at least two different clusters. The minimization of total entropy guarantees that the global compression gain is maximized, while the coverage requirement ensures that the impact of cluster failures on the decoding reliability is mitigated.

Chapter 4

Proposed Algorithm

The total number of transmissions to the sink in the existing approach is approximately double the number of clusters present in the network. This is because according to OCC problem, each camera node is covered by atleast two clusters.

To reduce the number of total transmissions, we device a new approach in which each cluster will transmit only once a session.

Initially the camera nodes in the network are distributed among clusters as per the Algorithm 1 of clustering for calculating the joint entropy of correlated camera nodes. As per this Algorithm 1, each cluster consists the nodes having higher correlation with each other. Thus the joint entropy of each cluster is minimum, as the camera nodes of the cluster are closely correlated. We use this computed total entropy as the measure parameter for transmission decisions of images by clusters.

Cluster Entropy : It is the joint entropy of all the camera nodes that for the cluster.

Consider S_i as a set of clusters having highly correlated camera nodes. Initially all the clusters send their cluster entropy to each other. Now each cluster computes its joint entropy with each of the received entropies of N-1 clusters individually. Now search for pairs of nodes having maximum joint entropy with each other.

Algorithm for Clustering

S_i=$S_1, S_2, ...S_n$ // Set of all sensor nodes

E_{jk}= $E_{12}, E_{13}, ..., E_{23}, ..., E_{n-1n}$ // Joint Entropies of all pairs of nodes

P: Number of clusters // should be even

$E_{max} = E_{12}$

$E_{min} = E_{12}$ for j=1 to n

for k=2 to n-1

for l=2 to p

if($E_{ij} < \delta$ && $E_{max} < E_{jk}$)

set $E_{max,p} = E_{jk}$

$C_i \leftarrow S_{jk}$

$M_i = S_i$ - C_i

for i=1 to n

for j=i+1 to n

if($E_{M_iC_i} < E_{M_iC_j}$)

$C_j \leftarrow M_i$

If the number of clusters are n, then the number of such pairs of nodes with maximum joint entropy will be n/2. The obtained pairs will then become cluster heads of the clusters to be formed for transmission.

The cluster members will be those nodes which have minimum joint entropy with the node. Thus each cluster consist of nodes having high correlation with cluster head and the correlation between any two clusterheads is very low.

This approach will increase the information gain at the sink as the images delivered at the sink will have minimum correlation with each other.

Thus the more diverse images will only be transmitted to sink and redundancy of images to be transmitted is reduced upto considerable extent.

Transmission Pattern

In the first round of transmission the cluster heads transmit of behalf of the cluster members forming the cluster. But if in the second round the cluster heads remain same as that in first round, then life of clusterheads reduces abruptly. To avoid this situation, in the second round another group of clusterheads are formed which transmit on behalf of entire cluster. This pattern of transmission distributes the responsibility of clusterhead among all the clusterheads equally, which increases the network life time considerably.

Chapter 5

Simulation Results

5.1 Paramater : Image Quality

The quality of image decoded at sink varies according to the technique used for aggregating the images at source before transmission.

5.1.1 Correlation Based Partial Image Capturing Technique:

The Correlation based Partial Image Capturing divides the image capturing task among multiple cameras, where each camera captures partial segments of the object under consideration, to obtain the complete image at sink, the sink node requires to fuse the partial images obtained from various nodes.

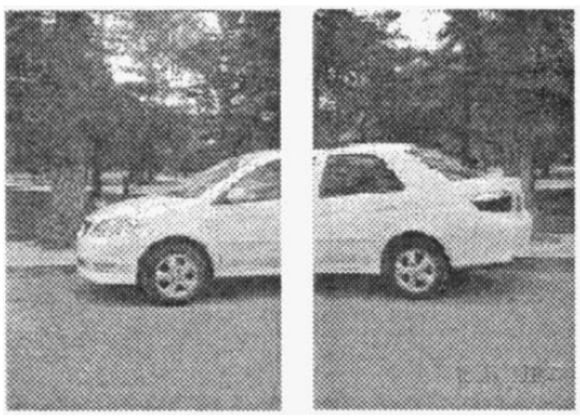


Figure 5.1: Fragmented image at source

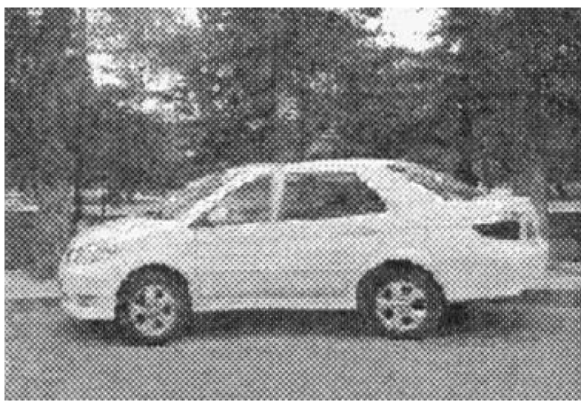


Figure 5.2: Fused image at sink

After fusion at sink there is a fusion mark in the fused image, thus there is degradation in quality of image received at sink. Such a fusion mark in image is acceptable only in cases where quality of image is not an important criteria and no further processing is required to be carried out on image.

5.1.2 Correspondence analysis and super-resolution based technique :

During super-resolution large number of low resolution images are merged together to form a single high resolution image. In order to obtain a high resolution image at sink, large number of low resolution images are required. Thus the quality of image received at sink depends on the number of low resolution images involved in formation of high resolution image at sink. In The original image form and its corresponding low resolution images are generated shown below

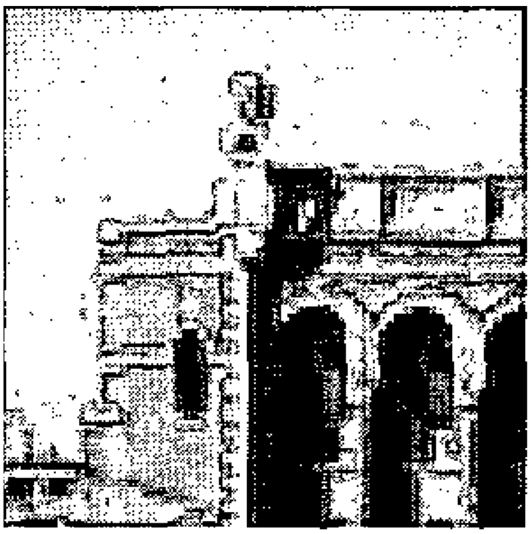


Figure 5.3: Original Image

Figure 5.4: Fused image at sink

While merging of images, large amount of energy is consumed by the processing at sink node. So if we consider large number of images, the life of sink will be reduced. Thus there is a trade-off between quality of image and life time of sink node. The image formed at sink after Super-resolution of several low-resolution images has low quality, which is shown below

Figure 5.5: Super-resolved image

5.1.3 Shape Matching based aggregation :

In shape matching based technique shape contexts are used to plot a image and comparison is carried out on the bases of these shape contexts. To describe an image with good resolution requires large number of shape contexts and while comparing two such high resolution images, the process of searching matching points in two images becomes lengthy and time consuming. Thus the quality of image depends upon the number of shape contexts used to describe the image. The Image at source and destination is shown below

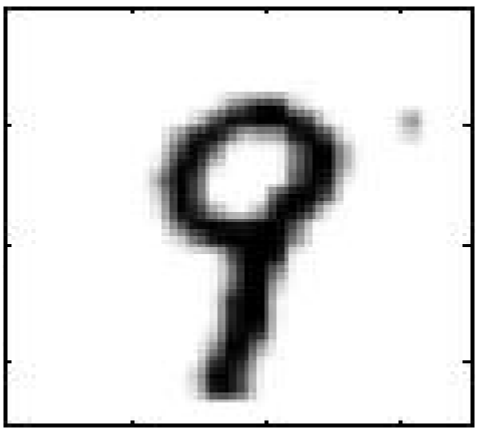


Figure 5.6: Original Image

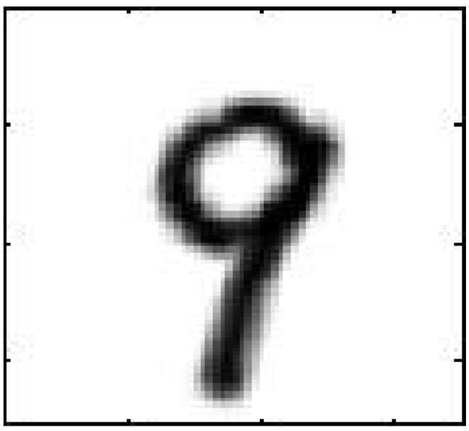


Figure 5.7: Image at sink

5.1.4 Entropy Based Aggregation Framework:

In entropy based framework, the nodes having highly correlated images are clustered together to send a single image on behalf of all the nodes forming a cluster. In quality of image received at sink is same as that of the image captured by source node. As there is no processing carried out on the image before transmission. But the information gain depends upon how correlated the images are. Two images captured by two different cameras are as shown below

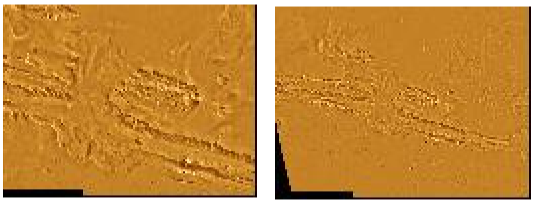


Figure 5.8: Images of two different cameras

The single image shown on behalf of these two images is as shown below

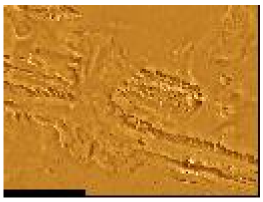


Figure 5.9: Image transmitted to sink

5.2 Network Topology

Below Figure displays the topology of network. The network consists of 15 nodes, where the 15th node is sink and rest of all are traffic generator and also act as relay nodes when they do not have their own data for transmission.

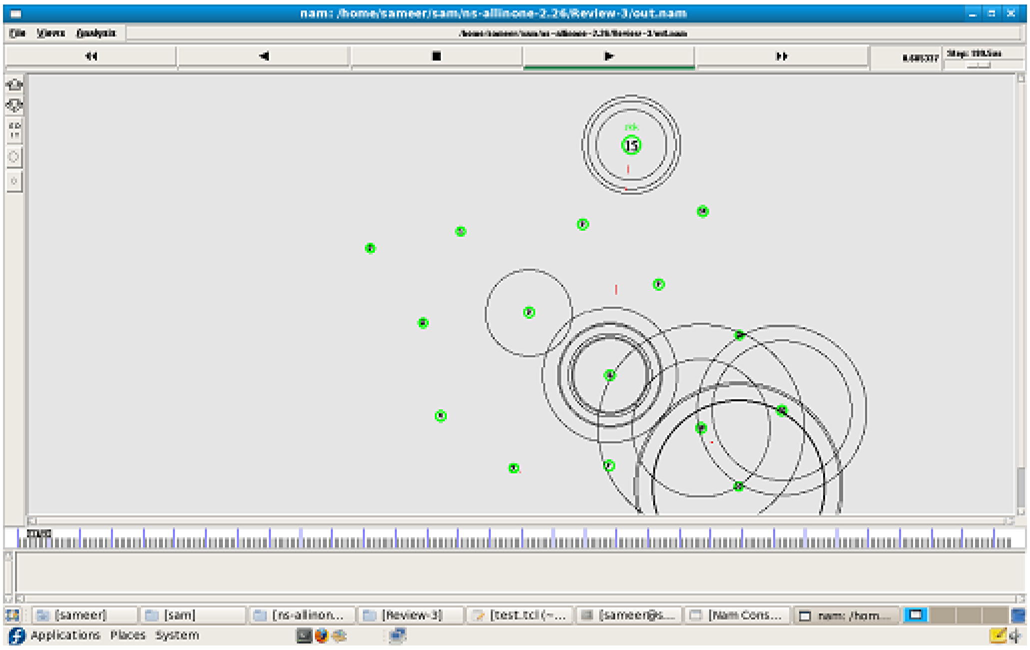


Figure 5.10: Network Topology

5.3 Parameter : Network Lifetime

The proposed EBAF has greater lifetime as compared to other techniques, as the cluster heads keep on changing in each round of transmission. The existing EBAF has large number of clusters as each node is included in atleast two clusters, thus each node will be sending data to two clusterheads and this will result into reduced network lifetime because of increased number of data transmissions. Here in each round the clusterhead keeps on transmitting on behalf of all nodes, thus after few rounds the clusterhead itself will die. As the number of nodes increase the network life time in each of the existing technique will be reduced due to the reasons specified in the explanation of the techniques. But it is clear in the below graph that as the number of nodes increase in the network the network life time is also enhanced.

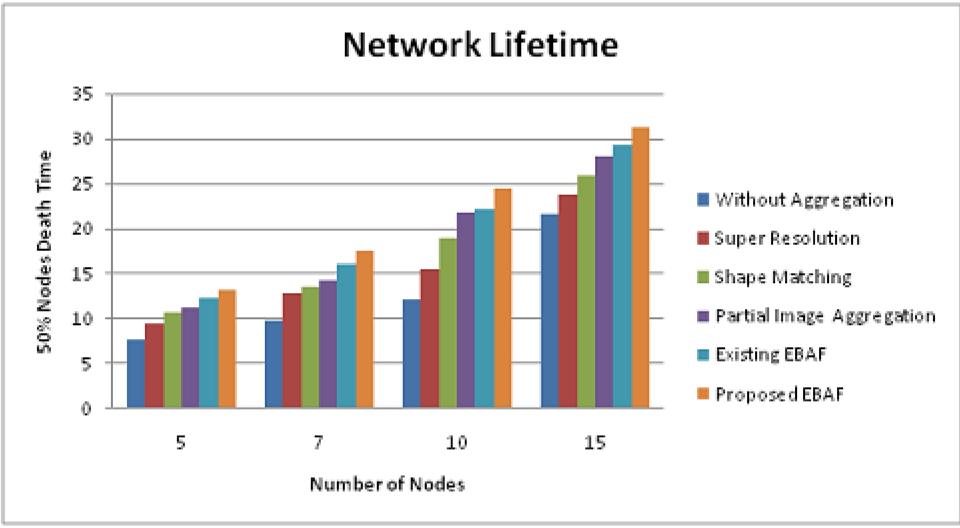


Figure 5.11: Network Lifetime

5.4 Parameter : Network traffic corresponding to each node

The existing EBAF has number of transmissions equal to twice the number of clusters. As each node is a member of atleast two clusters. This increases the number of clusters and as we know that each clusterhead transmits on behalf of entire cluster. Thus if the number of clusters increase then the number of transmissions per round will also increase. In the proposed EBAF the number of transmissions is equal to number of clusters. Thus as compared to the existing techniques the number of transmissions in the network will be low in proposed EBAF.

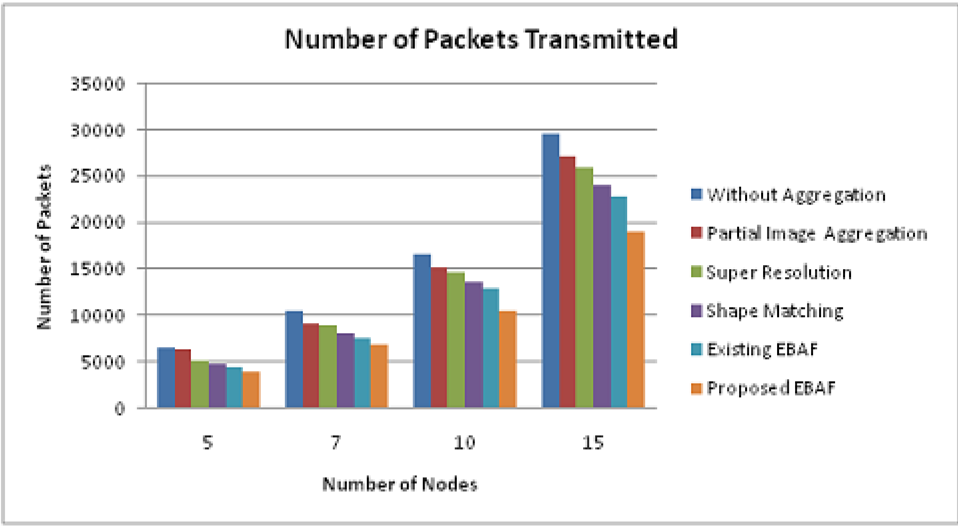


Figure 5.12: Network traffic

Chapter 6

Conclusion and Future Work

6.1 Conclusion

- The Entropy based Framework is independent of the specific coding algorithms and images types, thus providing a generic architecture that allows users to freely customize the WMSN applications based on them.
- In the existing Entropy based approach the number of transmission is twice the number of clusters in network.
- The proposed approach avoids this limitation of large number of transmissions. In best case the total number of transmissions is equal to number of clusters in network and in worst case the number of transmission is equal to number of sensor nodes in the network.
- Thus for a scenario with highly correlated sensor nodes, this approach will reduce the number of transmission to considerably small value.
- The reduction in number of transmissions will lead to increase in life time of network due to conservation of energy because of reduced number of transmissions.

6.2 Future Work

The proposed framework considers image aggregation for WMSN which can be extended to aggregation of other multimedia traffic, such as audio and video captured by different sensor nodes aggregated together to reduce the traffic.

Appendix A

Simulation Platform Details

A.1 Simulation Platform:

Network simulator is tool used to stimulate different network scenarios. We can build ns either from the the various packages (Tcl/Tk, otcl, etc.), or We can download an 'all-in-one' package. I start with the all-in-one package, especially if we're not entirely sure which packages are installed on your system, and where exactly they are installed.

The disadvantage of the all-in-one distribution is the size, since it contains some components that we don't need anymore after we compiled ns and nam. It's still good for first tests, and we can always switch to the single-package distribution later.

Network simulator is tool used to stimulate different network scenarios. Using this we need to achieve following goals.

- Optimization on Energy, Lifetime of network and packet delivery.
- Analysis of overhead delivery for different scenario
- Find out the gain achieved by traffic patterns.

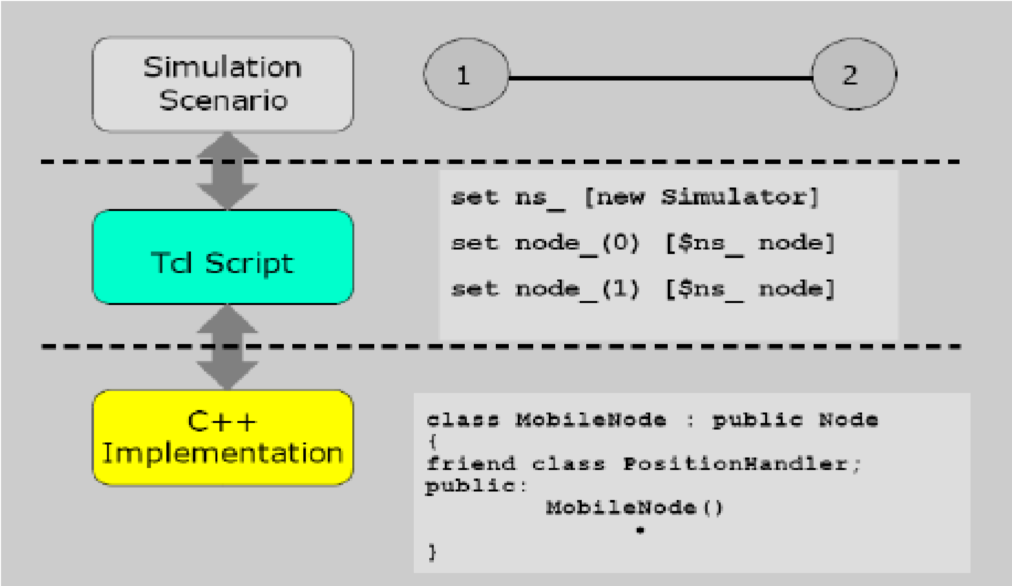


Figure A.1: NS-2 Overview

A.2 Installation Procedure

A.2.1 Installation of FEDORA 8

The various platforms which supports Network Simulator(NS) are Linux,Solaris, Windows XP using Cygwin. We have used Fedora 8 because Fedora 8 is open source. Following are the steps to install Fedora 8 OS.

- Remove data from one of the drives where fedora is to be installed.
- Load the CD in drive.
- Skip for disc checking.
- Then select the English US language.
- Select Custom layout for partitioning hard disk and click next.
- Where we want to install, make that disk space free rst (at least 5GB). The drives are named as dev/sda1 (which is C drive in the system) and so on for the other drives.
- Then select New for that free space.

- Give a mount point as '/'(root) and select ext3, give a space equal to total free space minus 2*RAM. This space is given in MB.
- Then make it ok.
- Then on the remaining free space select New and then select le format as Swap Give space to swap 2*RAM size. This space is to be given in MB.
- Select grub boot loader and make windows or other OS as by default.
- Select all the software if needed.
- Give the computer name.
- Select region as Asia/Calcutta.
- Enter Root password.
- Then installation will start (1103 packet should be installed for proper functioning of fedora).

A.2.2 Installation of Network Simulator

Network Simulator 2.26

Step 1: Prepare necessary files for installation:

1. NS-2.26 package: ns-allinone-2.26.tar.gz
2. Patch for compiling NS-2.26 with GCC 4.1.2: ns-2.26-gcc412.patch

Step 2: Download NS-2.26, apply ns-2.26-gcc412.patch, and install it
Download NS-2.26 from

1. http://www.isi.edu/nsnam/dist/ns-allinone-2.26.tar.gz or
2. tar zxvf ns-allinone-2.26.tar.gz
3. patch -p0 < ns-2.26-gcc412.patch
4. cd ns-allinone-2.26/
5. ./install

Step 3: Set the environment variables to make NS-2.26 works:

1. cd∼

2. gedit /.bashrc
   #LD LIBRARY PATH
   OTCL LIB= /ns-allinone-2.26/otcl-1.8
   NS2 LIB= /ns-allinone-2.26/lib
   X11 LIB=/usr/X11R6/lib
   USR LOCAL LIB=/usr/local/lib
   export
   LD LIBRARY PATH=$LD LIBRARY PATH:$OTCL LIB:$NS2 LIB:
   $X11 LIB:$USR LOCAL LIB
   #TCL LIBRARY
   TCL LIB= /ns-allinone-2.26/tcl8.4.5/library
   USR LIB=/usr/lib
   export TCL LIBRARY=$TCL LIB:$USR LIB
   #PATH
   PATH=$PATH: /ns-allinone-2.26/bin: /ns-allinone-2.26/tcl8.4.5/unix: /ns-allinone-2.26/tk8.4.5/unix

3. source /.bashrc

Step 4: Install Evalvid framework : Evalvid is a complete framework and tool-set for evaluation of the quality of video transmitted over a real or simulated communication network.

1. Download ffmpeg from : git clone git://git.ffmpeg.org/ffmpeg.git

2. ffmpeg -s cif -r 30 -b 64000 -bt 3200 -g 30 -i akiyo_cif.yuv -vcodec mpeg4 a02.m4v

3. MP4Box -hint -mtu 1024 -fps 30 -add a02.m4v a02.mp4

4. ffmpeg -i a02.mp4 a02_ref.yuv

5. $mp4trace - f - s192.168.0.212346a01.mp4\ st_a01$

Bibliography

[1] Hong-Hsu Yen. Energy-aware Image Aggregation in Wireless Visual Sensor Networks. IEEE 2011 978-1-4577-0253-2.

[2] Xiaoxia Ren, Zhigang Yang. Research on The Key Issue in Video Sensor Network. IEEE 2010 978-1-4244-5540-9.

[3] Cheng Chen, Bin Tian, Ye Li and Qingming Yao. Data aggregation technologies of wireless multimedia sensor networks: A survey. IEEE 2010 978-1-4244-7125-6.

[4] Huadong Mat and Yonghe Liu. Correlation Based Video Processing in Video Sensor Networks. International Conference on Wireless Networks, Communications and Mobile Computing, IEEE 2005.

[5] Pu Wang, Rui Dai, and Ian F. Akyildiz. Collaborative Data Compression Using Clustered Source Coding for Wireless Multimedia Sensor Networks. IEEE INFOCOM 2010 978-1-4244-5837-0.

[6] Raymond Wugnef; Robert Nowak, Richard Baruniuk. Distributed image compression for sensor networks using Correspondence analysis and super-resolution. IEEE 2003 0-7803-7750.

[7] Serge Belongie, Jitendra Malik, Jan Puzicha. Shape Matching and Object Recognition Using Shape Contexts. IEEE Transaction on Pattern Analysis and Machine Intelligence VOL 24, April 2002.

[8] Satyajayant Misra, Martin Reisslein, and Guoliang Xue. A Survey of Multimedia Streaming in Wireless Sensor Networks. IEEE COMMUNICATIONS SURVEYS and TUTORIALS, VOL. 10, NO. 4, FOURTH QUARTER 2008.

[9] Andre L. L. Aquino, Antonio A. F. Loureiro, Antonio O. Fernandes, and Raquel A. F. Mini. An In-Network Reduction Algorithm for Real-time Wireless Sensor Network Applications. ACM 2008 978-1-60558-238-2.

[10] http://mohittahiliani.blogspot.in/2010/02/installing-ns226-with-nrlsensorsim-on.html