Overhearing Source Coding in Wireless Multimedia Sensor Networks

Chang-Yu Song
Graduate Institute of Communication Engineering
National Taiwan University
Taipei, R.O.C
r02942122@ntu.edu.tw

***Abstract*—TODO**

I. Introduction

Nowadays, as the concept of Internet of Things (IoT) gradually becomes more and more important, IP cameras on road can be seen in many countries. These cameras are connected to the Internet and responsible for some surveillance applications such as traffic monitoring or crime prevention. However, the installation of these cameras are often around a small area. That is to say, image data collected from these cameras might be correlated to each other. Therefore, we argue in this paper that we can make use of this correlation and try to reduce the radio resource used to transmit the image data.

More specifically, if we consider two cameras allocated at a crossroad, we can analyze their correlation by letting one camera as a I-frame transmitter while the other as a P-frame transmitter. Under this assumption, higher correlation level means that we can reduce more transmission bits when serving the network. In this paper, we will work on how to serve the wireless multimedia sensor network (WMSN) better via the overhearing source coding. We generate the testing images by a 3-D modeling software and these images are passed through the H.264 reference software to estimate the correlation matrix. Based on the correlation matrix, we further present a scheduling algorithm to serve the network so that we can efficiently reduce the total transmission time needed.

The rest of this paper is organized as follows: Section II describes our multimedia network scenario and Scetion III-B states our problem formulation. Section **??** solves this problem and Section **??** shows the simulation results. Finally, Section V concludes our work.

II. Network Scenario

Figure 1 shows our scenario for the wireless multimedia sensor network. We assume that there is a data aggregator located at the center of a city, which is responsible for gathering image data collected from all the IP cameras in this city. All these images are transmitted through the wireless channel under a TDMA coding scheme. Therefore, a camera i can overhear the image data from other cameras if these cameras are transmitted before i and their transmission range cover i.

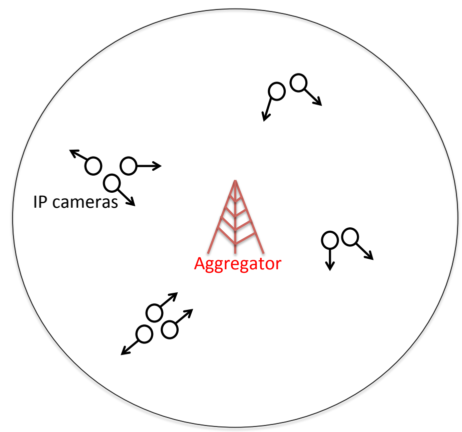

Figure 1. Network scenario

We now claim that different cameras might have different sensing direction, and the sensing direction of a given camera is shown by the arrow in Figure 1. Besides, several cameras might be grouped together in a small area for a real world WMSN application. For example, some companies might install more than one surveillance cameras on the top of its building so that they can safeguard their company [1]. These cameras is set for sensing different direction; however, the covered range of these cameras will overlap with each other. We focus on the overlapping region in this paper since overlapping means correlation and this correlation can be used for motion prediction in the image coding. In this paper, we try to schedule the WMSN for the sake of reducing transmission time for gathering image data.

III. Correlation Analysis and Problem Formulation

In this section, we first details how we analysis the correlation between cameras in the wireless multimedia sensor network and then formulate our problem.

A. Correlation Analysis

As we mentioned before, our purpose in this paper is to reduce the total transmission time of the wireless multimedia sensor network. To simplify our problem, we assume that all the cameras in the network need to transmit its collected image to the aggregator and all the cameras can overhear each other. Therefore, our problem is to determine the transmission order so that all cameras in the network can reference from the most correlated frame.

Given a set $V = \{1, 2, \cdots N\}$ of cameras placed in a city, we now assume that the number of transmission bits of camera i equals $H(X_i)$, and $H(X_i|X_j)$ is the number of transmission bits when camera i reference from camera j to provide H.264 coding. Note that since we assume that all the cameras in the city can hear others' transmission; therefore, whether camera j is an I-frame transmitter or not does not influence the value of $H(X_i|X_j)$. More precisely, if camera j is a P-frame transmitter referenced from camera k, since camera i can gather the data from both j and k, it can first decode the image of camera j and then perform the H.264 encoding. Hence, we can still obtain the same value of $H(X_i|X_j)$.

To proceed, we now analysis the correlation between two cameras i and j as:

$$c_{ij} = \max\{(1 - \frac{H(X_i|X_j)}{H(X_i)}), 0\}. \tag{1}$$

The reason why we take the maximum with 0 is that $H(X_i|X_j)$ can be larger than $H(X_i)$ when the scene gathered by camera i and j differs a lot. Therefore, in order to make the correlation level always non-negative, we set $c_{ij} = 0$ when $H(X_i|X_j) > H(X_i)$. Besides, c_{ii} will not equals to 1 since there still has some remaining header in H.264 coding scheme even when using camera i to predict itself.

B. Problem Formulation

The total transmission time needed for the WMSN is highly related to the number of encoded bits. That is to say, if we first focus on a camera i, we can write the required transmission time for camera i as:

$$T_i = \begin{cases} H(X_i)/C_{i0}, & \text{if camera } i \text{ transmits as an I-frame,} \\ H(X_i|X_j)/C_{i0}, & \text{if camera } i \text{ transmits as an P-frame} \\ & \text{and reference from camera } j, \end{cases}$$

where C_{i0} is the channel capacity from camera i to the data aggregator. Therefore, the total transmission time is to sum up all the individual T_i, $\forall i \in V$ written as:

$$T = \sum_{i=1}^{N} T_i. \tag{2}$$

The goal to minimize equation (2) can be divided into two sub-problems, including the I-frame selection problem and the P-frame scheduling problem. The I-frame selection problem is to choose part of the cameras in the set V, and these selected cameras, say subset S, can save the most transmission bits for the other cameras. For all the cameras in $V \setminus S$, the P-frame scheduling problem is to determine which camera to reference with.

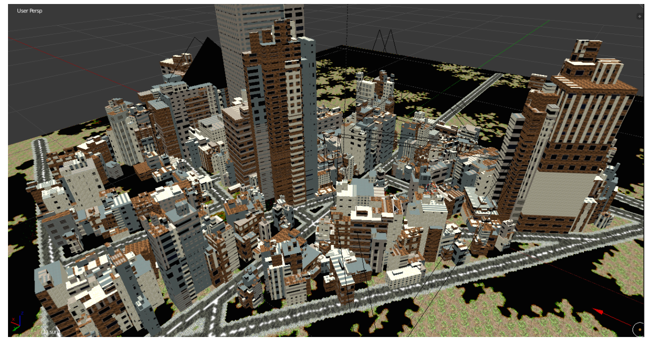

Figure 2. 3D city view in Blender

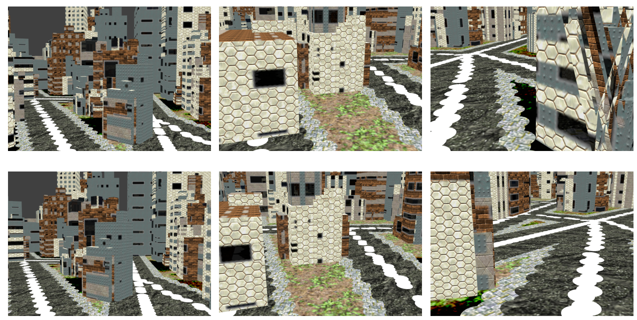

Figure 3. Gathered image from Blender

IV. Experiment Results

We used an open source 3-D modeling software called Blender to generate a city view for our simulation in this paper [2]. Figure 2 Figure 3

V. Conclusion

References

[1] Surveillance camera installation for your business. [Online]. Available: http://www.downtoearthcommunications.com/com-cctv.php

[2] Blender Online Community, *Blender - a 3D modelling and rendering package*, Blender Foundation, Blender Institute, Amsterdam. [Online]. Available: http://www.blender.org