# Joint Effect of Multiple Correlated Cameras in Wireless Multimedia Sensor Networks

**Rui Dai** and **Ian F. Akyildiz**

Broadband Wireless Networking Laboratory
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332
Email: {aprildai; ian}@ece.gatech.edu

***Abstract*—Wireless multimedia sensor networks (WMSNs) are interconnected devices that allow retrieving video and audio streams, still images, and scalar data from the environment. In a densely deployed WMSN, there exists correlation among the visual information observed by cameras with overlapped field of views. In this paper, the correlation characteristics of visual information are used to address two issues: 1) how to measure the amount of visual information provided by multiple cameras in the network, and 2) how to select a group of cameras to report their information to the sink under distortion constraints. An entropy-based analytical framework is developed to measure the amount of visual information provided by multiple correlated cameras first. Based on this framework, a correlation-based camera selection scheme is introduced. Simulation results show that, given a distortion bound at the sink, the correlation-based selection scheme requires fewer cameras to report to the sink than the random selection scheme.**

## I. Introduction

Wireless multimedia sensor networks (WMSNs) are interconnected devices that allow retrieving video and audio streams, still images, and scalar data from the environment [1]. WMSNs are widely used in applications such as video surveillance, environmental monitoring, home automation, and industrial process control. Compared with traditional wireless sensor networks that deal with scalar data, the design of WMSNs are even more challenging. The resource constraints of sensors such as energy constraints and limited processing capabilities still exist. Moreover, visual information, the dominating part of multimedia data, requires more sophisticated processing techniques and much higher bandwidth to deliver. Our study will be focused on the processing and communication of visual information in multimedia sensor networks.

In WMSNs, multiple cameras work together to provide multiple views, multiple resolutions and enhanced observations of the environment [3]. Cameras are directional sensors with limited field of views [6], and the image observed by a camera is directly related to its field of view. There exists correlation among the visual information observed by cameras with overlapped field of views. In many recent research efforts [6][8][9], correlation of visual information among cameras with overlapped field of views are exploited to design the collaborative processing, filtering and aggregation of visual information (multimedia in-network processing).

The correlation for visual information can be obtained by image processing methods. In [8], images from correlated views are registered using correspondence analysis. In [9], spatial correlation is obtained by image shape matching, while temporal correlation is calculated via background subtraction. Another approach to obtain correlation is through geometric methods: by studying camera's location and field of view, a correlation coefficient is derived as the portion of overlapped sensing area to the entire area of the field of view [6].

In this paper, we study the joint effect of multiple correlated cameras in WMSNs. In particular, we study how to measure the amount of visual information provided by multiple cameras in a WMSN. Intuitively, the visual information provided by multiple cameras should be related to the correlation characteristics of the observed images. If the images observed by these cameras are less correlated, they will provide more information to the sink. We develop an entropy-based framework to estimate the amount of information from multiple cameras.

In addition, we propose to study a camera selection problem. Since the delivery of visual information needs very high bandwidth, which may reduce the lifetime of the network, the communication load in WMSNs should be reduced as much as possible. If the application has a distortion bound for the observations, it may not be necessary for all the cameras to report their observed information to the sink. We investigate how many cameras are required to report to the sink under distortion constraints. Based on the study on the joint effect of multiple cameras, we design a correlation-based algorithm to select cameras under distortion constraints.

The remainder of the paper is organized as follows. Section II introduces the entropy-based framework for multiple correlated cameras. Based on this framework, in Section III, a correlation-based camera selection algorithm is proposed. Simulation results for the proposed framework and the camera selection algorithm are discussed in Section IV. The concluding remarks are given in Section V.

## II. Visual Information from Multiple Correlated Cameras

In a multimedia sensor network, multiple camera sensors are deployed to provide multiple views, multiple resolutions and enhanced observations of the environment [3]. A typical

978-1-4244-3435-0/09/$25.00 

scenario of WMSN is: the application specifies which area it is interested in, and the cameras that can observe this area will transmit their images to the sink. For a certain area of interest, suppose there are $N$ cameras that can observe the area of interest, we denote them as a group $S = \{S_1, S_2, ..., S_N\}$, and denote their observed images as $X_1, X_2, ..., X_N$. There exists correlation among the observations $X_1, X_2, ..., X_N$.

As introduced in Section I, the correlation characteristics for images observed by different cameras can be obtained through image processing methods [8][9] or geometric methods [6]. Take the geometric approach in [6] for example, for Camera $i$ and Camera $j$, a correlation coefficient $\rho_{ij}$ can be computed as the portion of the overlapped sensing area to a camera's entire field of view. For the group of camera sensors $S = \{S_1, S_2, ..., S_N\}$, the correlation among these cameras can be represented as a correlation matrix $C$, denoted as

$$C = (\rho_{ij})_{N*N}, \tag{1}$$

where $\rho_{ij}$ is the correlation coefficient of image $X_i$ and image $X_j$. The correlation coefficient satisfies: 1) $0 \le \rho \le 1$; 2) $\rho_{ij} = \rho_{ji}$; and 3) $\rho_{ii} = 1$. In the following study, we assume that this correlation matrix is obtained in advance.

### A. Entropy-based Approach

In information theory [2], 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

$$\widetilde{H} = -\sum_{k=1}^{L} p(r_k) log p(r_k), \tag{2}$$

where $L$ is the number of all possible gray-levels, and $p(r_k)$ is the probability of the $k$th gray-level [4]. It denotes the average amount of information per pixel in the image.

If a camera $S_i$ transmits its observed image $X_i$ to the sink, the amount of information gained at the sink is $H(X_i)$. (We do not consider the information loss caused by compression and transmission.) If the group of camera sensors, $S = \{S_1, S_2, ..., S_N\}$, transmit their observed images $X_1, X_2, ..., X_N$ to the sink, the amount of information gained at the sink will be the joint entropy $H(X_1, X_2, ..., X_N)$. Our objective is to estimate the joint entropy of multiple cameras.

### B. Joint Entropy of Two Cameras

We consider two cameras that can observe the area of interest. Suppose each camera has captured one image about the area of interest, denoted as image A and image B. The joint entropy of image A and image B is

$$H(A, B) = H(A) + H(B) - I(A; B), \tag{3}$$

where $I(A; B)$ is the mutual information of the two sources. $I(A; B)$ can be interpreted as the reduction in the uncertainty of one source due to the knowledge of the other source:

$$I(A; B) = H(A) - H(A|B) = H(B) - H(B|A). \tag{4}$$

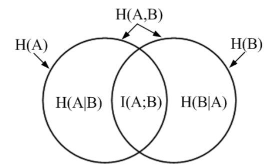


Fig. 1. Relationship between entropy and mutual information

The definition of $I(A; B)$ in probability form is given as

$$I(A; B) = \sum_a \sum_b p(a, b) log \frac{p(a, b)}{p(a)p(b)}. \tag{5}$$

where $p(a)$ and $p(b)$ are the probability distributions of the pixels in image A and image B, and $p(a, b)$ is the joint probability distribution of the two sources.

The relationship between entropy and mutual information is illustrated in Fig. 1, where the individual circles stand for the entropies of $A$ and $B$, while the overlapped area corresponds to the mutual information $I(A; B)$. Mutual information is a measure of dependence between two sources: the more $A$ and $B$ are correlated, the larger the mutual information $I(A; B)$.

In [7], a normalized form of mutual information, *entropy correlation coefficient (ECC)*, is defined as

$$ECC = \frac{2I(A; B)}{H(A) + H(B)}. \tag{6}$$

The entropy correlation coefficient ($ECC$) ranges from 0 to 1, where 0 indicates that source $A$ and $B$ are independent, while 1 indicates that source $A$ equals to source $B$. The larger the $ECC$ value, the more these two sources are correlated.

Based on (3) and (6), the joint entropy of $A$ and $B$ can be expressed as a function of $H(A)$, $H(B)$ and $ECC$:

$$H(A, B) = (1 - \frac{1}{2}ECC)(H(A) + H(B)). \tag{7}$$

Since $H(A)$ and $H(B)$ can be calculated at each camera using (2), if $ECC$ can be estimated, the joint entropy $H(A, B)$ will be obtained. However, to calculate $I(A; B)$ and $ECC$, the joint probability distribution of the two sources needs to be estimated (5). Due to the complexity of image contents and the difficulty in image modeling, it is difficult to get an accurate estimation of the joint probability distribution [7]. Besides, estimating the joint probability requires large bulk of computation [7]. If joint probability distribution is to be estimated in a sensor network, cameras at different locations must exchange their observed images, which will increase the communication burden in the network.

As introduced above, a normalized correlation coefficient, $\rho (0 \le \rho \le 1)$, can be calculated by simple geometric methods in [6]. Considering the limited processing capability of sensors, we propose to estimate $ECC$ by the correlation coefficient $\rho$. If we replace $ECC$ in (7) by $\rho$, we can obtain an estimation of the joint entropy as

$$H(A, B) \approx (1 - \frac{1}{2}\rho)(H(A) + H(B)). \tag{8}$$

Therefore, the amount of information that can be gained from image $A$ and $B$ 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.

### C. Joint Entropy of Multiple Cameras

In this section, we extend the study of joint entropy to the case of more than two cameras. Suppose there is a group of camera sensors $S = \{S_1, S_2, ..., S_N\}$ with their observed images $X_1, X_2, ..., X_N$. We are interested in estimating the joint entropy for this group of sensors. If $H(X_1, X_2, ..., X_N)$ is to be computed by its definition in probability, the joint probability distribution of these $N$ images needs to be estimated. However, it is difficult to estimate the joint probability distribution of multiple sources, especially when $N$ is large.

A feasible approach is to make use of the two cameras case in the last section. As there are $N$ individual elements in the group $\{X_1, X_2, ..., X_N\}$, we can merge two of them together, so that the joint entropy of these two elements can be calculated by (8). We treat these two elements as a whole element, then the number of elements in the group reduces to $N-1$. If we repeat this process, the $N$ individual sensors will be combined into a single element in the end. As the joint entropy of merged sensors are calculated along the merging process, the joint entropy $H(X_1, X_2, ..., X_N)$ can be obtained when the merging process is completed.

We design an algorithm to estimate the joint entropy of multiple cameras based on the idea of hierarchical clustering [5]. As long as the entropies of single images ($H(X_i), i = 1, 2, ..., N$) and the correlation matrix ($C = (\rho_{ij})_{N*N}$ (1)) are given, the joint entropy $H(X_1, X_2, ..., X_N)$ can be estimated through the hierarchical clustering process. The details of the estimation algorithm are presented in *Algorithm 1*, where $\chi$ denotes the set of clusters, and $\rho(\{X_i\}, \{X_j\})$ is the correlation coefficient between cluster $\{X_i\}$ and cluster $\{X_j\}$.

**Algorithm 1** Estimate the Joint Entropy of Multiple Cameras

$H(X_1, X_2, ..., X_N) = JointEntropy(H(X_i), (\rho_{ij})_{N*N})$
**begin**
$\chi = \{\{X_1\}, \{X_2\}, ..., \{X_N\}\}$, $\rho(\{X_i\}, \{X_j\}) = \rho_{ij}$.
**for** $k$ = 1 to $N-1$ **do**
  Find $(\{X_i\}, \{X_j\}) = arg \max_{\{X_i\},\{X_j\} \in \chi} \{\rho(\{X_i\}, \{X_j\})\}$
  Merge $\{X_i\}$ and $\{X_j\}$ into a new cluster $\{X_{N+k}\}$.
  $H(X_{N+k}) = H(X_i, X_j)$ (8).
  **for** $X_l \in \chi, l \neq i, l \neq j$ **do**
    Compute $\rho(\{X_{N+k}\}, \{X_l\})$. $(*)$
  **end for**
  Remove $\{X_i\}$ and $\{X_j\}$ from $\chi$; Add the new cluster $\{X_{N+k}\}$ into $\chi$.
**end for**
$H(X_1, X_2, ..., X_N) = H(X_{2N-1})$
**return** $H(X_1, X_2, ..., X_N)$
**end**

In step $(*)$ of *Algorithm 1*, the correlation coefficient between one cluster and another cluster can be obtained by the *greatest/shortest/average* correlation coefficient from any member of one cluster to any member of the other cluster [5], which are refered to as *single-linkage/complete-linkage/average-linkage* clustering.

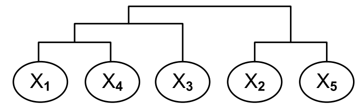


Fig. 2. An example of hierarchical clustering

TABLE I
HIERARCHICAL CLUSTERING STEPS

| Steps | Nodes for Clustering | Estimation of joint entropy (Relative value to $H(\cdot)$) |
|---|---|---|
| 1 | $\{X_1\}$,$\{X_4\}$ | $H(X_1, X_4) = 1.0557$ |
| 2 | $\{X_2\}$,$\{X_5\}$ | $H(X_2, X_5) = 1.2641$ |
| 3 | $\{X_1X_4\}$,$\{X_3\}$ | $H(X_1, X_3, X_4) = 1.7290$ |
| 4 | $\{X_1X_3X_4\}$,$\{X_2X_5\}$ | $H(X_1, X_2, X_3, X_4, X_5) = 2.9931$ |

The following is an example of the estimation of joint entropy. Suppose there is a group of five camera sensors. Without loss of generality, we assume that the entropy of a single image is a constant value, denoted as $H(X_i) = H(\cdot)(i = 1, ..., 5)$. A correlation matrix for these five sensors is given by

$$(\rho_{ij})_{5*5} = \begin{pmatrix} 1 & 0 & 0.2942 & 0.9443 & 0 \\ & 1 & 0 & 0 & 0.7359 \\ & & 1 & 0.3416 & 0 \\ & & & 1 & 0 \\ & & & & 1 \end{pmatrix}. \quad (9)$$

Apply *Algorithm 1* to this group of sensors, and use the *average-linkage clustering*[5] metric in step $(*)$. The clustering process is illustrated in Fig. 2, and the results in each step of clustering are shown in Table III. From the correlation matrix (9) and the illustration in Fig. 2, one can find that in every clustering step, sensors that contain the most correlated images are merged into one cluster. As can be seen from (8), the value of the joint entropy decreases as the correlation degree of the two images increases. Thus, the joint entropies obtained from the clustering process are always relatively small. The final result of the estimation algorithm is a conservative estimation of the joint entropy $H(X_1, X_2, ..., X_N)$.

## III. Correlation-based Camera Selection

### A. Correlation-based Camera Selection

Suppose for an area of interest in a WMSN, a total number of $N$ cameras can observe the area of interest. If network resources permit, all these cameras can transmit their observed images to the sink, so that the applications at the sink can gain comprehensive information about the area of interest. However, as the processing capabilities of sensors are limited, and the communication among sensors causes huge energy consumption, sometimes the network cannot support all these cameras to report their observations to the sink. Consequently, we propose a camera selection problem: if only $M$ cameras ($M \leq N$) 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.

As in Section II, we also assume that the entropy of a single image is a constant value here. The estimation of joint entropy in (8) indicates that the less correlated are the two sensors, the more information can be provided by the two sensors together. Thus, to maximize the joint entropy of $M$ cameras, we should try to minimize the correlation among the cameras to be selected. We propose a *correlation-based* algorithm to maximize the joint entropy of $M$ cameras. At each step of the algorithm, we select one camera that is least correlated with the cameras that have already been selected. The details are presented in *Algorithm 2*, where $\chi = \{X_1, X_2, ..., X_N\}$ is the set of images observed by these $N$ cameras, and $S$ denotes the set of cameras that are already selected.

**Algorithm 2** Correlation-based Selection of Cameras

$S = CorrSelection(\{X_1, X_2, ..., X_N\}, (\rho_{ij})_{N*N}, M)$
**begin**
$S = \emptyset$, $\chi = \{X_1, X_2, ..., X_N\}$, $\rho(X_i, X_j) = \rho_{ij}$.
Find $(X_i, X_j) = arg \min_{X_i, X_j \in \chi} \{\rho(X_i, X_j)\}$
Add the corresponding $X_i$ and $X_j$ into $S$. $\{M = 2\}$
**if** $M > 2$ **then**
  **for** $k$ = 1 to $M - 2$ **do**
    **for** $X_l \in \chi, X_l \notin S$ **do**
      $\rho(X_l, S) = \max_{X_j \in S} \{\rho(X_l, X_j)\}$.
    **end for**
    $X_m = arg \min_{X_m \in \chi, X_m \notin S} \{\rho(X_m, S)\}$; Add $X_m$ into $S$.
  **end for**
**end if**
**return** $S = \{X_{i1}, X_{i2}, ..., X_{iM}\}$
**end**

### B. A Distortion Function

For an area of interest in WMSN, we suppose a total number of $N$ cameras can observe this area, and denote their observed images as $\{X_1, X_2, ..., X_N\}$. The joint entropy of all these $N$ sensors, $H(X_1, X_2, ..., X_N)$, is the maximum amount of information that can be gained for the area of interest. If a subset of these sensors, denoted as $\{X_{i1}, X_{i2}, ..., X_{iM}\}$, are selected to report their observed images to the sink, the information gained at the sink is $H(X_{i1}, X_{i2}, ..., X_{iM})$.

We define a distortion function as the ratio of the decrease in the amount of information to the maximum amount of information, given by

$$D = \frac{H(X_1, X_2, ..., X_N) - H(X_{i1}, X_{i2}, ..., X_{iM})}{H(X_1, X_2, ..., X_N)}, \quad (10)$$

The value of $D$ satisfies $0 \leq D \leq 1$. It can be interpreted as the percentage of information loss due to network resource constraints. Applications of WMSNs can use this distortion function as a metric to describe its requirements. For example, an application may ask the network to transmit information within 10% of information loss. According to the derivation of joint entropy in *Algorithm 1*, we can find that the value of distortion is related to the number of selected cameras, as well as the correlation degrees among the selected cameras.

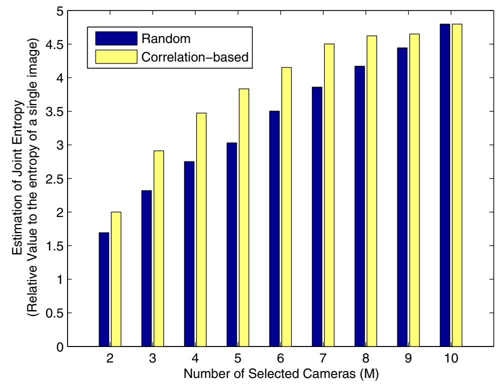


Fig. 3. Estimation of joint entropy.

## IV. PERFORMANCE EVALUATION

In this section, we present a set of simulations to evaluate the joint effect of multiple cameras and the correlation-based camera selection algorithm. In a field of 500m*500m, we set an area of interest that is located in the center of the field and has a radius of 10 meters. We randomly deploy $N$ cameras that can observe this area of interest. Let $M$ be the number of cameras to be selected by the sink to transmit their observed images. Suppose each camera obtains one image about the area of interest. Let $\{X_1, X_2, ..., X_N\}$ denote the images observed by these $N$ cameras, and let $\{X_{i1}, X_{i2}, ..., X_{iM}\}$ denote the images observed by the $M$ selected cameras.

Without loss of generality, we assume that the entropy of a single image is a constant value, denoted as $H(X_i) = H(\cdot)(i = 1, 2, ..., N)$. For these $N$ cameras, we can obtain a correlation matrix $(\rho_{ij})_{N*N}$ as introduced in Section II. So the joint entropy of $H(X_1, X_2, ..., X_N)$ and $H(X_{i1}, X_{i2}, ..., X_{iM})$ can be estimated using *Algorithm 1*.

We compare the following two camera selection schemes:

*1) Random selection:* Randomly select $M$ cameras out of the $N$ cameras. For each M, repeat the experiment for 50 times. Compute the joint entropy at each time, and take the average value of the 50 trials as the final joint entropy.

*2) Correlation-based selection:* This is the proposed scheme described in *Algorithm 2*. It makes use of correlation by selecting a group of $M$ cameras that are least correlated with each other, so that the amount of information from the selected cameras can be maximized.

In our first experiment, we randomly deploy 10 cameras in the field ($N = 10$), and let $M$ change from 2 to 10. The results of both schems are shown in Fig. 3. The value of joint entropy increases as the number of nodes increases, which indicates that if more cameras transmit their observed images to the sink, more information can be gained about the area of interest at the sink. When $M = 10$, all the cameras are selected to transmit their obsereved images, so both schemes

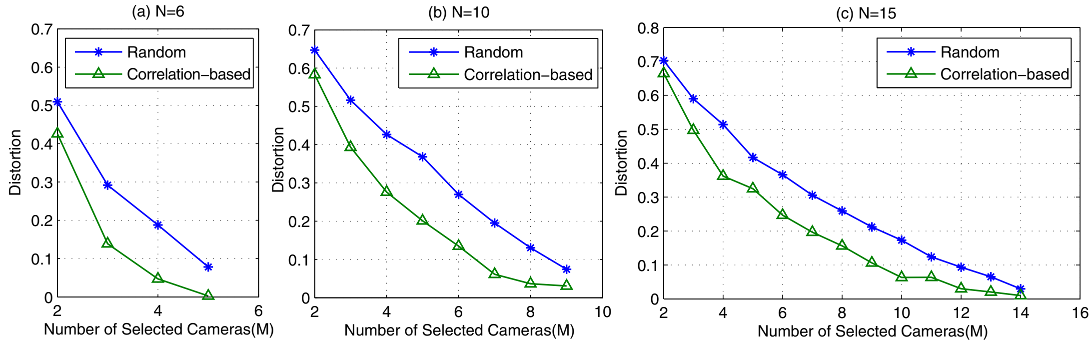


Fig. 4. The Distortion Function.

produce the same results. But for $M = 2$ to 9, the correlation-based algorithm always results in larger joint entropy than the random selection of cameras.

According to the numerical results, when the number of selected cameras are the same for these two schemes, the correlation-based algorithm can increase the joint entropy by $0.5466 * H(\cdot)$ in average (increase by 18.37% in average compared to the random selection algorithm). It should be noted that the values of joint entropy in our simulation are expressed as relative values to the entropy of a single image, $H(\cdot)$. We find in our experiments that a typical value of $H(\cdot)$ is 5-6 bits/pixel for images of 8-bits depth, so the correlation-based scheme can result in about 3 bits/pixel increase in joint entropy than the random selection scheme.

Next we introduce more simulations to evaluate the distortion performance of both schemes. We implement both schemes for three different network topologies, where the total number of cameras, $N$, equals to 6, 10, and 15, respectively. Fig. 4 plots the distortion performance of both schemes. The distortion decreases as the number of selected nodes increases. For the same number of selected cameras ($M$), the proposed correlation-based scheme results in lower distortion compared to the random selection scheme.

From another perspective, if a certain distortion bound is required at the sink, we may need less cameras to transmit their information using the correlation-based selection scheme. For example, in Fig. 4(b), a total number of 10 cameras are deployed to observe an area of interest. If the sink wants to obtain 80% of the total information, the maximum distortion is 0.2. As shown in Fig. 4(b), 7 cameras are needed on average when cameras are randomly selected, but only 5 cameras are needed when the correlation-based selection scheme is used. Therefore, if given a distortion bound at the sink, the correlation-based selection scheme requires fewer cameras to report to the sink than the random selection scheme.

## V. Conclusion

In this paper, we study the joint effect of multiple correlated cameras in WMSNs. We propose an entropy-based analytical framework to measure the amount of information provided by multiple cameras in WMSNs. A correlation-based camera selection scheme is also introduced to select cameras from sensor networks under distortion constraints. Simulation results show that, given a distortion bound at the sink, the correlation-based selection scheme requires fewer cameras to report to the sink than the random selection scheme.

The proposed camera selection algorithm aims to minimize the number of cameras to report to the sink under distortion constraints. In our future work, we can consider more factors for the camera selection problem, such as the residual energy of sensors, the locations of correlated cameras, the costs of communication between correlated cameras, and the costs for encoding correlated images.

## Acknowledgment

This material is based upon work supported by the US National Science Foundation under contract ECCS-0701559.

## References

[1] I. F. Akyildiz, T. Melodia, K. R. Chowdhury, "A Survey on Wireless Multimedia Sensor Networks," *Computer Networks (Elsevier)*, January 2007.

[2] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, NY, 1991.

[3] R. Cucchiara, "Multimedia Surveillance Systems," *Proceedings of ACM International Workshop on Video Surveillance and Sensor Networks*, pp. 3-10, November 2005.

[4] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*, Prentice Hall, 2004.

[5] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data Clustering: A Review," *ACM Computing Surveys*, vol. 31, no. 3, September 1999.

[6] H. Ma and Y. Liu, "Correlation Based Video Processing in Video Sensor Networks," *IEEE International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 2, pp. 987-992, June 2005.

[7] J. P. W. Pluim, J. B. Antoine Maintz, and M. A. Viergever, "Mutual-Information-Based Registration of Medical Images: A Survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, August 2003.

[8] R. Wagner, R. Nowak, and R. Baraniuk, "Distributed Image Compression for Sensor Networks Using Correspondence Analysis and Super-resolution," *IEEE International Conference on Image Processing ICIP 2003*, vol. 1, pp. 597-600, September 2003.

[9] M. Wu, C. W. Chen, "Collaborative Image Coding and Transmission over Wireless Sensor Networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 70481, 9 pages, 2007.