Chapter 9
Data Processing and Exchange Challenges in Video-Based Wireless Sensor Networks

Dan Pescaru
Politehnica University of Timisoara, Romania

Daniel-Ioan Curiac
Politehnica University of Timisoara, Romania

ABSTRACT

This chapter presents the main challenges in developing complex systems built around the core concept of Video-Based Wireless Sensor Networks. It summarizes some innovative solutions proposed in scientific literature on this field. Besides discussion on various issues related to such systems, the authors focus on two crucial aspects: video data processing and data exchange. A special attention is paid to localization algorithms in case of random deployment of nodes having no specific localization hardware installed. Solutions for data exchange are presented by highlighting the data compression and communication efficiency in terms of energy saving. In the end, some open research topics related with Video-Based Wireless Sensor Networks are identified and explained.

INTRODUCTION

Wireless Sensor Networks (WSNs) technology is nowadays widely used in various domains. It has applications in fields such as emergency rescue, environmental monitoring, military operations, at-home medical care or industrial systems. A wireless sensor network consists in a set of network nodes capable of sensing and wireless communication. It operates in the absence of a pre-deployed infrastructure and can work in hostile environments. Nodes are self-configurable, low power, low cost, and can be rapidly deployed in emergency situations. Their sensors interact with the physical environment by monitoring and measuring light, heat, position, movement, chemical presence, etc. The information from sensors is then delivered to the other nodes over the wireless

DOI: 10.4018/978-1-4666-5978-0.ch009

network. In many applications one more powerful node, known as central point (CP), gathers the information from sensor nodes, processes it, and interprets the results.

Special kinds of WSNs are represented by Video-based Wireless Sensor Networks (VWSNs), in which case large amounts of video data are sensed, processed in real-time and then transferred over the wireless networks (Sánchez, 2012). Among traditional applications, video monitoring for environment surveillance covers an area that focus attention nowadays due to more frequent threats posed by hurricanes, earthquakes or terrorist attacks.

One important problem in this context is related with data storage and exchange. Indeed, acquisition of a video sequence with reasonable frame rate implies significant amount of data that needs to be stored and transferred.

Handling video data usually required large storage buffers. These buffers help multi-frames video encoding/decoding process but also ensure temporary storage for multi-hop data transfer. Several hardware platforms where developed to provide large data buffers for such intensive data flows as eCAM (Chulsung, 2006), Cyclops (Mohammad, 2005) and RISE (Zeinalipour, 2005).

The wireless communication is characterized by noise, path loss, channel fading and interference. The result is a wireless channel having much lesser capacity than a wired one. Moreover, WSN multi-hop routing tends to generate more interference, delay, packet loss and higher number of errors during transmission. High packet loss rate on the path affects the bandwidth and delay values of transmission. Consequences depend on the application domain and on the kind of implemented system. All application has specific service requirements from the network usually expressed through a parameter named Quality of Service (QoS). Video surveillance using VWSNs in particular have a more constrained set of QoS requirements, aimed to sustain transmission of high quality data at a high bit-rate. Many of them require strict end-to-end delay, bandwidth and jitter guarantees. These parameters are hard to be satisfied not only due to mentioned communication issues but also because video encoding/decoding algorithms that involve significant processing time.

The aim of the chapter is to debate various solutions for data processing and exchange in video-based wireless sensor networks and to point out some open issues in this field.

VIDEO-BASED WIRELESS SENSOR NETWORKS

Combining video surveillance with wireless sensor networks brings important advantages in many fields. Resulting video-based wireless sensor networks have a large applicability especially in surveillance of critical zones to detect suspect activities. Beside obvious military applications, a lot of systems were developed for surveillance in subway and train stations, airports, hospitals, parking zones, stores, and other public places (Fernandez, 2013). Along common intrusion detection tasks, these systems can be used also to identify persons, vehicles or other kind of targets.

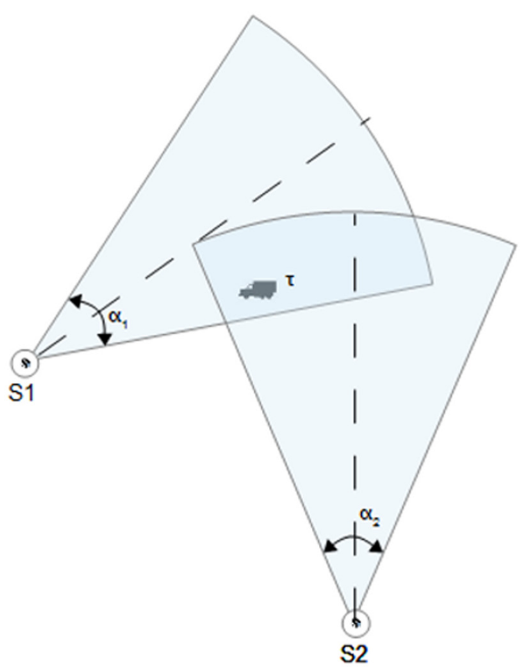
Another class of applications was designed for environmental monitoring in the case of areas subject to earthquakes, flooding or other natural disasters (Dawood, 2013). Sensor nodes can be deployed in the risk area to collect images over a wide surface. A disaster headquarter will use the information to take the best management decisions to overcome the situation.

Despite the fact that object sensed by the camera can be at arbitrary locations, information quality strongly depends on camera's resolution, size of the object and distance between camera and object. Depending on particular application and on the size of the smallest interesting object, we can determine experimentally the medium distance D_q, which provide adequate quality. This distance is known as camera range or depth of view.

Data collection from video sensor assumes a target entering field of view (FoV) coverage area. Due FoV overlapping, multiple sensors could sense the same target simultaneously as depicted by Figure 1. Indeed, two camera sensors can collect visual data of the same object even if they are far from each other. Same time, a very close object may not be viewed by a particular visual sensor if it isn't inside its FoV. Generally, the collected information depends on the sensors orientation, resolution and depth of view but also on light or environment conditions (Costa, 2010).

Target discovery and tracking become more complex in case of mobile sensors or cameras with pan, tilt, and zooming facilities. In these scenarios, sensors movement or cameras' parameters variation allow dynamic adjustment of the FoV (Desai, 2009).

Figure 1. Collecting visual data in case of FoV overlapping



CHALLENGES AND SOLUTIONS IN VWSN VIDEO DATA PROCESSING

Issues Regarding Deployment, Coverage and Video Processing

Wireless sensor networks are generally composed of a large number of tiny sensors densely deployed over the area when the phenomenon of interest is happening. The deployment could be deterministic or random. In case of deterministic deployment, the process implies a prior prepared plan. This approach allows maximization of the covered area with the minimum number of sensors. Sometimes a number of redundant nodes are added for higher availability to ensure optimal performances. However, there are many real-life situations when deterministic deployment cannot be a solution. When we take into consideration a harsh or a hard-to-reach environment the random deployment might be the unique feasible solution.

In case of a random deployment the sensors will be scattered over a target area. Approaches varies from airdropping the sensors from a plain, launching them using a rocket or releasing them on the ocean. The chosen approach is strongly depending on the application monitoring requirements and the characteristics of the area under investigation. This may result in regions densely or sparsely covered by the sensor nodes. The solution is to increase the density of the deployment in regions of interest with higher relevance for the application to enhance the overall monitoring quality. Higher density ensures also a certain level of sensing redundancy. The redundancy is valuable for a WSN as it increases the precision of environment observations and network lifetime (Pescaru, 2008). Benefits of redundancy in coverage recovery are demonstrated in (Istin, 2011), where FoV recovery problem was debated in context of traffic surveillance applications. The problem of redundancy was deeply analyzed in (Curiac, 2009). In addition to obvious physical sensing redundancy, temporal sensing redundancy, temporal communication redundancy, and information redundancy adds high relevant data for deployment planning.

Obviously, not all visual sources have the same relevancy for a particular VWSN application. The significance of each of them is weighted with the importance of the target, the overlap with some regions of interest or with observation conditions. A well-defined concept of sensing relevance in video-based wireless sensor networks is proposed in (Costa, 2013). Five different groups of relevance related to the overall significance of the source nodes are identified as irrelevant, low relevance, medium relevance, high relevance and maximum relevance. Using this classification, redundant nodes are considered only if they fall in the same group of relevance.

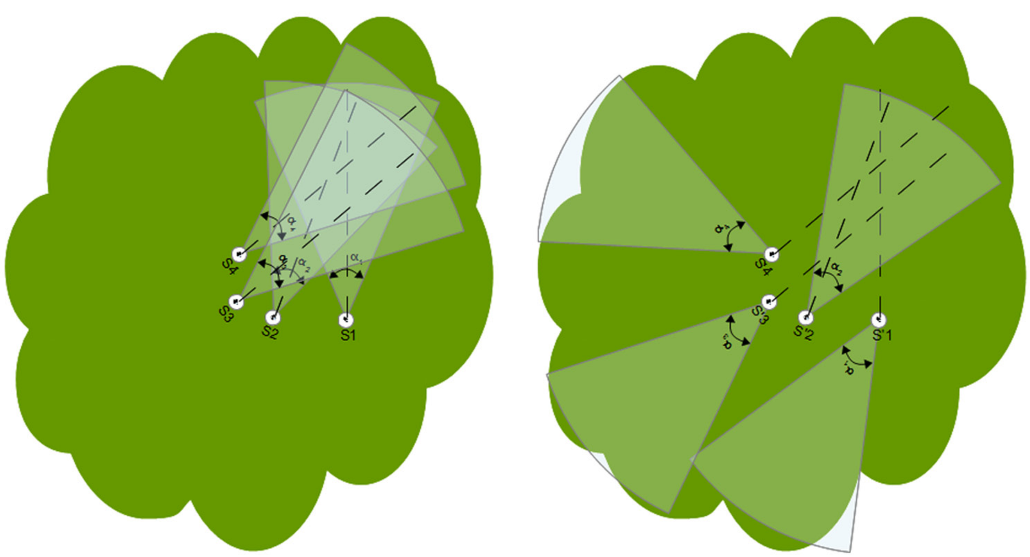
The success of a VWSN application relies many times on area coverage. Although the problem of WSN coverage was intensively studied, in the case of a VWSN it becomes more difficult. Unlike omni-directional disc sensing model of general sensors, the sensing model for video cameras induces complex deployment-related situations. Here the viewing direction of the sensor has significant impact on the quality of coverage over the target surface. Figure 2 demonstrates the large variation of coverage in two deployment scenarios that imply the same number of sensors.

After deployment, various classes of visual processing algorithms are required to fulfill the needs of VWSN applications. These include, but are not limited to:

- Image and video capturing and compression;
- Features extraction;
- Objects detection and tracking;
- Data aggregation, transfer and security;
- Distributed image processing.

Due to the limitation of communication bandwidth, it is not feasible to constantly stream video to a central server having high computation capabilities. On the other hand, many times it is unfeasible to move visual processing at the level of network nodes. Therefore, most solutions combine distributed processing, aggregation and central point processing to achieve desired result.

Figure 2. Area FoV coverage for two deployment scenarios involving four video sensors



Proposed Solutions for Deployment, Coverage and Video Processing

Deployment and Coverage

Most of the VWSN applications are employed in surveillance tasks. The success of a surveillance operation relies not only on some sophisticated visual processing algorithms, but also it depends on good coverage of the investigated area. The problem of coverage has to be solved during the deployment/redeployment phases.

In the case of video-based wireless sensor networks, the deployment plan or the density variation has to take care of FoV coverage. For deterministic deployments, various algorithms for optimal camera placement (Oasis, 2010) have been proposed. In all of them, the goal is to find the minimal number of nodes that can view the larger or most interesting area of the monitored environment. This can be rather difficult in the case of a complex environment with significant number of visual obstacles. The work in (Adriaens, 2006) proposed a polynomial time algorithm to compute worst-case coverage, which is related with the maximal distance between the mobile target and the sensors. In case of random deployment, the optimal deployment density can be determined using various probabilistic approaches.

In general, deployment strategies are based on coverage estimation. Coverage can be expressed using various metrics. A Directional K-Coverage – *DKC* – metric is proposed in (Liu, 2008), adapting the concept previously defined in (Huang, 2003) in order to consider directional visual monitoring. *DKC* is defined as a probability guarantee, since 100% coverage is very difficult to achieve for randomly deployed visual sensors with a uniform density. Reference (Istin, 2007) proposes a set of metrics, particularly relevant for surveillance systems. The first metric denotes the percentage of covered surface relative to the total deployment surface – *CS/S*. Its computation is straightforward and the conclusions drawn are useful for most applications. A refined variant of *CS/S* is the size of the Maximum Continuous Uncovered Surfaces over the monitored area – *MCUS*. It is especially important if we consider tracking applications. The aim is to reveal how much a target can move in the area without being noticed by the network. Several experiments demonstrate the saturation effect obtained for random deployments of VWSN with different number of nodes as presented in Table 1.

The deployment homogeneity could be analyzed using the total Number of Continuous Uncovered Surfaces – *NCUS*. To estimate the coverage closure, it is proposed a metric named Number of Crossing Paths – *NCP*. Here, a crossing path is considered a way between two borders of the guarded area uncovered by any sensor FoV. The *NCP* will count the number of different uncovered paths crossing the network. All paths, starting from the same uncovered surface and ending on other, have to be counted only once.

Table 1. The variation of Covered Surface and Maximum Continuous Uncovered Surface considering random deployments of wireless sensor networks with size variation between 0 and 10,000 nodes on a 1,000x1,000 m^2 monitored area

Network Size (# of nodes)	Covered Surface (%)	Maximum Continuous Uncovered Surface (%)
1,000	38.21	23.68
2,000	60.13	7.06
3,000	73.87	3.24
4,000	82.59	2.07
5,000	88.91	1.04
6,000	93.03	0.98
7,000	95.25	0.56
8,000	96.31	0.28
9,000	97.06	0.14
10,000	99.08	0.03

An analysis of the impact of coverage on surveillance quality after a random deployment over an area of interest is presented in (Pescaru, 2007). It is based on Relevant Camera Sensing Area – *RCSA* – parameter, which is defined as a sector resulting from the intersection of field o view and the monitored area. Based on that, the Network Relevant Sensing Area – *NRSA* – for whole video-sensor network is expressed as the union of all individual relevant cameras' sensing areas (1) as presented in Figure 3.

$$NRSA = \bigcup_{i=1}^{N} RCSA_i \qquad (1)$$

Using NRSA we can express an important deployment quality parameter calculated as the ratio between NRSA and network deployment area – *NDA*. We named it Deployment Coverage Quality – *DCQ*.

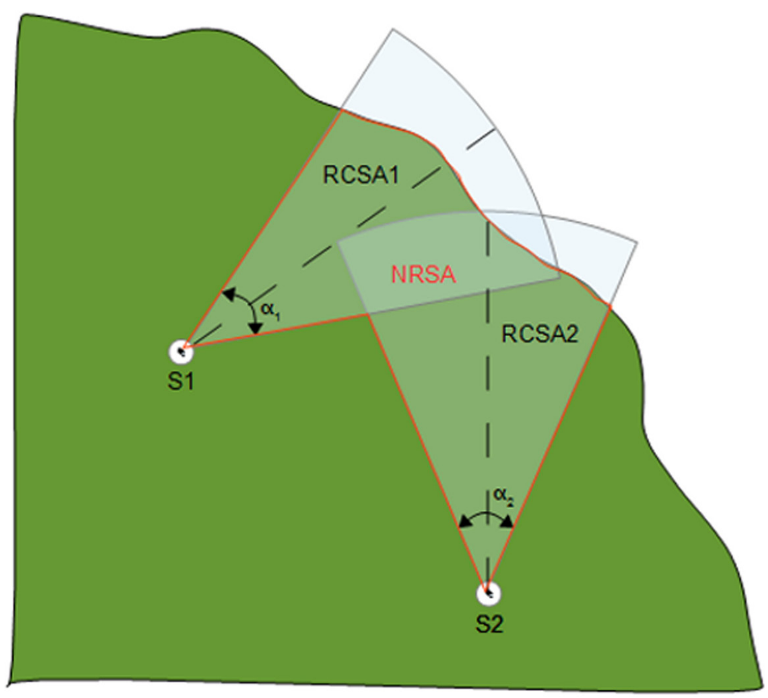
$$DCQ = \frac{NRSA}{NDA} \qquad (2)$$

The coverage can be used to switch off part of the redundant nodes. The lifetime of the sensor nodes is significantly longer in this case, as they do not have to operate the camera during idle time periods.

Localization

A wireless sensor network is often deployed in an ad-hoc manner in the absence of any knowledge on existing infrastructure or location characteristics. Moreover, redeployment of additional nodes could happen at any time when the upgrade of capabilities or replacements of malfunctioning nodes are necessary. Considering these situations, a central problem is the estimation of spatial-coordinates of the network nodes, known as topology extraction or localization. The solutions for this problem can

Figure 3. Network Relevant Sensing Area (NRSA)



be classified into coarse-grained methods (Bulusu, 2000) based on proximity to a reference point, and fine-grained localization (Savvides, 2001) based on timing or signal strength.

Popular localization methods make use of Global Positioning System (GPS). Unfortunately, this solution can work only outdoors and the cost of GPS sensor is high. Therefore, it is not suitable for development of small cheap sensor nodes as desired for a massive deployment. Other coarse-grained solutions consider the network to be organized as a mixed hierarchy, built using both complex nodes and cheap low-level nodes. The complex high-level nodes are considered to know their location, by using GPS or other techniques. These nodes act as beacons and transmit their position periodically via the network. The low level nodes can run different localization proximity-based algorithms, including sophisticated iterative multi-lateration or multi-angulation (Langendoen, 2005).

Most of the fine-grained localization solutions are based on the timing of arrival or on the signal strength. Methods from the first category use the distance between the node and a reference point determined by measuring the communication signal arrival time (Meghani, 2012). The methods based on signal strength consider the signal attenuation proportional to the traveled distance and use this information to approximate the relative location (Elnahrawy, 2004). More elaborate methods rely on signal pattern matching techniques and use pre-scanned coverage area transmitted with signals. A central system assigns a unique signature for each square in the location grid. The system looks for a match between the signal received and those from the pre-constructed database and determines the correct location. A better solution uses measurements of the Doppler shifts of the transmitted signal and radio interferometric techniques (Kusy, 2007).

Reference (Pescaru, 2006) proposes a novel solution for node localization and video-filed overlap estimation. It starts from video images acquired from network sensors and computes the video fields' superposition. Data are gathered and processed on a central point, which extracts parameters for coordinates translation, rotation and scaling. Parameters are sent back to all nodes and used to calculate FoV overlaps between pairs of sensors. Processing consists on image registration algorithm applied on each pair of images, coming from individual sensors. Registration involves searching for corresponding elements between source and target images. In practice, correspondence is established based on features extracted from images or based on similarity between regions. As the region based registration is prone to errors generated by segmentation and cameras' color sensitivity, feature based registration is considered more reliable. As all registration algorithms are hard computational, a distributed solution based on node resources is not feasible.

The initial images have to represent the monitored area at a certain moment of time. Therefore a setup protocol was designed to ensure quasi-simultaneous data collecting. It retrieves images on an initialization phase and allows the central point to calculate all necessary transformation parameters between each pair of cameras.

The protocol was adapted from the one presented in (Cosma, 2006), which demonstrates good performance in case of large networks. It starts with a setup message broadcasted by the central point. This message embeds an empty routing table for paths information. Each node that receives the message should test if it is included or not into the message routing table. If not, it has to add itself to the routing table of the original message and broadcasts the modified message. After broadcasting, it starts a timer that manages the image capture and transmission. Each node acts as a router by propagating the server requests to the entire network. When a node receives for the first time a setup message it will forward this message to its neighbors. After that, an image is sent to the central point in an energy-efficient manner that preserves network integrity. To increase efficiency each node keeps routing information that will describe each of its neighbors by its current energy level and the hop-count until the sink. All received messages will update also the energy information of the sender into the receiver routing table. As long the next hop node energy is over 50% the only restriction considered in electing it is to have minimum hop count to the sink. When energy node is between 20%-50%, the node will be elected as next hop only if there are no other neighbors with equal hop-count and higher energy level. In case of multiple candidates the possible next hop is computed based on a LRU like algorithm. This strategy ensures an optimal balance for power consumption along the network during images collection process.

After gathering all the images from the network the central point starts to calculate geometric transform parameters between all pairs of matching images. This operation is based on image registration algorithms. The goal of image registration is to overlay two images of the same scene captured at different times, from different viewpoints and by different cameras. To accomplish this goal, a transformation has to be found so that several points in one image, called the reference image, can be mapped to corresponding points in a second one. In other words, registration geometrically aligns two images in an optimal manner. One of the first proposed registration methods was RANSAC (Fischler, 1981). It is based on a robust estimator. The main drawback of this solution is represented by the substantial computation time needed in most of the cases. Practical solutions do not consider all the available data if corresponding reduction of its precision is acceptable in the application context. Besides of that, due to the diversity of images to be registered and due to the various types of visual degradations encountered, it is practically impossible to design a universal method applicable to all situations. In addition to possible geometric deformations, radiometric deformations and noise have to be considered.

The registration method presented in reference (Pescaru, 2006) makes use of a mean shift algorithm for robust parameters estimation (Comaniciu, 2002) to compute translation vector, rotation angle, and scaling factor. It relays on a semi-automated process to establish the set of corresponding features. Results are better than in case of using RANSAC, as this method considers all the available data samples. To deal with various errors that affect the algorithm in harsh situations, a post-processing phase could be applied.

The mean shift algorithm detects local maxima of a multivariate probability density. The computed parameters are selected from the values with the highest probability density in considered solution space. Estimation of the density vector δ starts from a sample of N k-dimensional data points χ_i, drawn from a distribution with multivariate probability density function $p(\delta)$

$$\hat{p}_B(\delta) = \frac{1}{N}\sum_{i=1}^{N} K_B(\delta - \chi_i)\,, \tag{3}$$

where K_B is the kernel function expressed by the equation

$$K_B\left(\delta\right) = \mid B \mid^{-1/2}\ K_B\left(B^{-1/2}\delta\right). \tag{4}$$

The kernel function depends on the bandwidth matrix B, which is a symmetric positive $k\times k$ matrix. Frequently B has a diagonal form $diag\left[b_1^2,\dots,b_d^2\right]$ or a proportional to the identity matrix $b^2 I$. Considering that, the profile of the radial symmetric kernel was defined as

$$K^R\left(\delta\right) = \xi\, K(\|\,\delta\,\|^2)\,, \tag{5}$$

where ξ represents a normalization constant, assumed strictly positive. A function $g(\delta) = -k'(\delta)$ could be now defined assuming existence of the derivate of the kernel profile for all $\delta \in \left[1,\infty\right)$, excepting a finite set of points. Considering this function, the mean shift algorithm is then used to find the location of the maxima of the estimated probability density function, which is closest with a starting location γ_i. The searching is conducted by iterating until reaching convergence the equation

$$\gamma_{j+1} = \frac{\sum_{i=1}^{n} \delta_i g\left(\left\| \frac{\gamma_j - \delta_i}{b} \right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\| \frac{\gamma_j - \delta_i}{b} \right\|^2\right)}, \quad j = 1, 2, \ldots \tag{6}$$

for all $\{\gamma_j\}_{j=1,2,\ldots}$ representing the sequence of successive locations of the kernel *G*. In practice the convergence is very fast involving a very small number of iterations.

The registration model is based on a set of 2D transformations having the propriety of shape preserving mapping. The model is defined by Equation (7), which relates a pair of corresponding pixels (α, β) from two images.

$$\begin{bmatrix} \beta_x \\ \beta_y \end{bmatrix} = \begin{bmatrix} \sigma & 1 \\ 1 & \sigma \end{bmatrix} \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} + \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix} \tag{7}$$

The σ, φ, τ_x and τ_y represents scaling, rotation and translation parameters. They can be unambiguously determined from the correspondence of two pairs of points. However, in real situations precision is affected by visual errors. A good solution exploits set of pairs of corresponding points for parameter estimation. As the angle between two line segments is not depending on translation or scaling, the rotation parameter φ can be estimated prior to translation or scaling parameters. Next, the scaling parameter can be computed prior to translation, based on the distances between known pairs of points. The last estimation concerns the translation vector, as it requires coordinates compensated accordingly to rotation and scaling parameters. Figure 4 presents an example of image registration on two sensor images in a typical outdoor environment.

The algorithm is affected by small errors due to noises and limited precision of point matching procedure. A fine-grade post-processing step could be used to reduce this error. In reference (Pescaru, 2006) the processing is based on chamfer matching (Borgefors, 1984). This technique relies on edges matching. The matching criterion is the correlation of a searched pattern with a distance map computed over a target image. It involves several steps. First, a distance transformation is computed and corresponding distance map is produced starting from the upper right corner of the second image. Then, a relevant pattern extracted from the first image is then moved over the relief

Figure 4. The results of registration process

defined by the distance map. Under the action of gravity, the pattern slides over the relief until it reaches the lowest possible altitude. If this altitude is close to zero, the result corresponds to an optimal matching pattern. The pattern is located where this correlation reaches an absolute minimum. Figure 5 depicts the result of post-processing in case of a complex view. Two sensors having different orientation have captured a pair images over the common area. To highlight the processing results, a pair of corresponding rectangular fragments from registered images was investigated. A Kenny edge detector (Kenny, 1986) was used to extract edges in considered fragments from original image coming from the first sensor and the image from the second sensor after registration process. Images from the figure present the binary difference between edges from the two fragments before, and after chamfer matching adjustment. It highlights minimization of non-aligned edges in case of post-processed image.

An improved version of discussed approach is presented in reference (Fuiorea, 2008). The aim was to replace user intervention in features selection through a full automatic procedure. The selection is based in this case on a Scale-Invariant Feature Transform – SIFT (Lowe, 1999). It is a computer vision algorithm designed to detect and describe local features in images. These local features are invariant to image scale and rotation. Detection is robust to changes in illumination, noise, occlusion, and minor changes in view area. Extracted local features allow easy objects identification with high probability. Features extraction involves four steps. The first one is the scale-space extrema detection, followed by key-point localization, orientation assignment and computation of a local image descriptor. Key locations are defined as maxima and minima of the result of the difference of Gaussians function. They are applied in scale space over a series of resampled images. Scale-space extrema detection implies a convolution applied on the image using Gaussian filters at different scales. Then the difference of successive Gaussian-blurred images is computed. Key-points are generated as maxima/minima of the Difference of Gaussians (*DoG*) occurring at multiple scales.

$$DoG(x,y,\sigma) = L(x,y,k_i\sigma) - L(x,y,k_j\sigma) \tag{8}$$

The L represents the convolution of the original image I with a Gaussian blur G as presented in Equation (9).

$$L(x,y,k\sigma) = G(x,y,k\sigma) \cdot I(x,y) \tag{9}$$

Low contrast candidate and poorly localized points along an edge are discarded during key-point localization step. Orientation assignment

Figure 5. Differences between registered images before and after chamfer matching

step allows only dominant orientations to generate localized key-points. During this step gradient magnitude and direction are computed for every pixel in a neighboring region around the key-point that belongs to the Gaussian-blurred image *L*. These steps ensure invariance to image location, scale and rotation. The last step, which aims computation of local image descriptors, generates key-points that are highly distinctive and partially invariant to some other variations as illumination or 3D orientation. Following this procedure, SIFT descriptors robust to local affine distortion are obtained by considering pixels around a radius of the key location, blurring and resampling of local image orientation planes. This approach generates large numbers of features that densely cover the image over the full range of scales and locations.

After local feature extraction a robust estimation method could be used to give the best estimation of the geometrical transform parameters through features mapping between set of image pairs. The registration approach using automatic feature extraction proves strong benefit in terms of execution time, while losing in precision versus semi-automatic procedure is acceptable.

Data Processing

Data processing in VWSN represents an interdisciplinary field that combines networking, distributed and embedded computing and computer vision. Unlike wired camera networks, streaming video to a central server is not feasible due to limited bandwidth. Therefore, visual processing should be distributed between network nodes having very limited computing power and energy and the central point.

The problem implies high efficient solution at every level of VWSN. Regarding appropriate hardware, the main components include a powerful digital signal processor, a low power camera sensor and an efficient wireless module. A large variety of processors prove already enough capabilities for such applications. As examples we can mention Cyclops, WiSN, WiCa, and Citric platforms.

Cyclops (Rahimi, 2005) was one of the first working solutions for VWSN applications. Cyclops consists of an Agilent Technology ADCM-1700 CMOS video sensor, an ATMEL ATmega128L micro-controller, a Xilinx XC2C256 CoolRunner complex programmable logic device, an external SRAM and a Flash memory. The platform was designed as an external attached sensor to a WSN mote such as one from the Mica family, and therefore it does not include a radio device. The video sensor is capable of 352×288 CIF resolution. The microcontroller operates at 7.37 MHz at 3.3V. The Xilinx XC2C256 device implements synchronization and memory control that is required for image capture. Cyclops technology was designed for minimal power consumption to enable large-scale deployment and extended lifetime. As a consequence, it has a strong limitation in processing capabilities and it was not designed for applications that require high-speed processing or high resolutions. However, it can be used for various applications that rely for example on detecting changes in size or shape of the objects.

The WiSN project (Downes, 2006) proposes flexible and expandable mote architecture for distributed image sensing and processing. The board is organized around an Atmel AT91SAM7S 32 bit processor, clocked at 48 MHz. It includes 64 MB of RAM and could operate two Agilent ADCM-1670 352×288 pixels at 15 fps sensor or up to four Agilent ADNS-3060 30×30 pixel image sensors at 100 fps. It has also a built-in IEEE 802.15.4 Chipcon CC2420 radio device. Target processing solutions includes the ability to learn from an environment and to control agents based on visual observations.

WiCa (Kleihorst, 2007) smart camera mote are based on a high performance single-instruction multiple-data processor. It contains a Xetal-II SIMD processor running at 84 MHz. The advantage of this solution is represented by the use of parallel processing to reduce the number of memory accesses, clock speed and instruction decoding. Despite the high computational capabilities, the power consumption is kept under 600 mW. The Xetal-II is coupled using a dual port RAM with a general purpose ATMEL8051 processor. The mote integrates one or two OM6802 camera sensors and a Texas Instruments CC2420 transceiver. Possible applications are Canny edge detection or real time gesture recognition.

The Citric platform (Chen, 2008) consists of an Intel XScale PXA270 processor running at up to 624 MHz clock speed, 32MB RAM, 32MB FLASH and an OmniVision OV9655 1.3 megapixel CMOS camera capable of SXGA 1280×1024 pixel resolution at 15 fps. The typical active power consumption of the camera is around 90mW at SXGA and the standby current is less than 20μA. The communication is ensured by a CC2420 radio component implementing IEEE 802.15.4 capable of 250 kbps transfers. This platform is powerful enough for medium intensive vision algorithms. It can ensure for image difference background subtraction and for bounding box computation a processing time per frame in the range of 0.2s – 0.4s at a resolution of 320×480.

CHALLENGES AND SOLUTIONS IN WSN MULTIMEDIA DATA EXCHANGE

Issues Regarding Data Exchange

Wireless Sensor Networks are used to sense the environment. Their applications require data to be collected in a central point in order to be processed and stored. Network nodes usually send these data when environment changes or when an event is detected. In case of VWSN systems, video frames captured by sensor camera represent the row data. A basic solution for data exchange implies in this case video streaming to the central point. However, real-time video streaming has stringent requirements for bandwidth, end-to-end delay and loss during transmission. These issues are hard to be solved by nodes wireless communication modules. The wireless communication is characterized by high path loss, channel fading, interference, noise disturbances, and high error rate. In general, wireless channels have much lesser streaming capacity than wired channels. But for many monitoring applications, transient faults during transmissions can be tolerated. In addition, some error recovery mechanism could to be adopted to reduce the impact of packet losses, providing some level of reliability (Qaisar, 2009). In case of higher requirements, various solutions were proposed in form of data compression or flow congestion control mechanisms.

Recent versions of the VWSNs have the ability to perform local processing computations and data aggregation. The aim is to send only the relevant part of the sensed data to the central point. This represents a valuable improvement over their predecessors as it has been shown that the network nodes typically spend most of their energy in transmitting data. Therefore, in-node processing, compression and data aggregation often results in a reduction in the overall energy consumption.

Proposed Solutions for Data Exchange

The ability of a WSN to provide support for video streaming is restricted due to the hardware, communication capabilities, and power limitations of the sensor nodes. Relatively few applications have been proposed for multimedia streaming in such systems.

Data Streaming

Solutions in VWSN data streaming are limited due communication hardware restrictions. Early platforms as WeC, René and Dot2000 have very low 10kbps bandwidth. More recently platforms as Crossbow MICA and MICA2 rely on 38.4 kbps ChipCon1000 wireless module. Today, most WSN motes adopt the IEEE 802.15.4 communication standard (Bougard, 2005), and a transmission rate around 250kbps. To obtain it, Crossbow MicaZ, Tmote Sky and TelosB motes use ChipCon2420 module, while Arduino Mega-2560 and Waspmote adopt XBee-802.15.4. However, 250 kbps are not suitable to transmit high data rate media streams. A better solution is Bluetooth radio with maximum data rate of 3Mbps, used by Intel iMote1 with the drawbacks of quite limited size of the network. More advanced platforms, as Panoptes (Feng, 2005) or Intel Stargate1 (Nachman, 2005), use instead more capable 802.11 networking. Even so the main problem related to IEEE 802.11 devices is the high power-consumption that makes them suitable mainly for wireless local area computer network implementation.

Based on existing platforms one important direction of research in VWSN video streaming covers improvements of video transmission mechanisms. The reference (Maimour, 2009) addresses the problem of congestion control for information-intensive flows in surveillance WSN applications. They propose a multi-path routing solution and efficient congestion detection in case of packet losses due buffer overflow and the contention of radio channel. It relays on a mix of several mechanisms with the aim of ensuring a better handling of video flows. To solve congestions they develop several load repartition strategies on top of the multipart support. Using those strategies the video flow is split on multiple paths based on these strategies. This help in maintaining the transmission rate unchanged to ensure the effectiveness of the surveillance application. To evaluate performance they calculate a *fairness* metric defined as

$$\frac{\left(\sum_{i=1}^{N_s} r_i\right)^2}{N_s \sum_{i=1}^{N_s} r_i^2}, \qquad (10)$$

where r_i represents the success rate achieved by the source i, and N_s is the total number of sources. The performance achieved in term of this metric for 250 nodes using an incremental approach to add new paths is around 80%. However, traffic distribution on multiple paths is not efficient in term of energy preservation.

The problem of loss recovery is treated in (Paek, 2007). They propose a Rate-Controlled Reliable Transport – RCRT protocol, which addresses emerging high-rate applications involving loss-intolerant multimedia data transfer. It uses end-to-end explicit loss recovery, but places all the congestion detection and rate adaptation functionality in the sinks. Sinks are able to achieve greater efficiency since they have a more comprehensive view of network behavior.

Other solutions use multiple paths for data transfer to alleviate the intensity of buffer usage at the intermediate sensor nodes and to reduce the required data rate on each wireless path. An example is COngestion Detection and Avoidance – CODA protocol described in (Wan, 2003). It allows a collection point to manage multiple sources associated with a single event in case of detecting network congestion. A drawback of the solution is the time delay in tacking action by the source.

Data Compression

One of the most commonly used communications standard for wireless sensor networks is the IEEE 802.15.4. Main advantages are low cost of equipments and power efficiency. This standard specifies a maximum data rate up to 250kB/s, which is relatively slow for video streaming. Hence, compression methods should be used to reduce the amount of data. They solve the problem of low data rate, but compression algorithms involved is high power demanding. Therefore the solution is most of the time comparable with the transmission cost of uncompressed image. Nevertheless some of them prove very reasonable power demands as for example JPEG compression using fixed-point discrete cosine.

A video clean sensor architecture based on IEEE 802.15.4 was proposed by (Shahidan, 2011). It consists in an Atmega644PV microcontroller unit operating at 3.3 V, an AT45DB321D data flash for data buffering, a C328-7640 VGA resolution camera module, and an XBee RF transceiver module for communication. Solution proves good performance in data compression and transfer for 640x480 video frames.

An alternative to classical video streaming is the image streams. In reference (Chiasserini, 2002) the solution is periodic transmission of compressed images. For efficiency they use JPEG with fixed-point discrete cosine transform for compression, in place of the commonly used floating-point transform. This approach provides good compression rates while computation complexity is not too high. The system runs on an Intel Strong-Arm 1110 platform at 59 MHz, the compression factor is around 8:1, and the maximum image transmission delay is 2s. However, the processed grayscale QCIF images are not suitable for complex applications.

Solution proposed in (Wu, 2004) adopts a multi-layer coding based on JPEG 2000 instead of JPEG change-difference coding. They use wavelet-based decomposition to create multiple bit-stream image encodings that are transmitted in small fragment bursts. The aim is to obtain an optimum balance between energy consumption for image coding and energy spent for wireless data transmission. Performance evaluation through simulations shows significant increasing of the system lifetime while satisfying application constraints related to image quality.

Routing Algorithms

Unlike routing in computers networks, in case of WSN several new issues have to be considered. The depletion of nodes battery power can result in broken links and affect the continuity of data transmission. Therefore, energy-aware routing is necessary to include policies for managing energy depletion. Low bandwidth, complex topology and harsh deployment environment should to be take into account by routing algorithms. Furthermore VWSN routing is expected to ensure also specific Quality of Service.

The reference (Wang, 2007) proposes an approach based on synchronized pipelined transmission for video data streaming. The route discovery process is based on a probabilistic method. The source node periodically sends out route probing packets. The probing packets are randomly relayed to a neighbor until they reach the central point. The subscriber node calculates the optimal path based on all received probing packets when a predefined timer expires. This algorithm is efficient if the source node is not significantly far from the central point. A certain level of energy conservation is achieved through reduction of packet retransmissions in the presence of node failures.

A location-based routing for video streaming is presented in (Cosma, 2006). The authors propose a topology extraction protocol using networks video cameras. There are two steps to accomplish the topology extraction. First the central point diffuse routing messages over the network and every node records routing information. After a path set-up phase, every node in the network captures an image using its video camera, and sends the image back to the central point. This node then performs image registration to extract the topology of the network. The authors further discuss possible optimizations for path routing and energy conservation. To accomplish that, every node maintains a record of their neighbors' energy level and hop count to the server. Any node with a relatively high energy (e.g. >20%) should be included in the candidate set for next hop selection during routing. The winner candidate will be the one with the smallest hop count.

The two-phase geographical greedy forwarding routing protocol (Shu et al. 2008) uses a greedy location-based scheme. In relays on a *step back and mark* process to explore possible paths to the sink. The aim is to find a route to the destination as if one exists. The method puts routing paths as close as possible to the centerline, and can cause very severe path coupling issues.

Other approaches are based on hierarchical schemes. Reference (Politis, 2008) describes a hierarchical solution for video data. The network architecture setup is derived from the architecture of Low-Energy Adaptive Clustering Hierarchy – LEACH – proposed by (Heinzelman, 2000). Instead of using a direct link between a cluster head and sink, cluster heads are permitted to establish links to each other. A video sensor node can select a number of available paths through other cluster heads in order to transmit its data. This improvement decreases the transmission power of a cluster head for shorter-range communication and saves energy.

Some interesting solutions are focused on energy saving. A possibility is to send data through the path of fewest hops and most longevity. This approach was proposed by (Li, 2007). The resulting global-energy-balancing routing – GEBR – scheme for real-time traffic is based on directed diffusion and balances node energy utilization to increase the network lifetime. The longevity of a path is measured with a metric called Minimum-Path-Energy, which express the minimum energy of all the nodes along a certain path. The path with fewer hops is considered to ensure the lowest data transmission delay.

CONCLUSION AND OPEN RESEARCH TOPICS

Advances in embedded systems, low power CMOS video sensors, and wireless communication have led to Video Wireless Sensor Networks. They add value to a large variety of application domains like civil or military surveillance, traffic management or environment monitoring. Challenges in designing VWSN applications are manly related with resource limitations. Video sensing is expensive in terms of energy consumption. In-node data processing as compression or feature extraction should be implemented on limited computational power and small memory buffers. Wireless data transmission over multiple hops is prone to errors and congestion, and overall distributed resource management is complex.

Among several solutions presented in this chapter we identify also some remaining open research issues concerned with various aspects of VWSN design and development.

As both video capturing and processing remains high power demanding, a main issue is related with node hardware design. A good perspective is offered here by development of specialized low power DSP chips or more complex controllers with vector processing capabilities.

More research should be done in embedded image and video processing to improve localization techniques, tracking and feature extraction. Good perspectives offer also 3D processing techniques and specialized stereovision devices.

Energy conservation is a critical design challenge for VWSN routing. An open issue here is how to ensure accurate network energy status measurement and to define accurate consumption models. Research should address improved protocols, distributed resource management and new low power/long range communication devices. High accurate simulation models and tools should be developed.

REFERENCES

Adriaens, J., Megerian, S., & Potkonjak, M. (2006). Optimal worst-case coverage of directional field-of view sensor networks. In *Proceedings of IEEE Sensor, Mesh and Ad Hoc Communications and Networks* (pp. 336 – 345). Reston, VA: IEEE.

Borgefors, G. (1984). An improved version of the chamfer matching algorithm. In *Proceedings of the 7th International Conference on Pattern Recognition*. Montreal, Canada: Academic Press.

Bougard, B., Catthoor, F., Daly, D. C., Chandrakasan, A., & Dehaene, W. (2005). Energy efficiency of the IEEE 802.15.4 standard in dense wireless microsensor networks: Modeling and improvement perspectives. In Proceedings of Design, Automation and Test in Europe – DATE'05 (pp. 196-201). IEEE Computer Society.

Bulusu, N., Heidemann, J., & Estrin, D. (2000). GPS-less low cost outdoor localization for very small devices. *Proceedings of the IEEE Personal Communications*, *7*(5), 28–34. doi:10.1109/98.878533

Canny, J. F. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *8*(6), 679–698. doi:10.1109/TPAMI.1986.4767851 PMID:21869365

Chen, P., Ahammad, P., Boyer, C., Huang, S., Lin, L., & Lobaton, E. … Sastry, S. (2008). Citric: A low-bandwidth wireless camera network platform. In *Proceedings of the Second ACM/IEEE International Conference on Distributed Smart Cameras ICDSC 2008*. Stanford, CA: ACM/IEEE.

Chiasserini, C., & Magli, E. (2002). Energy consumption and image quality in wireless video-surveillance networks. In *Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications – PIMRC'02* (pp. 2357–2361). IEEE.

Chiasserini, C. F., & Magli, E. (2002). Energy consumption and image quality in wireless video-surveillance networks. In *Proceedings of 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC'02* (pp. 2357–2361). IEEE Computer Society.

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(5), 603–619. doi:10.1109/34.1000236

Cosma, M., Pescaru, D., Ciubotaru, B., & Todinca, D. (2006). Routing and topology extraction protocol for a wireless sensor network using video information. In *Proceedings of the 3rd Romanian-Hungarian Joint Symposium on Applied Computational Intelligence SACI'06*. Timisoara, Romania: SACI.

Costa, D., & Guedes, L. (2010). The coverage problem in video-based wireless sensor networks: A survey. *Sensors (Basel, Switzerland)*, *10*, 8215–8247. doi:10.3390/s100908215 PMID:22163651

Costa, D., & Guedes, L. (2013). Exploiting the sensing relevancies of source nodes for optimizations in visual sensor networks. *Multimedia Tools and Applications*, *64*(3), 549–579. doi:10.1007/s11042-011-0961-4

Curiac, D., Volosencu, C., Pescaru, D., Jurca, L., & Doboli, A. (2009). View upon redundancy in wireless sensor networks. In *Proceedings of the 8th Wseas International Conference on Signal Processing, Robotics and Automation – ISPRA'09, Mathematics and Computers in Science and Engineering* (pp. 341-346). Cambridge, UK: ISPRA.

Dawood, M. S., Suganya, J., Devi, R. K., & Athisha, G. (2013). A review on wireless sensor network protocol for disaster management. *International Journal of Computer Applications Technology and Research*, *2*(2), 141–146. doi:10.7753/IJCATR0202.1011

Desai, P., & Rattan, K. S. (2009). Indoor localization and surveillance using wireless sensor network and pan/tilt camera. In *Proceedings of the IEEE National Aerospace & Electronics Conference NAECON'09*. Fairborn, OH: IEEE.

Downes, I., Rad, L. B., & Aghajan, H. (2006). Development of a mote for wireless image sensor networks. In *Proceedings of COGnitive Systems with Interactive Sensors – COGIS'06*. Paris, France: COGIS.

Elnahrawy, E., Li, X., & Martin, R. P. (2004). The limits of localization using signal strength: A comparative study. In *Proceedings of the 1st IEEE International Conference on Sensor and Ad Hoc Communications and Networks SECON'04*. Santa Clara, CA: IEEE.

Feng, W. C., Kaiser, E., Feng, W. C., & Le Baillif, M. (2005). Panoptes: Scalable low-power video sensor networking technologies. *ACM Transactions on Multimedia Computing*. *Communications and Applications*, *1*(2), 151–167.

Fernandez, J., Calavia, L., Baladrón, C., Aguiar, J. M., Carro, B., & Sánchez-Esguevillas, A. et al. (2013). An intelligent surveillance platform for large metropolitan areas with dense sensor deployment. *Sensors (Basel, Switzerland)*, *13*(6), 7414–7442. doi:10.3390/s130607414 PMID:23748169

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communication Magazine of the ACM*, *24*(6), 381–395. doi:10.1145/358669.358692

Fuiorea, D., Gui, V., Pescaru, D., Paraschiv, P., Istin, C., Curiac, D., & Volosencu, C. (2008). Sensor node localization using SIFT algorithm. In *Proceedings of the 9th International Conference on Automation and Information* (pp. 436-442). Bucharest, Romania: Academic Press.

Heinzelman, W., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. IEEE.

Huang, C., & Tseng, Y. (2003). The coverage problem in a wireless sensor network. In *Proceedings of 2nd 909 ACM International Workshop on Wireless Sensor Networks and Applications* (pp. 115-121). San Diego, CA: ACM.

Istin, C., & Pescaru, D. (2007). Deployments metrics for video-based wireless sensor networks. *Transactions on Automatic Control and Computer Science, 52*(66).

Istin, C., Pescaru, D., & Doboli, A. (2011). Stochastic model-based heuristics for fast field of view loss recovery in urban traffic management through networks of video cameras. *IEEE Transactions on Intelligent Traffic Systems*, *12*(3), 895–907. doi:10.1109/TITS.2011.2123095

Kleihorst, R., Abbo, A., Schueler, B., & Danilin, A. (2007). Camera mote with a high-performance parallel processor for real-time frame-based video processing. In *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance* – AVSS'07 (pp. 69-74). IEEE.

Kusy, B., Ledeczi, A., & Koutsoukos, X. (2007). Tracking mobile nodes using RF doppler shifts. In *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems SenSys'07*. Sydney, Australia: ACM Press.

Langendoen, K., & Reijers, N. (2005). Distributed localization algorithm. In *Embedded systems handbook*. Boca Raton, FL: CRC Press.

Li, P., Gu, Y., & Zhao, B. (2007). A global-energy-balancing real-time routing in wireless sensor networks. In *Proceedings of the 2nd IEEE Asia-Pacific Service Computing Conference* (pp. 89 - 93). Tsukuba, Japan: IEEE.

Liu, L., Ma, H., & Zhang, X. (2008). On directional k-coverage analysis of randomly deployed camera sensor networks. In *Proceedings of IEEE International Conference on Communications* (pp. 2707-2711). Beijing, China: IEEE.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proceedings of the 7th IEEE International Conference on Computer Vision*, (Vol. 2, pp. 1150–1157). IEEE.

Maimour, M., Pham, C., & Hoang, D. (2009). A congestion control framework for handling video surveillance traffics on WSN. In *Proceedings of the 2009 International Conference on Computational Science and Engineering (CSE '09)*, (Vol. 2, pp. 943-948). Washington, DC: IEEE Computer Society.

Meghani, S. K., Asif, M., & Amir, S. (2012). Localization of WSN node based on time of arrival using ultra wide band spectrum. In *Proceedings of the 13th IEEE Annual Wireless and Microwave Technology Conference WAMICON'12*. IEEE.

Mohammad, R., Rick, B., Obimdinachi, I. I., Juan, C. G., Jay, W., Deborah, E., & Mani, S. (2005). Cyclops: In situ image sensing and interpretation in wireless sensor networks. In *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems SenSys '05* (pp. 192-204). ACM.

Nachman, L., Kling, R., Adler, R., Huang, J., & Hummel, V. (2005). The Intel® mote platform: A bluetooth-based sensor network for industrial monitoring. In *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks – IPSN '05*. Piscataway, NJ: IEEE Press.

Osais, Y. E., St-Hilaire, M., & Fei, R. Y. (2010). Directional sensor placement with optimal sensing ranging, field of view and orientation. *Mobile Networks and Applications*, *15*(2), 216–225. doi:10.1007/s11036-009-0179-0

Paek, J., & Govindan, R. (2007). RCRT: Rate-controlled reliable transport for wireless sensor networks. In *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems – SenSys'07* (pp. 305-319). New York, NY: ACM.

Park, C., & Chou, P. H. (2006). eCAM: Ultra compact, high data-rate wireless sensor node with a miniature camera. In *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems SenSys '06* (pp. 359-360). New York, NY: ACM.

Pescaru, D., Ciubotaru, B., Chiciudean, D., & Doboli, A. (2005). Experimenting motion detection algorithms for sensor network video surveillance applications. *Transactions on Automatic Control and Computer Science*, *50*(64), 39–44.

Pescaru, D., Fuiorea, D., Gui, V., Toma, C., Muntean, G. M., & Doboli, A. (2006). Image-based node localization algorithm for wireless video sensor networks. In *Proceedings of the IT&T Conference*. Carlow, Ireland: Institute of Technology.

Pescaru, D., Gui, V., Toma, C., & Fuiorea, D. (2007). Analyses of post-deployment sensing coverage for video wireless sensor networks. In *Proceedings of the 6th International Conference RoEduNet-2007* (pp. 109-113). SITECH.

Pescaru, D., Istin, C., Curiac, D., & Doboli, A. (2008). Energy saving strategy for video-based sensor networks under field coverage preservation. In *Proceedings of the IEEE-TTTC International Conference on Automation, Quality and Testing, Robotics – AQTR'08* (pp. 289-294). IEEE.

Politis, I., Tsagkaropoulos, M., Dagiuklas, T., & Kotsopoulos, S. (2008). Power efficient video multipath transmission over wireless multimedia sensor networks. *Mobile Networks and Applications*, *13*(3-4), 274–284.

Qaisar, S., & Radha, H. (2009). A reliability framework for visual sensor networks. In *Proceedings of Picture Coding Symposium*. Chicago: Academic Press.

Rahimi, M., Baer, R., Iroezi, O., Garcia, J., Warrior, J., Estrin, D., & Srivastava, M. (2005). Cyclops: In-situ image sensing and interpretation in wireless sensor networks. In *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems – SenSys'05* (pp. 192-204). New York, NY: ACM.

Sanchez, J., Benet, G., & Simó, J. E. (2012). Video sensor architecture for surveillance applications. *Sensors (Basel, Switzerland)*, *12*, 1509–1528. doi:10.3390/s120201509 PMID:22438723

Savvides, A., Han, C. C., & Srivastava, M. B. (2001). Dynamic fine-grained localization in ad-hoc networks of sensors. In *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom'01),* (pp. 166-179). ACM Press.

Shahidan, A. A., Fisal, N., Fikri, A. H., Ismail Nor-Syahidatul, N., & Yunus, F. (2011). Image transfer in wireless sensor networks. In *Proceedings of the International Conference on Communication Engineering and Networks IPCSIT'2011*, (Vol. 19, pp. 158-165). Singapore: IACSIT Press.

Shu, L., Zhang, Y., Zhou, Z., Hauswirth, M., Yu, Z., & Hyns, G. (2008). Transmitting and gathering streaming data in wireless multimedia sensor networks within expected network lifetime. *Mobile Networks and Applications*, *13*(3-4), 306–323.

Wan, C. Y., Eisenman, S. B., & Campbell, A. T. (2003). CODA: Congestion detection and avoidance in sensor networks. In *Proceedings of the ACM Conference on Embedded Networked Sensor Systems –SenSys*. Los Angeles, CA: ACM Press.

Wang, J., Masilela, M., & Liu, J. (2007). Supporting video data in wireless sensor networks. In *Proceedings of the 9th IEEE International Symposium on Multimedia* (pp. 310-317). Los Alamitos, CA: IEEE.

Wu, H., & Abouzeid, A. (2004). Energy efficient distributed JPEG 2000 image compression in multi-hop wireless networks. In *Proceedings of the 4th Workshop on Applications and Services in Wireless Networks – ASWN'04* (pp. 152–160). ASWN.

Zeinalipour, Y. D., Neema, S., Kalogeraki, V., Gunopulos, D., & Najjar, W. (2005). Data acquisition in sensor networks with large memories. In *Proceedings of the 21st International Conference on Data Engineering Workshops ICDEW '05*. IEEE Computer Society.

ADDITIONAL READING

Akyildiz, I. F., Melodia, T., & Chowdhury, K. R. (2007). A Survey on Wireless Multimedia Sensor Networks. [Elsevier USA.]. *Computer Networks*, *51*(4), 921–960. doi:10.1016/j.comnet.2006.10.002

Horster, E., Lienhart, R., Kellermann, W., & Bouguet, J. Y. (2005). *Calibrating Visual Sensors and Actuators in Distributed Platforms. Computer Vision for Interactive and Intelligent Environment* (pp. 39–50). Lexington, Kentucky, USA: IEEE Computer Society. doi:10.1109/CVIIE.2005.2

Molina, J., Mora-Merchan, J., Barbancho, J., & Leon, C. (2010). Multimedia Data Processing and Delivery in Wireless Sensor Networks, Wireless Sensor Networks: Application - Centric Design, Yen Kheng Tan (Ed.), InTech.

Radke, R. (2010). A Survey of Distributed Computer Vision Algorithms. In H. Nakashima, H. Aghajan, & J. C. Augusto (Eds.), *Handbook of Ambient Intelligence and Smart Environments* (pp. 35–55). Springer, US. doi:10.1007/978-0-387-93808-0_2

Tarique, M., Tepe, K. E., Adibi, S., & Erfani, S. (2009). Survey of Multipath Routing Protocols for Mobile Ad-hoc Networks. *Journal of Network and Computer Applications*, *32*(6), 1125–1143. doi:10.1016/j.jnca.2009.07.002

Vieira, M. A. M., Coelho, C. N., Da Silva, D. C., & Da Mata, J. M. (2003). Survey on Wireless Sensor Network Devices. *In Proceedings of the IEEE Conference in Emerging Technologies and Factory Automation – ETFA '03*, vol.1 (pp. 537-544). IEEE.

KEY TERMS AND DEFINITIONS

CP: Central point, base station or sink is a special node of a WSN that have access to external energy source, and incorporates significant memory resources and processing power. Its role is to collect, process, and sometimes store data gathered from network sensors. It provides user interface to the WSN and/or transfers date to external systems and applications.

FoV: Field of View represents the entire angular expanse visible through a video camera objective at a given moment of time. The field of view is determined by the focal length of the lens and the size of the image sensor. Obstacles, illumination and weather conditions can affect it significantly.

GPS: Global Positioning System is a space-based satellite navigation system that provides location and time information. It can be used for WSN nodes localization under certain conditions.

Mote: A mote is a low power node in a wireless sensor network that is capable of reading sensory information, performing processing, and communicating with other nodes using wireless connection.

QoS: Quality of Service is a measure of overall performance of a network system. Quality of service is particularly important for the transport of traffic with special requirements as video streams in case of VWSNs.

VWSN: Video Wireless Sensor Networks is a WSN that incorporates video sensors and gather and transmit video information over wireless multi-hop connection.

WSN: Wireless Sensor Networks consists of distributed sensor nodes over a deployment area. The applications are related to monitoring physical events and conditions. They involve a large variety of sensors such as temperature, sound, pressure, light etc.