

Adaptive Compressive Sensing for Target Tracking within Wireless Visual Sensor Networks-based Surveillance applications

Salema Fayed · Sherin M.Youssef ·
Amr El-Helw · Mohammad Patwary ·
Mansour Moniri

Received: date / Accepted: date

Abstract Wireless Visual Sensor Networks (WVSNs) have gained significant importance in the last few years and have emerged in several distinctive applications. The main aim is to design low power WVSN surveillance application using adaptive Compressive Sensing (CS) which is expected to overcome the WVSN resource constraints such as memory limitation, communication bandwidth and battery constraints. In this paper, an adaptive block CS technique is proposed and implemented to represent the high volume of captured images in a way for energy efficient wireless transmission and minimum storage. Furthermore, to achieve energy-efficient target detection and tracking with high detection reliability and robust tracking, to maximize the lifetime of sensor nodes as they can be left for months without any human interactions. Adaptive CS is expected to dynamically achieve higher compression rates depending on the sparsity nature of different datasets, while only compressing relative blocks in the image that contain the target to be tracked instead of compressing the whole image. Hence, saving power and increasing compression rates. Least mean square adaptive filter is used to predicts target's next location to investigate the effect of CS on the tracking performance. The tracking is achieved in both indoor and outdoor environments for single/multi targets. Results have shown that with adaptive block CS up to 20% measurements of data are required to be transmitted while preserving the required performance for target detection and tracking.

Keywords Adaptive Compressive Sensing · Compressive sensing · LMS · Surveillance applications · Target tracking · WVSN

S. Fayed^{a*} · S. M.Youssef^a · A. El-Helw^b

^a Computer Engineering Department, ^b Electronics and Communication Department,
College of Engineering and Technology, AAST, Alexandria, Egypt

* E-mail: fw010203@student.staffs.ac.uk

M.Patwary^c · M. Moniri^c

^c Faculty of Computing, Engineering and Technology
Staffordshire University, Stoke on Trent, UK

1 Introduction

Wireless Visual Sensor Networks (WVSNs) have gained significant importance in the last few years and have emerged in several distinctive applications where sensor nodes are deployed with cameras to capture and transmit visual data [1–4]. Due to the evolvement of new technologies and techniques, there are immediate needs for automated energy-efficient surveillance systems. WVSN has targeted various surveillance applications in commercial, law enforcement and military purpose as well as traffic control, security in shopping malls and amusement parks. Systems have been developed for video surveillance including highway, subway and tunnel monitoring, in addition to remote surveillance of human activities such as elderly or patients care. Visual sensor nodes are resource constraint devices bringing the special characteristics of WVSNs such as energy, storage and bandwidth constraints which introduced new challenges [2, 5]. In WVSN large data sets such as video, and still images are to be retrieved from the environment. Since the size of visual data is much larger than scalar data, visual data processing in WVSNs is more complex and it requires more hardware resources such as CPU power and memory buffer. Moreover, the transmission of visual data requires more network bandwidth resources, which are all quite costly in terms of energy consumption. Hence, in order not to consume so many hardware resources, one possibility is to reduce the size of the visual data [1]. Furthermore, wireless channels in surveillance applications are subject to noisy conditions; therefore, detection and tracking reliability within such resource constrained condition be the main challenge when designing WVSN surveillance applications. energy efficient processing and efficient compression techniques are the strongest candidates to overcome such constraints while transmitting data for WVSN applications and hence minimize energy expenditure [4, 6].

Much work is present in the literature for surveillance applications within WVSNs. In [7], the authors introduced a multi-view visual-target-surveillance system in WVSN, which implements target classification and tracking with collaborative online learning and localization. In [8], a practical target tracking WSN system is proposed based on the auto regressive moving average (ARMA) model in a distributed peer-to-peer (P2P) signal processing framework. Wireless sensor nodes act as peers that perform target detection, feature extraction, classification and tracking, whereas target localization requires the collaboration between wireless sensor nodes for improving the accuracy and robustness. A distributed multi-view tracking system using collaborative signal processing in distributed wireless sensor networks is proposed in [9]. In the tracking system, target detection and classification algorithms are based on single-node processing and target tracking is performed in sink node, whereas target localization algorithm is carried out by collaboration between multi-sensors. A progressive distributed data fusion is proposed to overcome the disadvantages of client/server based centralized data fusion. In [10], Sherin M.Youssef et al proposed a model integrating the human visual characteristics of video motion, in the frequency multi-resolution wavelet domain, with multi-

dimensional fuzzy inference perceptual model. However, due to the integration of fuzzy inference, the model is slow and not suitable for real time WVSNs application.

There is significant literature for target tracking surveillance applications in WVSNs. Template correlation matching as proposed in [11] is a method for tracking where a template is taken from previous frames and correlated with regions in next frame to find which region best accommodates the template. However, template correlation is not robust in the presence of changes in the targets appearance. Kalman filtering [12, 13] is relatively the best linear estimator for target tracking. Kalman filters are robust under optimal conditions, otherwise adaptive approaches are needed to solve these problems which can be either computational expensive or not always be applicable in real time tracking.

Particle filtering which is known to be suitable for real time tracking and non-linear non-Gaussian processes, it relies on motion parameter estimation and probability estimates [14]. Subsequently, the performance of the particle filter in terms of tracking reliability decreases with noisy or low resolution frames and with false positive detection of target.

Adaptive filters such as least mean square (LMS) algorithm is relatively simple, has much lower computational complexity than the original Kalman filters and other adaptive algorithms; LMS works satisfactorily in the absence of apriori information, in contrast to Kalman filtering. They have been widely used in several applications such as motion estimation, signal prediction and tracking time variations in signals and vision applications. Moreover, it is suitable for real time image applications [15, 16].

Based on the above literature, to attain a trade off between computational complexity and detection and tracking accuracy in the context of energy constrained WVSN, an image processing scheme is required with optimal pre-processing and post-processing can provide intended target detection and tracking accuracy within energy constraint nature of WVSN. Moreover, as for high volume data sets are acquired with WVSN surveillance applications, should be represented in such a way that it requires optimum storage, energy, and allow reliable transmission due to the constraint on the physical and radio resources. Suppose for a surveillance application within WVSN, an image is captured and required to be sampled for storage as well as to be transmitted through wireless channel. According to Shannon-Nyquist sampling theory the minimum number of samples required to accurately reconstruct the signal without losses is twice its maximum frequency [17]. It is always challenging to reduce this sampling rate as much as possible, hence reducing the computation energy and storage. Within the scope of the authors knowledge, recently proposed CS [17] is expected to be a strong candidate to provide this and overcome the above mentioned limitations where CS has been considered for different aspects of surveillance applications due to its energy efficient and low power processing as reported in [18, 19]

CS is a useful imaging tool under various noise conditions when the underlying image is compressible in a known basis or representation. It has been

widely used in several applications such as image processing, steganography and image watermarking [20, 21]. Moreover, CS offers an advantage and shows a better performance than conventional sampling and compression techniques such as JPEG, JPEG2000 and DCT, where it achieved higher PSNR and lower reconstruction error [20–23].

In [18], compressive sensing for background subtraction and multi-view ground plane target tracking are proposed. A convex optimization known as basis pursuit or orthogonal matching pursuit is exploited to recover only the target in the difference image using the compressive measurements to eliminate the requirement of any auxiliary image reconstruction. Other work in compressive sensing for surveillance applications has been proposed in [24], where an image is projected on a set of random sensing basis yielding some measurements where at the receiver end the image is reconstructed by minimizing the weighted version of the L2 norm. However, further research is required to address the selection of the weights and fully understand their impact on the reconstruction problem while taking into account the energy-efficiency parameter.

The H.264 and MPEG4 video coding are the most recent coding standard of video stream. However, they are based on complex encoders and simple decoders as the encoder performs intra-frame coding and exploits statistical dependence between frames in the source video signal to perform inter-frame coding. This configuration is suitable for many applications such as video broadcasting but for WVSNs it is different due to the limited energy and computational capabilities [25]. Moreover, although H.264 achieves high coding performance but at the expense of huge computational complexity as predictive encoding such as H.264 and MPEG4 requires complex processing algorithms, which lead to high energy consumption [26]. In contrast to CS which is designed to aim simple encoders with very low computational complexity whereas complex computations are left at the decoder side which is not battery-powered. A comparative study between CS and traditional coding techniques is carried out in [25, 26], CS measurements can tolerate a fairly large amount of bit errors before the received video quality is affected. This is certainly not the case for predictive video encoding, and not even for transform-based image compression standards such as JPEG. This could result in significant transmission power savings or a significant decrease in the amount of forward error correction. Furthermore, the processor load is significantly lower for CS than for H.264. This results in a reduction in the energy needed to encode the video. In addition, for H.264 and MPEG4, PSNR is affected by the frame type, (i.e. I, P, or B frame) due to inter-frame coding. However, for CS, size of measurement matrix satisfying the $M \geq K \log N$ lower bound guarantees more satisfactory quality.

The WVSNs have strict resource limitations hence make it desirable to design energy-efficient target tracking technique. In this paper, the impact of low-power adaptive CS is investigated in designing robust and reliable target detection and tracking techniques for WVSNs-based surveillance applications, without compromising the energy constraint which is one of the main char-

acteristics of WVSNs. Adaptive CS is compared to conventional CS in terms of energy consumed and size of data transmitted. Adaptive CS is expected to reduce the size of sampled data by adaptively choosing compression rates according to the sparsity nature of different scenes with low complexity processing due to its low power simple process [27]. Hence, an adaptive block compressive sensing-based single/multi target tracking is proposed which is expected to reduce energy consumption, space requirements and channel bandwidth, with acceptable tracking reliability which will be represented as minimal Mean Square Error (MSE).

The rest of the paper is organized as follows, Introduction to CS is presented in Section 2. Section 3 presents the proposed system and energy models. The proposed technique for CS-based target detection and tracking is given in Section 3.2. Simulations and results are provided in Section 4 and finally the conclusion in Section 5.

2 Compressive Sensing

CS theory shows that a signal can be reconstructed from far fewer samples than required by Nyquist theory as it is always challenging to reduce the sampling rate as possible, provided that the signal is sparse (where most of the signal's energy is concentrated in few non-zero coefficients) or compressible in some basis domain [27]. CS is a simple and low energy consumption process which is suitable for power constraint sensor nodes where complex computations are just done at the Base station (BS).

CS exploits the sparsity nature of frames, so it compresses the image using far fewer measurements. Although, it is not necessary for the signal itself to be sparse but compressible or sparse in some known transform domain Ψ according to the nature of the image (i.e., the original image has approximate sparse representations in some linear transformations), smooth signals are sparse in the Fourier basis, and piecewise smooth signals are sparse in a wavelet basis [27–30].

Suppose image \mathbf{X} of size $(N \times N)$ is K -sparse that either sparse by nature or sparse in Ψ domain, Ψ is the basis invertible Orthonormal function of size $(N \times N)$ driven from a transform such as the DCT, fourier, or wavelet, where $K \ll N$, that is, only K coefficients of \mathbf{x} are nonzero and the remaining are zero, thus the K -sparse image \mathbf{X} is compressible. CS then guarantees acceptable reconstruction and recovery of the image from lower measurements compared to those required by shannon-Nyquist theory as long as the number of measurements satisfies a lower bound depending on how sparse the image is. Hence, \mathbf{X} can be recovered from measurements of size M where $M \geq K \log N \ll N$. Eq.(1) shows the mathematical representation of \mathbf{X}

$$\mathbf{X} = \Psi \mathbf{S} \quad (1)$$

Where, \mathbf{S} is a matrix containing the sparse coefficients of \mathbf{X} of size $(N \times N)$, $s_i = \langle \mathbf{X}, \psi_i^T \rangle = \psi_i^T \mathbf{X}$, $\mathbf{S} = \mathbf{\Psi}^T \mathbf{X}$. The image is represented with fewer samples from \mathbf{X} instead of all pixels by computing the inner product between \mathbf{X} and $\mathbf{\Phi}$, namely through incoherent measurements \mathbf{Y} in Eq.(2), where $\mathbf{\Phi}$ is a random measurement matrix of size $(M \times N)$ where $K \ll M \ll N$. $\mathbf{y}_1 = \langle \mathbf{x}, \phi_1 \rangle$, $\mathbf{y}_2 = \langle \mathbf{x}, \phi_2 \rangle, \dots, \mathbf{y}_m = \langle \mathbf{x}, \phi_m \rangle$.

$$\mathbf{Y} = \mathbf{\Phi} \mathbf{X} = \mathbf{\Phi} \mathbf{\Psi} \mathbf{S} = \mathbf{\Theta} \mathbf{S} \quad (2)$$

Since $M < N$, recovery of the image \mathbf{X} from the measurements \mathbf{Y} is undetermined, However, if \mathbf{S} is K -sparse, and $M \geq K \log N$ it has been shown in [27] that \mathbf{X} can be reconstructed by ℓ_1 norm minimization with high probability through the use of special convex optimization techniques without having any knowledge about the number of nonzero coefficients of \mathbf{X} , their locations, neither their amplitudes which are assumed to be completely unknown a priori [29–31]

$$\min \|\hat{\mathbf{X}}\|_{\ell_1} \text{ subject to } \mathbf{\Phi} \hat{\mathbf{X}} = \mathbf{Y} \quad (3)$$

Convex optimization problem can be reduced to linear programming known as Orthogonal Matching Pursuit (OMP) which was proposed in [32] to handle the signal recovery problem. It is an attractive alternative to Basis Pursuit (BP) [33] for signal recovery problems. The major advantages of this algorithm are its speed and its ease of implementation. As seen, the CS is a very simple process as it enables simple computations at the encoder side (sensor nodes) and all the complex computations for recovery of frames are left at the decoder side or BS.

3 System model

This work proposes an adaptive block compressive sensing model which is expected to reduce energy consumption, space requirements and communication overhead. Consider for a surveillance application a WWSN model composed of N visual sensor nodes and one or more base station (BS). Each sensor node i is required to capture images from a video sequence and detect the presence of objects. Each sensor node is assumed to be in 'wake-up' state according to a duty cycle $\beta_s \in [0, 1]$ over a period t_s , hence each sensor is awake for an interval of length $\beta_s t_s$ and sleep for an interval $(1 - \beta_s) t_s$. At the time where a sensor node enters a 'wake-up' state, the time reference for the frame count is assumed to be $t = 0$. Hence, a single snapshot at $t = 0$ is expected to be stored within the memory allocated at the sensor node; that is assumed to be the background for the intended target tracking; denoted as X_b . The following frames are the subsequent captured frames X_t with $t > 0$. Hence, X_b and X_t are the background and test images respectively of size $(N \times N)$ each. Let us assume most features of the targets are known to the monitoring center. However, the existence and the location of targets are required for monitoring. The receiver or BS also has prior explicit information of the background.

To achieve higher compression rates, the foreground target is extracted first by background subtraction resulting in the difference frame. Hence, assuring sparsity as the difference frame is always sparse regardless the sparsity nature of original image frames. Within the image frame, The extraction of foreground target X_d is achieved at each sensor node where CS is then applied for transmission through the wireless channel. At the BS side, the receiver decompresses the received compressed data obtaining \hat{X}_t to predicts the intended target's next location for tracking. The system model for the proposed WVSN is shown in Fig. 1, rest of the phases involved are described later in subsequent sections.

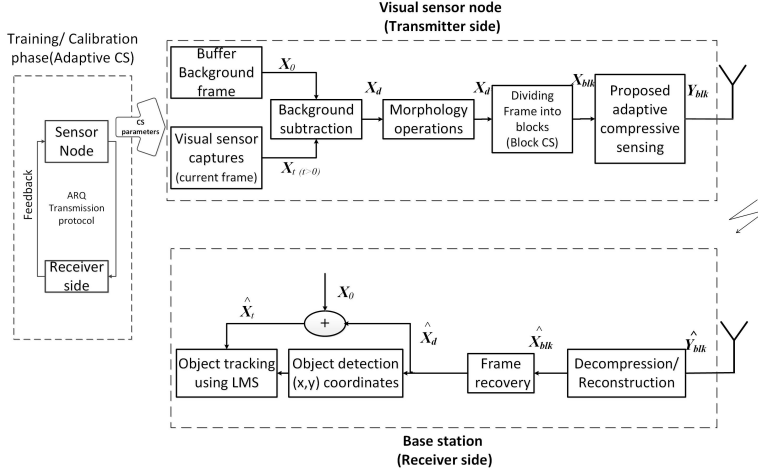


Fig. 1 The proposed model for WVSN-based surveillance application

3.1 Energy model

Currently, there is a great research in the area of low-energy radios. In our work, the same energy model as in [34] is used, where energy cost dissipated by a node over a distance d is denoted by E_{tx} as shown in (4).

$$E_{tx} = E_{elec} * k + e_{amp} * k * d^2 \quad (4)$$

Where, k is size of data (samples) transmitted, $E_{elec} = 50nJ/bit$, is the energy being used to run transmitter and receiver circuit, $e_{amp} = 100pJ/bit$ for the transmitted amplifier.

In WVSNs, most energy dissipated is during the transmission and reception, in our case the reception is the base station node which is assumed not to be battery-powered. Hence minimizing transmission energy can have more impact in energy saving [35, 36] where the energy consumed for processing is

very low as compared with the transmission energy. The energy needed to transmit 1 KB over a 100m distance is approximately equivalent to the energy necessary to carry out 3 million instructions [37–39].

3.2 Proposed detection technique

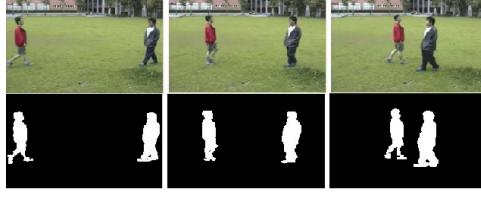
At each sensor node, after each image frame is being captured, some preprocessing might be required. In our case, to assure sparsity within the image frame, the foreground target is extracted first by background subtraction. Hence, instead of producing the compressed measurements for X_b and X_t separately, the compressed measurements are produced directly for X_d , as the difference frame is always sparse regardless of the sparsity nature of original images. Assuming the visual sensor node has captured an image denoted as \mathbf{X}_t . The target afterwards is detected based on thresholding the absolute difference between current frame \mathbf{X}_t and background frame \mathbf{X}_b , $\mathbf{X}_d = |\mathbf{X}_t - \mathbf{X}_b| > \gamma$, where γ is a given threshold to extract the foreground target as in Eq.(5),

$$X_d(i) = \begin{cases} |X_t(i) - X_b(i)| & \text{(foreground pixel) } |X_t(i) - X_b(i)| > \gamma \\ 0 & \text{(background pixel) } \textit{otherwise} \end{cases} \quad (5)$$

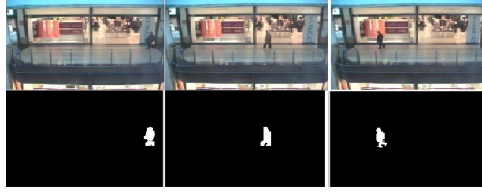
The value for γ is chosen in order to reduce scattered noise that could exist in the background due to many factors such as rain, dust, illumination changes, trees movements, etc. Its value is determined as in [40], where the image is divided into blocks (as will be illustrated in Sec.3.3.2 for block CS) and each block should be identified by a unique value that allows to properly choose the threshold. Statistic functions are commonly used such as mean, median, mean of minimum and maximum values of the local intensity distribution as they largely depends on the input images. In this paper, mean value is used as threshold to describe each image block. The threshold value helps reduce unwanted background subtraction noise and at the same time without causing disconnected targets as possible. Where, lower values of γ will result in more noise and higher values will result in disconnected targets. In both cases, resulting in either lower probability to detect the target or more preprocessing to overcome this problem. Once the foreground is detected, morphology operations [41] such as erosion and dilation operations are then applied for noise removal and blob formation respectively. Fig.2 shows the blob formation after background subtraction and morphological operations.

3.3 Proposed adaptive block Compressive sensing

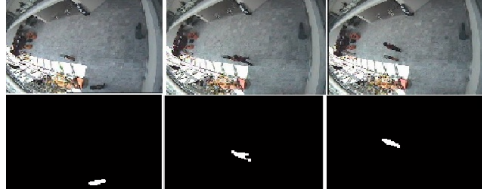
After the foreground blob \mathbf{X}_d is being extracted from the difference image, the proposed adaptive block CS is then applied. The proposed adaptive and block techniques are illustrated in subsequent sections.



(a) Walking men



(b) Shopping center 1



(c) Shopping center 2

Fig. 2 First row in (a)(b) and (c) shows test frames and background subtraction results in second row

3.3.1 Proposed Adaptive CS

For any given scheme, different M and Φ are needed, as stated earlier the value of M is inversely proportional to the degree of sparsity of an image. If the same value of M is used for all different schemes, it is expected that the reliability of target detection will be different as the degree of sparsity varies from one image to another. For this reason there is a great challenge for adaptive CS by making M variable depending on how sparse the image is. For the adaptive CS, the CS process is preceded by a calibration phase. During that phase an Automatic Repeat Query (ARQ) transmission protocol is used between sensor nodes and the receiver side, as a feedback is needed for the adaptation phase. Initially, an arbitrary value of M is chosen according to a sparsity measure and is used to obtain the compressed measurements

\mathbf{Y}_d . The sensor node is then set to transmit \mathbf{Y}_d to the receiver side where the image is to be reconstructed, and based on the reconstruction error a decision is made whether the reconstruction is satisfactory or not. In case the reconstruction results are satisfactory, the receiver node sends a 'zero' flag through the feedback channel ending the calibration phase; otherwise a 'one' flag is to be sent. While the sensor node receives a 'one' flag, it is expected to change the value of M and change Φ accordingly, the sensor node repeats the search for an optimum value of M at the CS adaptation process till it receives a zero feedback from the receiver. At this point, the optimum values for M and Φ obtained are used next in the CS process. Fig.3 shows a flow chart summarizing the entire adaptive CS process.

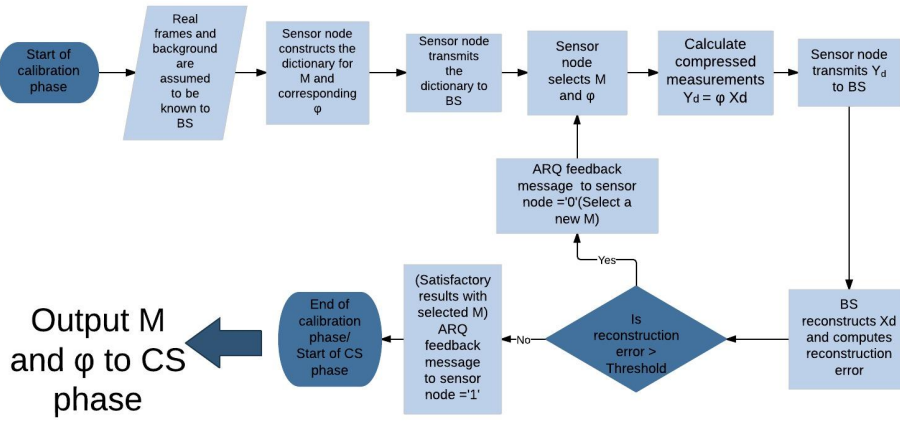


Fig. 3 Flowchart for the adaptive CS process



Fig. 4 the first image on the left is the background subtracted frame, the rest images are blocks containing the targets (with non-zero pixels)

3.3.2 Proposed block CS

To exploit the fact that the difference frame is always sparse, instead of compressing the whole frame, the image is divided into blocks and only blocks with non-zero pixels (containing the target) are proposed to be compressed and transmitted. This strategy is expected to help reduce the required value of M , subsequently, save the communication bandwidth and preserve the energy at visual sensor nodes. To illustrate this process, Fig.4(a) shows a (256×256) background subtracted frame which is then divided into 16 blocks (64×64) . It can be found that only 7 blocks have non-zero pixels as shown in Fig.4 and the rest of the blocks are all zeros, hence do not required to be processed. An index for each block is transmitted with each compressed measurements, such that the receiver side can reconstruct the whole image correctly again in the correct order. It has to be noted that the missing blocks (which has not been transmitted) are to be reconstructed as pixels sets with all zero values.



Fig. 5 The reconstructed original image for "Walking men"

Below are the steps undertaken during the entire process of block CS applied in the proposed work.

- **Step 1:** A frame with dimension $(N \times N)$ denoted as \mathbf{X}_d is divided into B blocks of size $(N_b \times N_b)$ each, where $\frac{N^2}{N_b^2} = B$, each block is denoted as \mathbf{X}_{blk}
- **Step 2:** for each block \mathbf{X}_{blk} with non-zero pixels, perform the following steps
 - **Step 3:** Φ is a randomly chosen sensing matrix of size $M_b \times N_b$, where $M_b \ll M$ and $M_b \ll N_b \ll N$
 - **Step 4:** produce the compressed measurements $\mathbf{Y}_{\text{blk}} = \Phi \mathbf{X}_{\text{blk}}$
 - **Step 5:** sensor nodes transmits the compressed block \mathbf{Y}_{blk} together with an index of the block number through the wireless channel
 - **Step 6:** at the receiver, Φ is assumed to be known for the decompression of \mathbf{Y}_{blk} . $\hat{\mathbf{X}}_{\text{blk}}$ is reconstructed from the compressed measurements \mathbf{Y}_{blk} .

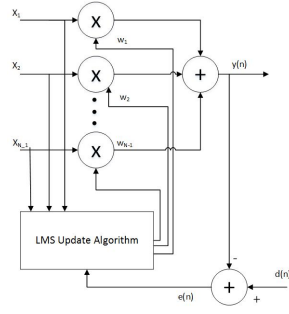


Fig. 6 An N-tap LMS adaptive filter

- **Step 7:** using the index number transmitted with every block all received blocks are placed together in the correct order resulting in a frame $\hat{\mathbf{X}}_d$ with only the foreground target present.
- **Step 8:** the original image $\hat{\mathbf{X}}_t$ is then obtained by adding $\hat{\mathbf{X}}_d$ (the reconstructed background subtracted image) to the masked background frame \mathbf{X}_b (masking the targets locations) which is also assumed to be known to the receiver side apriori as in Fig.5.
- **Step 9:** the targets locations are obtained after reconstructing the original image producing a trajectory for the complete path of each moving target

3.4 Proposed tracking model

3.4.1 Least mean square (LMS)

The LMS algorithm, is referred to as adaptive filtering algorithm since the statistics are estimated continuously, hence it can adapt to changes. LMS incorporates an iterative procedure during the training phase where it estimates the required coefficients to minimize the mean square error (MSE). This is accomplished through successive corrections to the expected set of coefficients which eventually leads to the minimum MSE. LMS algorithm is relatively simple, has much lower computational complexity than other adaptive algorithms; it does not require correlation function calculation nor does it require matrix inversions [15, 16]. LMS has been widely used in several real time image applications such as motion estimation and target tracking, where it showed robustness on fast moving targets and non-linear moving targets even in noisy environments as reported by the authors in [42–45]. The LMS implementation process has been illustrated in Fig.(6).

where, $\mathbf{x}(n)$ is the input signal, μ is the step size parameter and is selected by the autocorrelation matrix of the filter inputs, $\mathbf{e}(n)$ is the MSE between the predicted output $\mathbf{y}(n)$ and the reference signal $\mathbf{d}(n)$ which is given by (6). The LMS algorithm is initiated with an arbitrary value $\mathbf{w}(0)$ for the weight vector at $n = 0$. The LMS process involves estimation of $\mathbf{y}(n)$ from the received signal

$\mathbf{x}(n)$, by minimizing the error between the reference signal $\mathbf{d}(n)$, which closely matches or has some extent of correlation with the desired estimate signal and the output. The successive corrections of the weight vector eventually leads to the minimum value of the mean squared error.

$$\mathbf{e}(n) = (\mathbf{d}(n) - \mathbf{y}(n))^2 \quad (6)$$

the output $\mathbf{y}(n)$ is calculated as follows

$$\mathbf{y}(n) = \mathbf{x}(n)\mathbf{w}(n) \quad (7)$$

And the weight update can be given by the following equation

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{x}(n)\mathbf{e}(n) \quad (8)$$

There are several variants of the LMS algorithm present in the literature that deal with the shortcoming of its basic form and aim for lower computational complexity and faster adaptation processes as it is required for high speed communication as well as to be applicable in real time applications where the time is critical. A simple modification of LMS is called the Sign LMS algorithm [46], it uses the **sgn** function to clip the error in the weight update calculation. In clipped LMS [43, 47], clipped input is used to update the weight instead of the input itself. Another version of the LMS is the Normalized LMS (NLMS) [47], it forces the input samples to have a constant norm. Hence, it improves the convergence speed in a non-static environment by introducing a variable adaptation rate.

For the proposed model, LMS is used to predict target's locations, an iterative quantized clipped LMS technique is used with threshold values chosen to use the proposed tracking model. To guarantee least MSE an iterative method is proposed with a defined threshold of acceptable MSE, in addition to a threshold on the maximum number of iterations to maintain the algorithm's applicability for real time applications which is one of the main WVSNs properties.

$$mqsgn(\mathbf{x}(n)) = \begin{cases} 1 & \mathbf{x}(n) > D_1 \\ 0 & -D_2 < \mathbf{x}(n) < D_1 \\ -1 & \mathbf{x}(n) < -D_2 \end{cases}$$

Where, D_1 and D_2 are threshold values used to clip the input data. The modified iterative quantized clipped LMS algorithm consists of two main phases;

- **Learning Phase:** The LMS algorithm learns the targets locations to estimate new updates for the filter's weights till minimizing the MSE.
- **Prediction Phase:** The updated weights from the previous phase are then used to predict the target's next locations. The MSE will start rising again if the target changes its direction or speed, in that case the LMS needs to undergo the learning phase for further weight updates before next predictions.

The summary of the application of the LMS algorithm in the proposed tracking model:

- Prepare the input data and set the filter length
- Learning phase
 - While $MSE < \text{error's threshold}$ and $\text{number of iterations} < \text{iteration's threshold}$
 - Determine the output data using modified iterative quantized clipped LMS algorithm
 - Calculate the MSE
 - Update the filters weights according to the MSE
- Repeat the above steps till finishing the learning phase
- Predict the next locations using the updated weights
- If the MSE fell below some defined threshold repeat the learning phase

4 Simulations and results

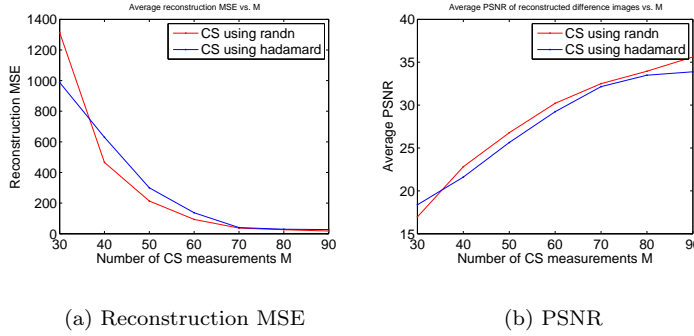


Fig. 7 Comparing reconstruction MSE and PSNR using randn and walsh sensing matrices for "Walking men"

Based on the system model proposed, simulations and experiments are conducted to evaluate the performance of the CS-based target detection and tracking algorithm. Simulations are performed for the WVSAN-based surveillance application in both outdoor and indoor scenes for single and multi-target tracking. Background and target's appearance are assumed to be static to investigate the effect of CS on the detection and tracking algorithms, hence schemes are chosen to reflect this assumption. Moreover, to illustrate the relation between the number of measurements required for CS to guarantee reconstruction and how sparse the image is. Simulations are performed on different schemes with different sparsity levels; "Walking men" is chosen to resemble an outdoor scenes for multi target tracking captured by [48]. While "Shopping

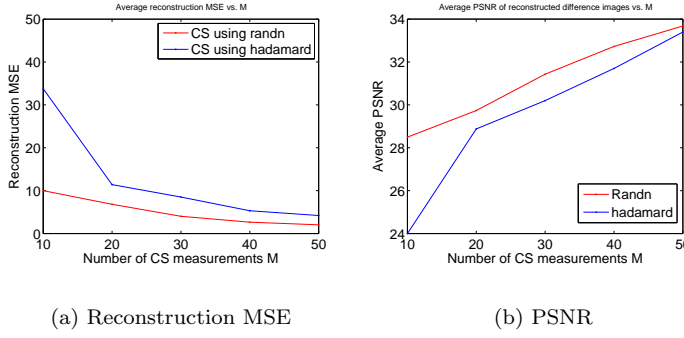


Fig. 8 Comparing reconstruction MSE and PSNR using randn and walsh sensing matrices for "Shopping center 1"

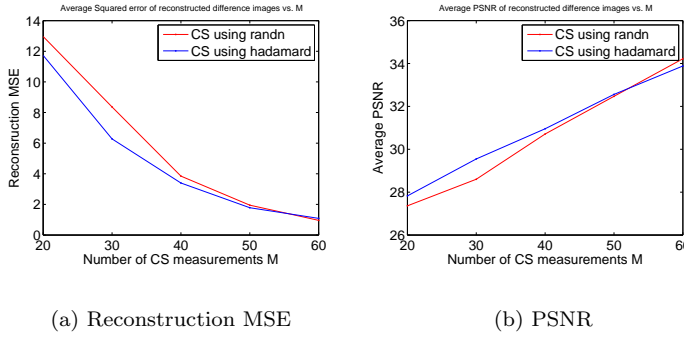


Fig. 9 Comparing reconstruction MSE and PSNR using randn and walsh sensing matrices for "Shopping center 2"

center 1" and "Shopping center 2" filmed for the EC funded CAVIAR project found in [49] for indoor scenes tracking a single target.

Mean square error (MSE) and peak signal to noise ratio (PSNR) are used as performance indicators to test the reliability of CS. MSE and PSNR are compared for different number of CS measurements M , where the MSE is the reconstruction error measured between real and reconstructed frames and PSNR is measured after frames recovery to reflect the quality of image reconstruction which will later on reflects the ability of reliable tracking. The background frame and Φ are known to the receiver node. Two candidate sensing matrices have been compared; normally distributed random numbers using Matlab function "randn" and a walsh-hadamard. Although the measurements are defined by a matrix multiplication, the operation of matrix-by-vector multiplication is seldom used in practice, because it has a complexity of $O(MN)$ which may be too expensive for real time applications. When a randomly permuted Walsh-Hadamard matrix is used as the sensing matrix, the mea-

surements may be computed by using a fast transform which has complexity of $O(K \log(N))$ [50]. The Hadamard matrix, is an $(N \times N)$ square matrix whose entries are either +1 or -1 and whose rows are mutually orthogonal, the matrix is first randomly reordered then, M samples are randomly chosen to construct the $(M \times N)$ random sensing matrix Φ .

The ability of reliable tracking depends on acceptable recovery of images. In other words, if CS fails in image reconstruction the targets location can not be detected. Hence, M is adaptively chosen depending on the sparsity nature of images as choosing the right value of M is critical in image reconstruction and afterwards tracking. It is clear from the results in Fig.7, 8 and 9, that for different sparsity levels different values of M and compression rates are required. When reaching optimum value of M least MSE and 33dB PSNR are successfully achieved. For illustration, MSE decreases and PSNR increases as M increases till reaching the optimum value, it has been shown that the lower bound on M is depending on how sparse the difference frame X_d is or in other words proportional to the ratio between the number of non-zero coefficients and the total number of pixels in a frame. For "Walking men", CS sets M to 70 as in Fig.7(a) to achieve satisfactory results. While for "Shopping center 1" and "Shopping center 2", it is obvious from Fig.8(a) and Fig.9(a) respectively, that for single-target tracking (where there is lower number of non-zero coefficients), better MSE is achieved with lower M , reduced to 50 and 60 for Shopping center "1" and "2" respectively compared to multi-target tracking while maintaining least MSE and 33dB PSNR. As a result, making CS adaptive helps in increasing the compression rate and avoiding the waste of using a higher value of M at the times where the image is sparse allowing for lower M . The above discussion reflects the reduction in channel bandwidth using CS by 72%, where instead of transmitting the whole (256×256) image, the compressed measurements of size (70×256) are transmitted. Whereas for more sparse images the reduction reaches 82% of the total image size.

The above simulation were carried out using two different sensing matrices, Randn and walsh-Hadamard. They are compared with respect to MSE and PSNR as in Fig.7, 8 and 9. It is clear from the results that when reaching the optimum value of M both sensing matrices perform nearly the same except in some cases in Fig.8 shows that Randn gives slightly a better performance than Hadamard. But this can be negligible when compared to the reduction in complexity gained by using Hadamard matrix which helps in accomplishing the main objective to save sensor nodes power and as a result maximizes their lifetime.

In the simulation results presented, the performance degradation due to channel impairments was not considered as we were examining the impact of CS on the tracking performance and MSE and PSNR are used as performance indicators (which are widely used quality measures of Image reconstruction). However, considering channel impairments, we have generated the simulation again for one "Shopping center1" using a 30dB PSNR during wireless transmission of 2Mbps bitrates and tested the performance of CS reconstruction along with additive white Gaussian noise in the channel. As seen in Fig.10,

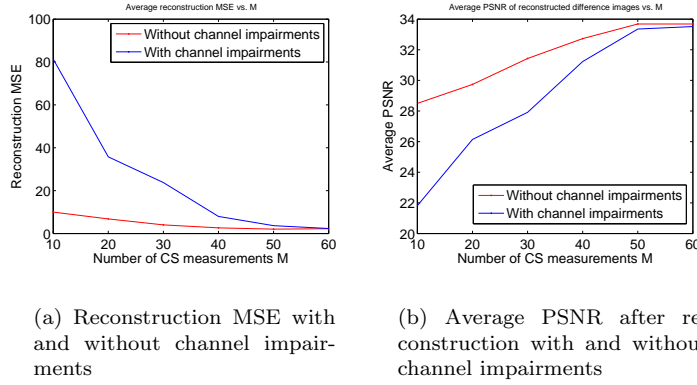


Fig. 10 Comparing reconstruction MSE and PSNR with and without considering channel impairments for "Shopping center 1"

reconstruction MSE reaches the same level as that without considering channel noise with the same PSNR by a slight increase in M (starting from 55) compared to the previous results without channel noise M reaches 50 for a guaranteed reconstruction.

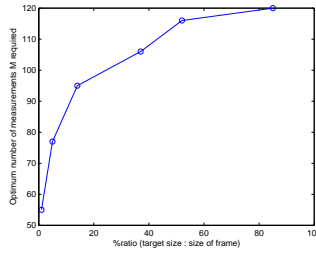


Fig. 11 Relation between the percentage ratio of target size:frame size vs. M

Fig.11 summarizes and demonstrates the effect of the target size ratio on the number of measurements M needed (the target size ratio is expressed as a ratio between non-zero pixels representing the target and the total size of the image frame, which reveals how much space the target acquires and how sparse the image is). It is clear from Fig.11 that for smaller target sizes, lower values of M are used. While as the target size increases, the required M increases to achieve the same performance in terms of MSE and PSNR. These results reflect the constraint on the lower bound of M and give a key to the problem when M is required to be kept as small as possible. Where in that case the size of targets is controlled by zooming or changing the location of sensor nodes during the calibration phase while bearing in mind to keep the scene of interest in the camera's field of view. By taking snapshots from a further location the

total space acquired by the target is hence reduced and as a result M can be reduced, and the goal of reducing the size of transmitted data is met .

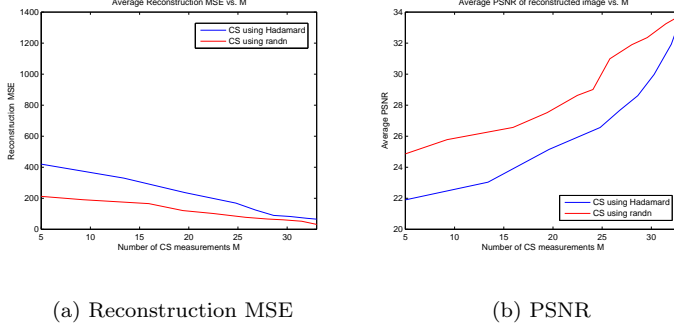


Fig. 12 Comparing reconstruction MSE and PSNR using randn and walsh sensing matrices for block CS

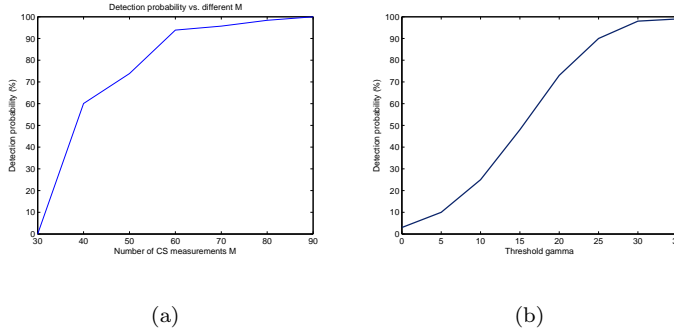


Fig. 13 Probability of detection vs. (a) different values of M and (b) different values of background subtraction threshold γ

Fig.12(a) illustrates the reduction in MSE and the number of measurements M required when dividing each frame into blocks. According to the sparsity level, the number of blocks are chosen. For "Walking men" each frame is divided into 16 blocks (64×64) each, and compressing only those with non-zero pixels. Compared to Fig.7(a) (compressing the whole frame), $\approx 70\%$ reduction in MSE is achieved without compromising an adequate PSNR of $\approx 33dB$ attained in Fig.7(b), PSNR versus the number of measurements M for block CS is shown in Fig.12(b). Demonstrating the reduction of the number of measurements needed, as seen in figures, for the normal scenario, M is set to 70

yielding $(70 \times N)$ measurements. Whereas, for the blocks scenario $\approx (30 \times 64)$ measurements are required per block which yields an extra communication bandwidth reduction by 40% for the total blocks transmitted compared to the normal CS-scenario which yields a total compression rate of 82%. This saves the communication bandwidth and resulting in faster transmission while saving energy at sensor nodes. Comparing CS with traditional H.264, H.264 compression rates can reach 90% for a PSNR of 31dB [51], whereas for CS compression rates can reach 82% and 33dB. However, this reduction of H.264 compression rates is at the expense of complex computations as compared to simple encoders building the CS process [26].

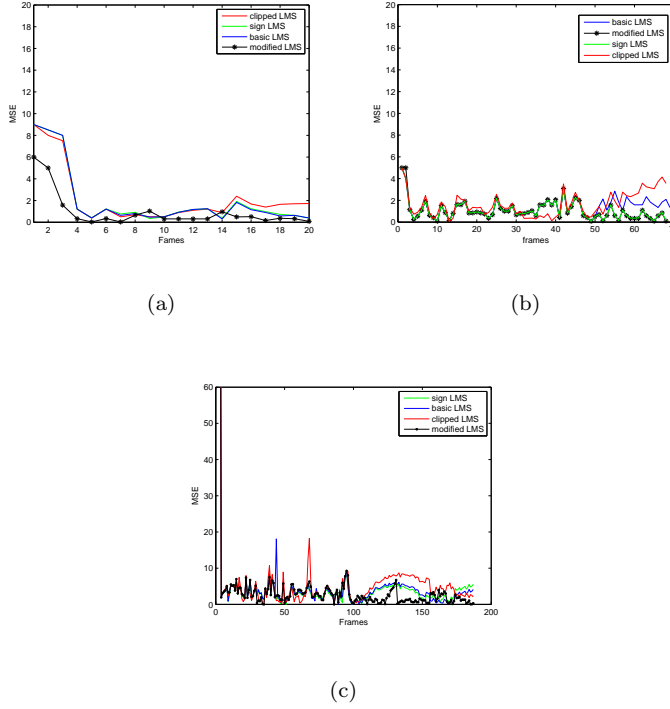


Fig. 14 Comparing MSE for different variants of LMS for (a) dataset 1 (b) dataset 2 and (c) dataset 3

Fig.13 shows the probability of detection versus different parameters, Fig.13(a) for different values of measurements M , it is clear from the graph that for lower values of M the target is misdetected. This reflects the fact that the reconstruction can not be guaranteed with lower values of M . The probability of detection increases till reaching 100% as M increases to its optimum value selected during the CS process. Fig.13(b) demonstrates the effect of background

subtraction threshold γ on the detection problem, where for lower values of γ , low probability of detection is achieved as the target may be misdetected due to unwanted noise.

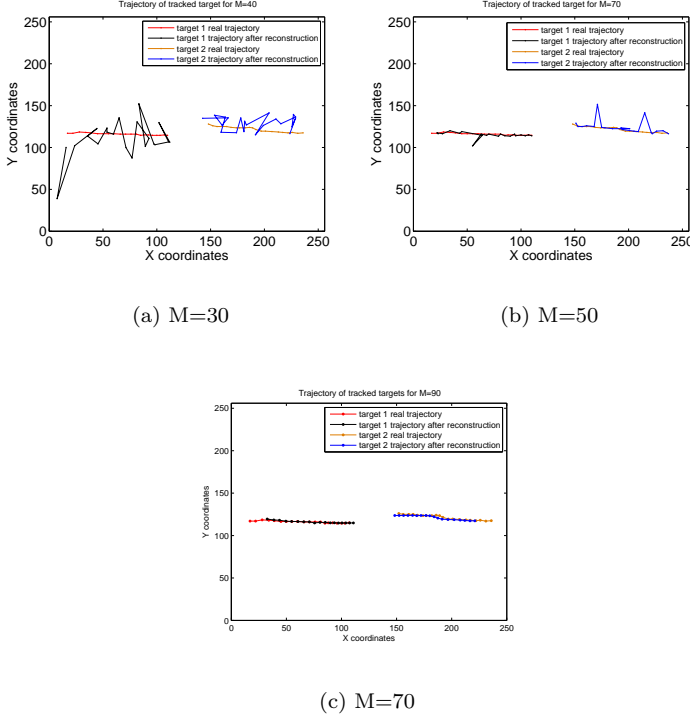


Fig. 15 Comparing predicted trajectory of multi-targets using LMS for "Walking men" (using different M for CS)

Several variants of LMS are implemented and compared, the basic LMS, clipped LMS, sign lms, and the proposed iterative quantized clipped LMS. MSE is also used as an indicator to test tracking reliability which is the error between the real target locations and the predicted target locations by the different LMS algorithms after image recovery for n consecutive frames. The thresholds are chosen for the proposed LMS based on experiments and targets locations in each dataset. They are set as follows; for the outdoor scheme "Walking men" $D_1 = 200$ and $D_2 = 120$, for the "Shopping center 1" $D_1 = 160$ and $D_2 = 140$, and for the "Shopping center 2" $D_1 = 240$ and $D_2 = 200$. Fig.14 show the MSE for the different variants of LMS for the 3 schemes. Experiments have shown that the clipped LMS did not perform better than the basic LMS as all input data were clipped to value one. The iterative modified LMS have the least MSE due to the lower bound constraint on the MSE, at the

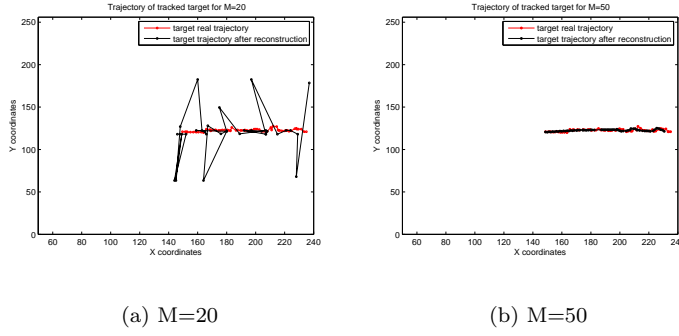


Fig. 16 Comparing predicted trajectory of single target using LMS for "Shopping center 1" (using different M for CS)

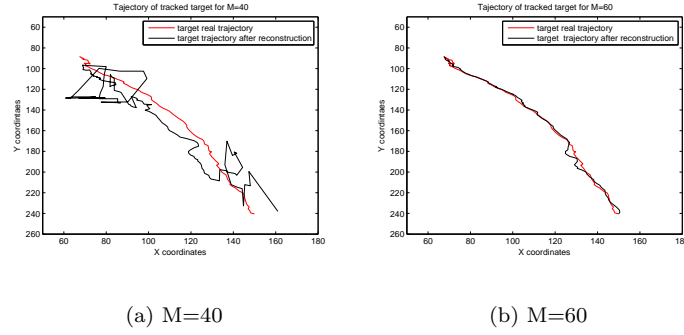


Fig. 17 Comparing predicted trajectory of single target using LMS for "Shopping center 2" (using different M for CS)

times when the MSE rises, the algorithm goes into the learning phase again before estimating new locations. In Fig.14(b), the signed LMS gave the same performance as the modified LMS, whereas the MSE for some datasets as in Fig.14(a) and 14(c) is the same as the basic LMS.

CS states that when enough measurements are used for compression, the reconstruction is done with high accuracy depending on a lower bound of M . Trajectory tracking of moving targets is considered to reflect the degree of reconstruction accuracy. Tracking reliability is tested by comparing the moving target's real and predicted trajectories using the proposed iterative quantized clipped LMS. Fig.15, 16 and 17 show the (x,y) position plots of the path tracked for the targets in the camera's scene. Fig.15(a) and 15(b) show that (for the outdoor "Walking men") for lower values of $M < \text{optimum value}$ (30 and 50 respectively), frames can not be reconstructed properly and as a result the targets tracks are not matching their real trajectories, whereas for optimum values of M reaching 70, LMS accurately predicted the target's

locations and the results are closely matching the real target trajectory before compression. Fig.16 and 17 illustrate the same for the indoor schemes.

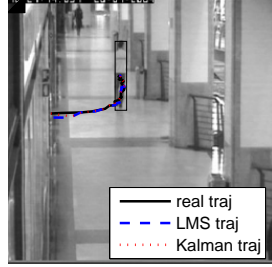


Fig. 18 Comparing predicted trajectory using LMS and Kalman filter

The performance of the proposed LMS algorithm is compared with state-of-the-art Kalman filter [52], both algorithms are applied on a standard surveillance video 'OneStopMoveNoEnter1cor' from CAVIAR [49] (same dataset used by the authors in [52]). In the video a man is selected for tracking and tracked in subsequent frames using both LMS algorithm and Kalman filter. Fig.18 shows output frame number 997 with trajectories of the target moving in the corridor since its appearance in the video, the figure shows that both LMS and Kalman filter matches the real trajectory of the target.

4.1 Computational complexity

Assuming all sensor nodes have the same unit distance d from the receiver side, Table.1 shows the energy dissipated during transmission for different k (number of samples transmitted). As illustrated, according to different k (which varies depending on compression rates due to sparsity levels), there is an 82% energy saving as compared to transmitting the captured image without CS. In addition, using block CS will result in 20% more energy saving compared to traditional CS.

Table.2 summarizes the computational time for the traditional CS process, block CS and the LMS tracking technique. As stated in sec.1 the LMS algorithm is relatively simple, has much lower computational complexity than the original Kalman filters and other adaptive algorithms and suitable for real time applications due to its fast convergence as demonstrated.

Table 1 Transmission energy using CS, block CS and without CS for different k

			Transmission Energy Etx
Without CS		k=64K	3.3mJ
CS	"Walking men"	k=17K	0.85mJ
	"Shopping center 1"	k=15K	0.7mJ
	"Shopping center 2"	k=12K	0.6mJ
Block CS		k=11K	0.5mJ

Table 2 Computational time for CS, block CS and LMS

	Computational time
CS process	0.03s
Block CS	0.002s/block
LMS	0.002s

5 Conclusion

WVSNs are characterized as resource constraints due to limited battery power, memory space and communication bandwidth. These constraints brought new implementation challenges to investigate adaptive CS in designing robust target detection and tracking techniques for surveillance applications without compromising the tracking performance as well as energy constraint. CS is investigated as it has been expected to be a strong candidate to achieve high compression rate using simple computations.

An adaptive CS technique has been proposed and has achieved high compression rates with minimum reconstruction error. In addition, block CS is energy efficient where it divides the image frame into blocks and only blocks containing the target is compressed and transmitted in contrast to basic CS where the whole frame is compressed, hence saving power of processing.

Experiments were carried out to evaluate the performance of the adaptive CS and its effect on target detection and tracking. Results have shown that using adaptive CS, the reconstruction MSE decreases till reaching the lower bound on the number of compressed measurements while preserving the acceptable PSNR and addressing the problems of WVSN, such as energy, memory and bandwidth constraints. In addition, for different datasets where the sparsity nature of each image differs, CS adaptively chooses the compression rates accordingly reaching a relation between the the number of compressed measurements and ratio of non-zero pixels to the total number of pixels. As a result, when higher compression rates are required, one control the targets size by zooming out or changing the location of sensor nodes during the CS calibration phase while bearing in mind to keep the scene of interest in the camera's field of view. Moreover, block CS achieved higher compression rates with lower reconstruction MSE saving the communication bandwidth and resulting in faster transmission. After image reconstruction, the impact of adaptive CS on target tracking is investigated where the proposed iterative

quantized LMS is performed for target tracking and is compared with other variants of LMS. Results have demonstrated that the proposed LMS technique achieved the least MSE. Target's trajectory tracking has been used as another performance indicator for the LMS algorithm, it is shown that the predicted path closely matches the target's real path which illustrates the accuracy of LMS and that CS has not degraded the performance of target detection and tracking.

References

1. F.G.H.Yap and H.H.Yen, "A survey on sensor coverage and visual data capturing/processing/transmission in wireless visual sensor networks," *Sensors*, vol. 14, pp. 3506–3527, February 2014.
2. T.Winkler and B.Rinner, "Security and privacy protection in visual sensor networks: A survey," *ACM Computing Surveys (CSUR)*, vol. 47, no. 1, July 2014.
3. A.Sharif, V.Potdar, and E.Chang, "Wireless multimedia sensor network technology: A survey," in *Proceedings of Industrial Informatics, 7th IEEE International Conference, 2009. INDIN 2009.*, June 2009, pp. 606–613.
4. I.F.Akyildiz, T.Melodia, and K.R.Chowdhury, "A survey on wireless multimedia sensor networks," *computer Networks*, vol. 51, pp. 921–960, March 2007.
5. S.Soro and W.Heinzelman, "A survey on visual sensor networks," *Hindawi publishing corporation, Advances in Multimedia*, vol. 2009, no. 640386, pp. 1–21, May 2009.
6. Y.Charfi, B.Canada, N.Wakamiya, and M.Murata, "Challenges issues in visual sensor networks," in *IEEE on wireless Communications*, April 2009, pp. 44–49.
7. X.Wang, S.Wang, and D.Bi, "Distributed visual-target-surveillance system in wireless sensor networks," *IEEE Transactions on Systems, MAN, and Cybernetics*, vol. 39, no. 5, pp. 1134–1146, October 2009.
8. X.Wang, S.Wang, D.W.Bi, and J.J.Ma, "Distributed peer-to-peer target tracking in wireless sensor networks," *MDPI, open access journal on the science and technology of sensors and biosensors*, vol. 7, pp. 1001–1027, 2007.
9. X.Wang and S.Wang, "Collaborative signal processing for target tracking in distributed wireless sensor networks," *Elsevier journal on Parallel and distributed computing*, vol. 67, p. 501–515, 2007.
10. S. M.Youssef, A. A. ElFarag, and N. M.Ghatwary, "Adaptive video watermarking integrating a fuzzy wavelet-based human visual system perceptual model," *Multimedia Tools and Applications, Springer*, August 2013.
11. H.S.Parry, A.D.Marshall, and K.C.Markham, "Region template correlation for the flir target tracking," in *British Machine Vision Conference*.
12. D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *The Institution of Engineering and Technology, Control theory applications*, vol. 4, no. 8, pp. 1303–1318, 2010.
13. J.C.Noyer, P.Lanvin, and M.Benjelloun, "Non-linear matched filtering for object detection and tracking," *Elsevier Pattern Recognition Letters*, vol. 25, pp. 655–668, 2004.
14. A.C.Sankaranarayanan, A.Veeraraghavan, and R.Chellappa, "Object detection, tracking and recognition for multiple smart cameras," in *Proceedings of the IEEE*, vol. 96, no. 10, October 2008, pp. 1606–1624.
15. S.Haykin, *Adaptive Filter Theory*. Prentice Hall, 2002, vol. 0-13-048434-2, ch. Least mean square adaptive filters, pp. 231–247.
16. P.S.R.Diniz, *Adaptive Filtering*. The Springer International Series in Engineering and Computer Science, January 1997, vol. 399, ch. The Least-Mean-Square (LMS) Algorithm, pp. 79–135.
17. R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, pp. 118–124, July 2007.
18. V.Cevher, A.Sankaranarayanan, M. Duarte, D.Reddy, R. Baraniuk, and R.Chellappa, "Compressive sensing for background subtraction," 2008.

19. E. Wang, J. Silva, and L. Carin, "Compressive particle filtering for target tracking," in *IEEE/SP 15th Workshop on Statistical Signal Processing, SSP*, September 2009, pp. 233–236.
20. M.Zhao, A.Wang, B.Zeng, L.Liu, and H.Bai, "Depth coding based on compressed sensing with optimized measurement and quantization," *Ubiquitous International Journal of Information Hiding and Multimedia Signal Processing*, vol. 5, no. 3, pp. 475–484, July 2014.
21. C.Patsakis and N.G.Aroukatos, "Lsb and dct steganographic detection using compressive sensing," *Ubiquitous International Journal of Information Hiding and Multimedia Signal Processing*, vol. 5, no. 4, pp. 20–32, January 2014.
22. J.Haupt and R.Nowak, "Compressive sampling vs. conventional imaging," in *IEEE International Conference on Image Processing*, October 2006, pp. 1269–1272.
23. M.S.Asif, F.Fernandes, and J.Romberg, "Low-complexity video compression and compressive sensing," in *In Asilomar Conference on Signals, Systems, and Computers*, 2013.
24. A.Mahalanobis and R.Muise, "Object specific image reconstruction using a compressive sensing architecture for application in surveillance systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1167–1180, July 2009.
25. H.R.ALZoubi, "Video coding and routing in wireless video sensor networks," *AASRI Conference on Parallel and Distributed Computing and Systems*, vol. 5, no. 0, pp. 48 – 53, 2013.
26. S.Pudlewski, A.Prasanna, and T.Melodia, "Compressed-sensing-enabled video streaming for wireless multimedia sensor networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 6, pp. 1060–1072, June 2012.
27. J.Romberg, "Imaging via compressive sampling," *IEEE Signal Processing Magazine*, pp. 14–20, March 2008.
28. H. C. Huang and F. C. Chang, "Robust image watermarking based on compressed sensing techniques," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 5, no. 2, pp. 275–285, April 2014.
29. E.J.Candes and M.B.Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, pp. 21–30, March 2008.
30. E.J.Candes, "Compressive sampling," in *Proc. of the International Congress of Mathematicians*, 2006.
31. A.Hormati, O.Roy, Y.M.Lu, and M.Vetterli, "Distributed sampling of signals linked by sparse filtering: theory and applications," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1095–1109, March 2010.
32. J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, December 2007.
33. D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
34. Ms.V.MuthuLakshmi, "Advanced leach protocol in large scale wireless sensor networks," *International Journal of Scientific and Engineering Research*, vol. 4, no. 5, pp. 248–254, May 2013.
35. A.Redondi, D.Buranapanichkit, M.Cesana, M.Tagliasacchi, and Y.Andreopoulos, "Energy consumption of visual sensor networks: Impact of spatio-temporal coverage," *IEEE transaction on Circuits and Systems for Video Technology*, vol. 24, no. 12, pp. 2117–2131, December 2014.
36. P.Mohanty, M.R.Kabat, and M.K.Patel, "Energy efficient transmission control protocol in wireless sensor networks," in *Wireless Networks and Computational Intelligence*, ser. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2012, vol. 292, pp. 56–65.
37. G.J.Pottie and W.J.Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, may 2000.
38. T.Melodia, D.Pompili, and I.F.Akyildiz, "A communication architecture for mobile wireless sensor and actor networks," in *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks, SECON*, vol. 1, Sept 2006, pp. 109–118.

39. F.Shebli, I.Dayoub, A.O.M'foubat, A.Rivenq, and J.M.Rouvaen, "Minimizing energy consumption within wireless sensors networks using optimal transmission range between nodes," in *IEEE International Conference on Signal Processing and Communications, ICSPC*, Nov 2007, pp. 105–108.
40. E.M.Martn and .P.Pobil.
41. N.Efford, *Digital Image Processing: A Practical Introduction Using JavaTM*. Pearson Education, 2000, ch. Morphological image processing.
42. Y.Zheng, H.Wang, and Q.Guo, "A novel mean shift algorithm combined with least square approach and its application in target tracking," in *IEEE 11th International Conference on Signal Processing (ICSP)*, vol. 2, Oct 2012, pp. 1102–1105.
43. H.S.Yazdi, M. Lotfizad, and M.Fathy, "Car tracking by quantised input lms, qx-lms algorithm in traffic scenes," *Vision, Image and Signal Processing, IEE Proceedings*, vol. 153, no. 1, pp. 37–45, 2006.
44. C.-H. Hsia, Y.-P. Yeh, T.-C. Wu, J.-S. Chiang, and Y.-J. Liou, "Low resolution method using adaptive lms scheme for moving objects detection and tracking," in *Intelligent Signal Processing and Communication Systems (ISPACS), 2010 International Symposium on*, 2010, pp. 1–4.
45. T.W.Bae, F.Zhang, and I.S.Kweon, "Edge directional 2d lms filter for infrared small target detection," *Infrared Physics and Technology, Sciencedirect*, vol. 55, no. 1, pp. 137–145, 2012.
46. S.B.Jebara and H.Besbes, "A variable step size filtered sign algorithm for acoustic echo cancellation," in *IEEE electronic letters*, vol. 39, 2003, pp. 936–93.
47. S.Dhull, S.Arya, and O. Sahu, "Performance variation of lms and its different variants," *International Journal of Computer Science and Security, (IJCSS)*, vol. 4, 2010.
48. F. Cheng and Y. Chen, "Real time multiple objects tracking and identification based on discrete wavelet transform," *Elsevier Pattern Recognition Journal*, vol. 39, p. 1126 1139, 2006.
49. "Caviar datasets," Dataset: EC Funded CAVIAR project/IST 2001 37540, <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>, 2001.
50. H. Jiang, W. Deng, and Z. Shen, "Surveillance video processing using compressive sensing," *arXiv preprint arXiv:1302.1942*, 2013.
51. A.RAVI and K.R.RAO, "Performance analysis and comparison of the dirac video codec with h.264/mpeg-4 part 10 avc," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 09, no. 04, pp. 635–654, 2011.
52. H.A.Patel and D.G.Thakore, "Moving object tracking using kalman filter," *International Journal of Computer Science and Mobile Computing, IJCSMC*, vol. 2, no. 4, pp. 326–332, April 2013.