IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 20, Issue 3, Ver. IV (May. - June. 2018), PP 49-53
www.iosrjournals.org

Block Compressed Sensing of Videos: A Case Study

Divya R. Bora, Mrs. Smita Jangale
Department of Information Technology, V.E.S. Institute of Technology, Mumbai, India
Corresponding author: Divya R. Bora

Abstract: *Compressive sensing is a technique for the data sampling and compressing simultaneously. In this paper, we study block compressive sensing for videos. Gaussian Mixture Model subtract the background from the video frame to get the detected object. The detected object acquisition is conducted using block-by-block mechanism. The block compressed sensing framework sufficiently capture the video frames. Several video samples demonstrate that the block CS technique reconstruct the video accurately and efficiently with the less complexity. Higher subrate values and larger dimension of sensing matrix improves the quality of the reconstructed frames.*
Keywords: *Block compressed sensing, Gaussian Mixture Model, Video frames.*

Date of Submission: 20-06-2018 Date of acceptance: 04-07-2018

I. Introduction

Image and video processing with object detection and pattern recognition are massive topics in the field of research concerning the automatic analysis of images and image sequence. It has a broad spectrum of applications in remote sensing, medical diagnosis, human-computer interaction or video compression etc. The aim of video processing is to help human administrators in checking Closed Circuit Television (CCTV) camera systems for better surveillance.

Compressed sensing (CS) is an extensive and rapidly growing field that has huge applications in signal processing, statistics, and computer science. It has been brought to our attention only a few years ago, thousands of papers have been published in this area and more than hundreds of conferences and workshops have been organized for this growing research field.

Compressive Sensing framework achieves compression by changing the sensing prototype, hence emerging as a revolutionary idea. The aim of CS is to achieve sensing and compression in a single step. It has been shown that a sparse signal can be reconstructed by a small number of linear, non-adaptive measurements. Basically, we capture the features of the signal as measurements of the data and not the complete signal itself. CS theory states that the perfect reconstruction of this redundant data is possible with these limited features of the captured signals. CS is based on the premise of redundancy in information just like most compression techniques. However, it can in principle, outperform some compression techniques in memory consumption, simplicity of encoding and cost of operation. For example, in the sampling context, CS allows redundant data to be represented by far fewer samples than the Nyquist Limit. This hold promises for the development of cheaper and more robust high frequency systems [1]-[4].

II. Background Subtraction

Background modelling of a video frame is an important task in the field of computer vision due to the uncontrolled nature of the environment. The background subtracted frame is then further used for pre-processing step for image or video transmission and compression. Background subtraction is usually two steps procedure, framing statistical representation of the background scene and detecting the foreground by subtracting the background from the scene.

Gaussian Mixture Model (GMM) is one of the techniques of background subtraction [5], [6]. The GMM divides every pixel into Gaussian distribution. Let's assume that $X_1, .., X_t$isthe value of the intensity of each pixel from 1 to time *t*.

$$X_1, .., X_t = I(x_0, y_0, i): 1 \le i \le t \quad (1)$$

Where I is the image sequence. Each pixel is modelled in different times using K Gaussian distribution and the parameters of the Gaussians are updated at each time. These distributions consist of both foreground and background. For K Gaussian distribution, the probability of every pixel is computed regardless of foreground or background using the equation [7]

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2)$$

where, X_t - current pixel in frame t

K - the number of distributions in the mixture

$\omega_{i,t}$ - the weight of the kth distribution in frame t

$\mu_{i,t}$ - the mean of the kth distribution in frame t

$\Sigma_{i,t}$ – the standard deviation of the kth distribution in frame t

$\eta(X_t, \mu_{i,t}, \Sigma_{i,t})$ - probability density function (pdf):

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp -\frac{1}{2}(X_t - \mu)\Sigma^{-1}(X_t - \mu) \quad (3)$$

The covariance matrix is formulated as

$$\Sigma_{i,t} = \sigma_{i,t}{}^2 I \quad (4)$$

The bigger Gaussian distribution than the designated threshold is termed as background. The other distribution is categorized as foreground.

$$B = argmin_b \left(\sum_{i=1}^{b} \omega_{i,t} > T\right) \quad (5)$$

If the pixel of the frame matches with one of the K Gaussian, the value of the ω, μ and σ is updated. If every parameter has been calculated, then the foreground detection can be implemented.

III. Compressed Sensing

Conventional approaches to sampling signals or images follow Nyquist-Shannon theorem: the sampling rate must be at least twice the maximum frequency n present in the signal. Using the Nyquist rate of sampling, n samples are generated and then compressed to m samples by removing n – m samples. The decompression of the signal data is achieved at the receiver end to retrieve n samples from m samples. But the implementation of high sampling rates may not be practical for circuitry. To achieve a target level of accepted distortion with minimal computations, the concept of compression is adopted. The compression theory has the basis of finding a frame that provides compressible or sparse representations for the signal of interest [2].

Compressed Sensing (CS) overcome the limits of traditional sampling theory. The compressed sensing theory states that recovery of certain signals and images from far fewer samples or measurement can be done. If the signal which is being sensed has low information rate in original or some transform domain, then this recovery becomes precise. The reconstruction algorithm decides the least number of samples required. The recovered signal is best reconstruction if the signal is not sparse. CS has emerged as a new framework for sensor design and signal acquisition which leads to large reduction in computation costs and sampling for sensing signals that have a compressible or sparse representation [2], [8].

CS relies on two principles [1], [2]:

1) Sparsity: Continuous time signals such as sound, image can be stored in compressed form on suitable basis. The choice of proper basis Ψ reflects the omission of zero or small projection coefficients. If a

signal has only k non-zero coefficients, it is said to be k-Sparse. The signal is compressible when a large number of projection coefficients are ignored as they are small enough.

2) Incoherence: The maximum correlation between any two elements of two different matrices is termed as Coherence. These two matrices represent different basis. Let $n \times n$ matrix, Ψ with columns $\Psi_1, .. \Psi_n$ and an
$m \times n$ matrix, Ψ with rows $\Psi_1, .. \Psi_n$. Then coherence μ is computed as:

$$\mu(\Phi, \Psi) = \sqrt{n} \max |\Phi_k, \Psi_j| \quad (6)$$

For, $1 \leq j \leq n,\ 1 \leq k \leq m.$

The incoherence of matrix helps to reconstruct the sensing signal in CS framework.

Consider an n-dimensional signal f, which has a representation in some orthonormal basis $\Psi = [\Psi_1, \Psi_2, .., \Psi_n]$ where

$$f = \sum_{i=1}^{n} x_i \psi_i = \Psi x \quad (7)$$

In the above, x_i and ψ_i (a column vector) are the i^{th} coefficient and i^{th} basis, respectively. The CS theory says that if *f* is sparse in the basis Ψ, which needs not be known apriori, then, under certain conditions, taking *m* nonadaptive measurements of *f* suffices to recover the signal exactly, where $m \ll n$. Each measurement y_j is a projection of the original signal. The *m*-dimensional measurement can be represented by

$$y = \Psi x \quad (8)$$

where *y* is the measurement vector and Ψ is an $m \times n$ sensing matrix.

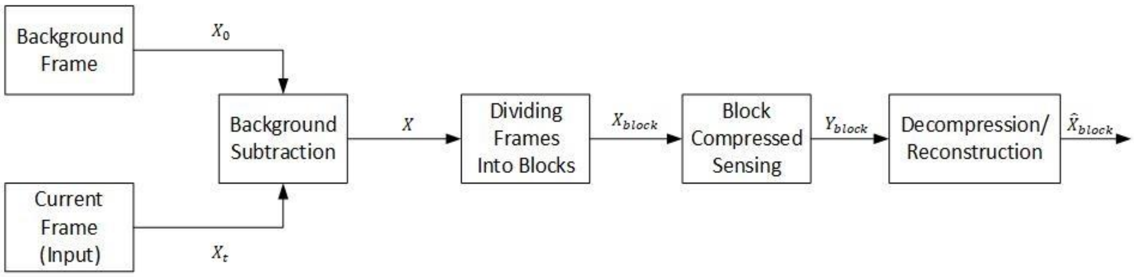


Fig. 1: Block Diagram of Block Compressed Sensing

A. BLOCK COMPRESSED SENSING

In block CS, to use the fact that the difference frame is always sparse, the image or a video frame is divided into blocks. The blocks with non-zero pixels which contains target are compared and transmitted. The Fig. 1 represents the block diagram of the block CS. The Input consists of two frames, a background frame X_0 and the current frame X_t. The background frame is fixed and always compared with the current frame. The background subtraction using GMM subtracts the background frame from the target frame and the resultant frame *X* is the input frame for block CS [9]. Consider a frame *X* of dimension $N \times N$ is divided into blocks *B* of size $N_{block} \times N_{block}$ each and *M* is the required number of samples. The number of blocks, *B* is computed as

$$B = \frac{N^2}{N_{block}^2} \quad (9)$$

Let X_{block} denotes each block and Y_{block} denotes the respective measurement. For each block with non-zero pixels, the measurement matrix, Φ with dimension $M_{block} \times N_{block}$ is selected where $M_{block} \ll M$ and $M_{block} \ll N_{block} \ll M$. Thus, the compressed measurement is

$$Y_{block} = \Phi \hat{X}_{block} \tag{10}$$

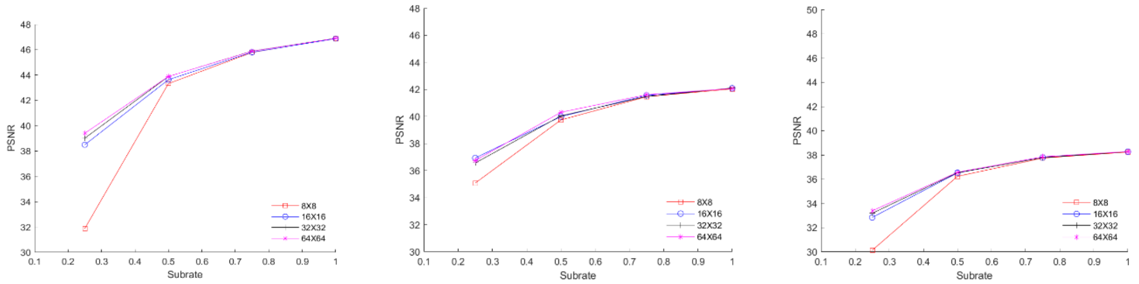
IV. Results

The block compressed sensing, sampling and reconstruction algorithms were implemented using MATLAB on a 2.10 GHz laptop computer. The three videos used for the simulation are retrieved from the standard database [10]. These videos are (a) a walk by shop, (b) enter exit crossing path and (c) two persons leave the shop. The results after applying the block compressed sensing are based on the calculation of PSNR MSE values. For each of the sample videos, the PSNR MSE values are calculated by altering the subrate from 0.25 to 0.99 and block size from 8×8 to 64×64. Subrate denotes the ratio of the length of sample M to the length of signal N. Table I tabulates the Mean Square Error (MSE) results for all three sample videos by changing the subrates for the block size 64×64. It can be deduced from the Table 1 that accuracy of the reconstructed frame after the block compressed sensing increases with the higher subrate values.

Table no 1: The Mean Square Error (MSE) values for all three sample videos (block size = 64×64)

Sr. No	Subrate	Sample Video 1	Sample Video 2	Sample Video 3
1	0.25	7.04759	13.7518	29.43014
2	0.5	2.57334	6.04964	14.11907
3	0.75	1.59914	4.4687	10.74578
4	0.99	1.2992	3.98355	9.65485

Fig. 2 represents the PSNR results for the sample videos, by differing the subrate values for each block size i.e. 8×8, 16×16, 32×32, 64×64. Increasing the size of the sensing matrix provides better quality in the recovered frame in terms of PSNR. For example, at subrate 0.1 with a block size 8×8, the recovered frame of the sample video 1 has a PSNR 31.8769dB. However, this value can be increased upto 39.4 dB with a block size of 64×64. Moreover, the object in the frames which is detected and reconstructed from the sample video 1 are illustrated in Figs. 3 and 4. The subrate is kept as 0.99 and the block size as 64×64. The PSNR value computed for the reconstructed frames is 46.89 dB. It is showed that the reconstructed frames are very much exact to the detected frames and hence, good performance of the block compressive sensing.



(a) Sample video 1 **(b) Sample video 2** **(c) Sample video 3**

Fig. 2: PSNR versus various block size (8×8, 16×16, 32×32 and 64×64) for (a) Sample video 1 (a walk by shop), (b) Sample video 2 (enter exit crossing path) and (c) Sample video 3 (two persons leave the shop)

V. Conclusion

This paper presents the application of the block compressive sensing framework for videos that can recover the videos with high-quality performance. Gaussian Mixture Model is used to get the detected object from the video frame. Due to the block-by-block processing mechanism, the algorithm has less complexities at each level. Our simulation results demonstrated that the quality of the reconstructed frames can be improved by higher subrate values with the large dimension of sensing matrix.

a) frame 16 **b) frame 18** **c) frame 20** **d) frame 24** **e) frame 27**

Fig. 3: Detected object frames

a) frame 16 **b) frame 18** **c) frame 20** **d) frame 24** **e) frame 27**

Fig. 4: Reconstructed object frames
(Subrate = 0.99, block size = 64×64, PSNR = 46.89 dB)

References

[1] R. Baraniuk, “Compressive sensing [lecture notes],” IEEE Signal Processing Magazine, vol. 24, no. 4, p. 118121, 2007.
[2] E. Candes and M. Wakin, “An introduction to compressive sampling,” IEEE Signal Processing Magazine, vol. 25, no. 2, p. 2130, 2008.
[3] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, “Compressive sensing: From theory to applications, a survey,” *Journal of Communications and Networks*, vol. 15, no. 5, p. 443456, 2013.
[4] T. V. Chien, K. Q. Dinh, B. Jeon, and M. Burger, “Block compressive sensing of image and video with nonlocal lagrangian multiplier and patch-based sparse representation,” *Signal Processing: Image Communication*, vol. 54, p. 93106, 2017.
[5] P.-M. Jodoin, “Comparative study of background subtraction algorithms,” *Journal of Electronic Imaging*, vol. 19, no. 3, p. 033003, Jan 2010.
[6] M. Yazdi, M. A. Bagherzadeh, M. Jokar, and M. A. Abasi, “Block-wise background subtraction based on gaussian mixture models,” *Applied Mechanics and Materials*, vol. 490-491, p. 12211227, 2014.
[7] A. Nurhadiyatna, W. Jatmiko, B. Hardjono, A. Wibisono, I. Sina, and P. Mursanto, “Background subtraction using gaussian mixture modelenhanced by hole filling algorithm (gmmhf),” *2013 IEEE InternationalConference on Systems, Man, and Cybernetics*, 2013.
[8] L. Gan, “Block compressed sensing of natural images,” *2007 15thInternational Conference on Digital Signal Processing*, 2007.
[9] S. Fayed, S. M.youssef, A. El-Helw, M. Patwary, and M. Moniri, “Adaptive compressive sensing for target tracking within wireless visual sensor networks-based surveillance applications,” *Multimedia Tools and Applications*, vol. 75, no. 11, p. 63476371, Jul 2015.
[10] “Caviar test case scenarios,” https://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/.

Divya R. Bora "Block Compressed Sensing of Videos: A Case Study." IOSR Journal of Computer Engineering (IOSR-JCE) 20.3 (2018): 49-53.