

Improving Video Streams Summarization Using Synthetic Noisy Video Data

Nada Jasim Al-Musawi
Computer Science, University of Babylon
Baghdad, Iraq

Saad Talib Hasson
Computer Science, University of Babylon
Baghdad, Iraq

Abstract—For monitoring public domains, surveillance camera systems are used. Reviewing and processing any subsequences from large amount of raw video streams is time and space consuming. Many efficient approaches of video summarization were proposed to reduce the amount of irrelevant information. Most of these approaches do not take into consideration the illumination or lighting changes that cause noise in video sequences. In this work, video summarization algorithm for video streams has been proposed using Histogram of Oriented Gradient and Correlation coefficients techniques. This algorithm has been applied on the proposed multi-model dataset which is created by combining the original data and the dynamic synthetic data. This dynamic data is proposed using Random Number Generator function. Experiments on this dataset showed the effectiveness of the proposed algorithm compared with traditional dataset.

Keywords—Video summarization; Histogram of Oriented Gradient (HOG); Correlation coefficients (R); key frames; illumination changes; noise; Random Numbers Generator function

I. INTRODUCTION

Public places may contain many of stationary cameras (such as banks, transport, airports, etc.) for security requirements. This data has rich information; it should be analyzed in order to get the useful information. Processing and storing this huge data is very difficult. It is very important to summarize this data in order to facilitate many tasks such as data mining tasks. Summarization task is a basic key in data mining. Summarization techniques find a compact description of dataset transforming it to a smaller and suitable form for stream data analysis with the maximum information content [1], [2].

Video summary can take two forms: a static summary [3], [4], which is a set of selected key-frames, or a dynamic video which is a short video constructed by concatenating short video segments [5]. Several efficient video summarization approaches have been proposed for surveillance video stream such as [6]-[9]. For real-time, generally video summarization approaches utilized motion object detection and extraction as essential process to extract motion information from video sequence [10]-[12]. Many of summarization approaches generated a video synopsis or summary for a single video stream such as [13]-[15]. Summarization approaches for single camera do not give generalization to multiple cameras and they do not take into account the relationship between the different cameras. Thus, some recent approaches have been proposed to handle the problem of multiple stationary cameras to produce a video summary for related scenes.

Xu et al. (2015) in [16] developed a new video summarization framework using clustering techniques that produces a video summary for multiple videos observing the same scene by computing a shared activity among all scenes. Gygli et al. (2015) in [17] proposed a new dynamic video skimming in which a supervised approach was used in order to learn the useful global information of a summary. The result of the proposed method is an optimal video summary that maintain the diversity of the original video. Kuanar et al. (2015) proposed a video summarization approach using a graph theoretic method. The steps can be summarized: Shot boundary detection is achieved depending on Bag of Visual Words and the global feature such as color, texture and shape to remove the redundant frames. The video summary is constructed using Gaussian entropy algorithm [18]. Sigari et al. (2015) proposed a fast video summarization using an on-demand feature extraction and a fuzzy inference system. Based on an on-demand feature extraction, the input video is partitioning to highlight and analyze video content. Each highlight is assigned a score using a Fuzzy Inference System. The score value indicates the importance of the events occurred in the highlight [19]. Bian et al., (2015) proposed a video event summarization method which comprised three stages: (1) noise removal. (2) Discovering sub events from multiple data types. (3) Generating visualized summary from the microblog streams of multiple data types [38]. Fu et al. proposed a summarization technique using the spatio-temporal shot graph, then the shot graph is divided and clusters of event-centered shots with similar contents are constructed. The video summary is produced by solving a multi-objective optimization problem by shot importance evaluated using a Gaussian entropy fusion technique [39]. Zheng et al. (2015) proposed a novel surveillance videos summarization. The motion feature is extracted using graphics processing units GPU to reduce running time. Then, the result of this step is smoothed to reduce noise, and finally, the video summary is created by selecting frames with local maxima of motion information [40].

Most of these approaches for video summarization do not take into account temporal noise which occurs in scenes under illumination changes or light changes. Noise is a common problem in digital cameras due to some of errors may occur in one of two sensor cameras, or ambiguity in some of the sensor data that is exposed to noise. The video summarization that based on motion detection with noisy video produces wrong motion object vectors. For real world applications, the feature extraction in computer vision and image processing should be robust to brightness or illumination changes or to frame

distortion such as noise or blur. The illumination or brightness changes of some points between consecutive frames in video frames sequence often occur due to variations in parameters of different video cameras, or moving of objects from one part to another part of the scene can be changed with different illuminations [20], [21]. Presence of these issues will cause an inaccurate processing of the video stream. Most of the video summarization methods for a single static camera or multi-camera video do not take into account to illumination changes or to the existing noise signals which are occurred in some of video frames. They depended on the assumption that noise or illumination values are static along video frames.

In this research, multi-sensor video summarization algorithm has been proposed based on Histogram of Oriented Gradient (HOG) procedure which is used as feature extraction and robust similarity or dissimilarity measure which is Correlation Coefficients approach. Unlike some video summarization approaches in the literature, the proposed algorithm framework is not operated directly on the raw pixels. The algorithm uses the feature vector for each frame in video sequence in order to improve the motion detection accuracy under illumination variance and shadowing. Thus, (HOG) is selected for this purpose.

The availability of real or representative data is an important issue for evaluating data analysis algorithms. Because of some of real data are lack or difficult to obtain, synthetic data becomes alternative data. In many research areas such as data mining, image processing, computer vision, sensor networks, and artificial intelligence developed different synthetic data generation schemes for different applications [22], [23]. Corruption of data may come from noise or blur which sometimes comes from different atmospheric conditions. Many approaches used synthetic multi-temporal data generated by Gaussian noise in order to test and evaluate their proposed approaches such as [24]-[26]. These techniques produced the traditional Gaussian noise which is identically distributed noise. Meaning that, the noise values at all pixel locations in all sequenced data are generated from the probability density function with the static mean and the standard deviation values. In this work, the developed synthetic data generation has been proposed. The proposed synthetic data generation method generates sequenced frames using Random Number Generator (RNG) function in order to simulate the original video sequence containing variant noisy frames.

The paper rest is organized as follows: in Section II, background theories that are related to this work are presented. The proposed methodology is provided in Section III. The experimental results and performance assessment are provided in Section IV and V. Finally, Section VI draws conclusions of this work.

II. THEORY

A. Histogram of Oriented Gradient

Histogram of Oriented Gradient HOG is features extraction technique and the most successful used to extract low-level features for object detection and recognition. HOG, can be found in [27], and was originally designed for human

detection [28], [29]. It has low and computational time [30] and robust against shadow and illumination changes [21], [31]. The HOG algorithm for an image can be implemented by four main steps [33]: 1) gradient computation, 2) orientation binning 3) descriptor blocks, and 4) block normalization.

B. Correlation Coefficients

The *correlation coefficient* which is a measure of the calculating the score of relationship similarity between two variables x and y can be defined as [32]:

$$R = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \sqrt{\sum(y_i - \bar{y})^2}} \quad (1)$$

Where x_i and y_i are gray level values of i -th pixel in the first and second frames respectively and \bar{x} and \bar{y} are the means of gray level values of x and y . The values of R always fall in $[-1, +1]$. When R is near to 0 meaning that the relationship between the two variables is little and when R is near to 1 the relationship is greater.

C. Noise Modeling

Real environments are often exposed to unexpected situations which are considered as noise. Gaussian noise is the most natural type of noise which is normally distributed. In MATLAB, noisy signal corrupted by Gaussian noise can be obtained by using the following:

$$g(x, y) = f(x, y) + \sigma * randn(\text{size}(f)) + \mu,$$

or

$$(x, y) = f(x, y) + \text{sqr}t(\sigma^2) * randn(\text{size}(f)) + \mu \quad (2)$$

Where $g(x, y)$ is the signal with additive Gaussian noise; $f(x, y)$ is the original signal; σ is the standard deviation; σ^2 is the variance; and μ is the mean. $randn()$ is MATLAB function for generating random numbers with a Gaussian normal distribution. The probability density function for a Gaussian distribution with mean μ and variance σ^2 can be defined as:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3)$$

D. Performance Evaluation Metrics

The performance evaluation of the proposed video summarization algorithm has been achieved using three types of metrics: (1) Data compression ratio (DCR) which is the ratio between number of frames in the original video and number of frames in the summary video [34], (2) Space savings ($S_{\text{saving}} = 1 - (\frac{1}{DCR})$) [35], and (3) Condensed Ratio (CR) which is the ratio between number of frames in the summary video and number of frames in the original video [36].

III. METHODOLOGY

A. The Proposed Method for Synthetic Data Generation

The standard deviation or the variance is the power of Gaussian noise signal. The classical Gaussian noise generator has the same approximated linear power for all given video frames. In dynamic environment, noise features are usually changed over time, it is necessary to use a simulation technique in order to adjust noise features. In this paper, the developed algorithm for generating synthetic video sequence

uses MATLAB random number generator function, the output of this function is added in (2) instead of the global σ values among frames. The new proposed equation can be defined as:

$$g(x, y) = f(x, y) + randn(d) * randn(size(f)) + \mu \tag{4}$$

Where d ($d=1$) for a single generated random number, $g(x, y)$ is noisy frame signal, $f(x, y)$ is the original frame.

Fig. 1 illustrates an example the visual comparison between the classical synthetic noisy data and the proposed synthetic noisy data.

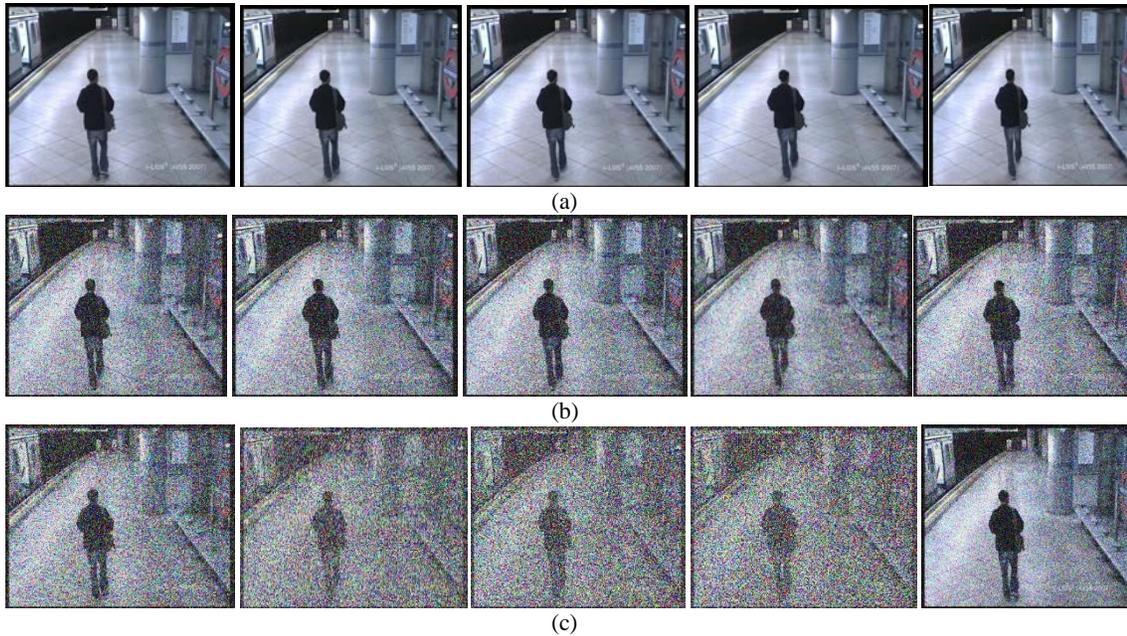


Fig. 1. Comparison between the classical synthetic noisy data and the proposed synthetic noisy data (a) The original video sequence, (b) The traditional noisy video sequence, (c) The proposed noisy video sequence

At each iteration, $randn(1)$ generates a new single random number which is considered as the value of standard deviation. The output of the algorithm is noisy video containing frames with non-linear or dynamic noise components. Mean Square Error (MSE) is used as an error measure between the original video signals and the noisy video signals. Table 1 and Fig. 2 show the comparison between Mean MSE's of the classical synthetic noisy data and MSE's of the proposed synthetic noisy data for ten sequenced frames.

TABLE I. COMPARISON BETWEEN MSE'S OF THE CLASSICAL SYNTHETIC NOISY DATA AND MSE'S OF THE PROPOSED SYNTHETIC NOISY DATA

MSE for traditional noisy sequenced frames	MSE for developed noisy sequenced frames
8	21.86
7.94	23.31
7.92	10.61
7.87	18.83
7.78	21.77
7.76	17.91
7.73	14.3
7.88	6.29
7.78	23.42
7.76	22.9

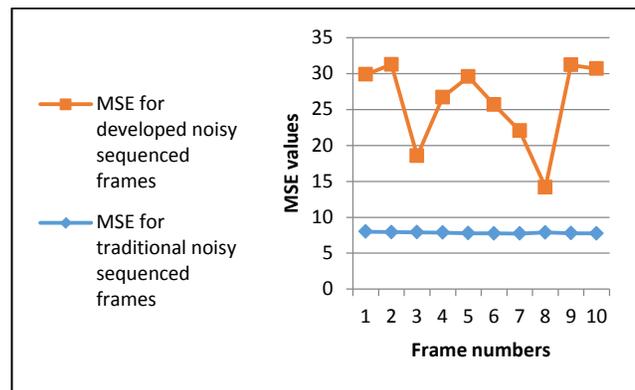


Fig. 2. Graph of comparison between MSE's of the classical synthetic noisy data and MSE's of the proposed synthetic noisy data

From Fig. 2, notice that the MSE's of the proposed synthetic video have nonlinear model, while MSE's of the proposed synthetic noisy video have approximately linear model. Multi-model dataset is constructed by using the original video dataset and the corresponding synthetic corrupted video dataset. In this research, this multi-model dataset is used for testing and evaluating the proposed summarization approach.

B. The Proposed Multi-Model Video Summarization

Multi-model video summarization using Histogram Oriented Gradient (HOG) algorithm for features extraction and Correlation Coefficient as a measure that quantifies the dependency (independency) between two video sequences. This algorithm has been applied on the proposed multi-model video dataset (for two videos). The videos have equal length (n frames). To increase efficiency of the proposed algorithm, HOG is used, which is a faster process and has low small features space. HOG procedure is computed for the current frame f_i ($i = 1..n$) from vid_1 (original video or reference video) and the current frame g_i ($i = 1..n$) from vid_2 (the corresponding noisy video of the original video). The results are two feature vectors. Correlation Coefficients (R) is computed between these feature vectors. Algorithm (1) illustrates the proposed HOG-Correlation Coefficients algorithm.

Algorithm (1) : HOG-Correlation based Summarization

Input: Vid_1 is the original video sequence

Vid_2 is the noisy video sequence

n is the number of video frames.

Output: video summary

Steps:

1. Given two video streams, Vid_1 and Vid_2 .
2. For $i = 1$ to n
3. Read the current frame (f_i) from the video stream (Vid_1).
4. Read the current frame (g_i) from the video stream (Vid_2).
5. Compute $H_1 = HOG(f_i)$ and $H_2 = HOG(g_i)$
6. Compute correlation coefficient $R = \text{correlation}(H_1 \text{ and } H_2)$
7. If $R \geq 0.9$ then
8. Store the current original frame into summary file.
9. else go to step 2 (disregard the current frame)
10. end

To detect redundancy (noisy) frames, we put a constraint on the correlation coefficients that are computed from the step 3. If $R \geq 0.9$ then the current frame from the original video sequence is stored into a summary file, otherwise go to step 2 (the current frames is dropped), and then continuously detect redundant frames for all frames.

IV. EXPERIMENTAL RESULTS

The original samples video datasets for metro and road scenes were selected from [37]. Using these videos, the synthetic data has been generated. The original and the synthetic video were used to construct the multi-model dataset for testing and evaluating the proposed summarization method.

The results analysis of the comparison between the classical synthetic video generation algorithm and the proposed synthetic video generation algorithm can be illustrated as follows:

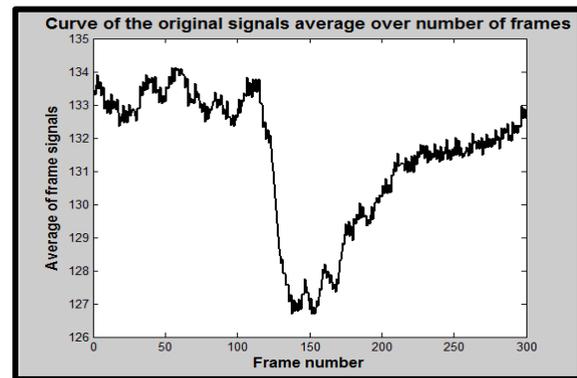
1) The random noise is in normal distribution case $N(0,1)$, $\text{mean}(\mu) = 0$ and $\text{variance}(\sigma^2) = 1$, the result contains high noise values (Mean Square Error (MSE) = $6.4784e+03$).

2) The random noise is in general distribution case $N(u, \sigma^2)$, the result contains very high noise values.

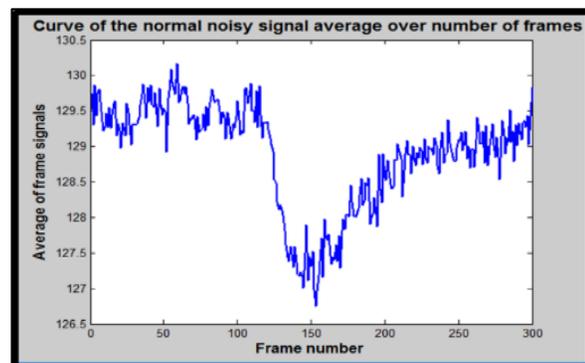
3) μ and σ^2 have values in the range $[0, 1]$, the result will contain lower noise values. The noise values close to zero when σ^2 is near or equal to zero and the result will become similar to the original signal ($\sigma^2 = 0.1$, $MSE = 923.0849$ and $\sigma^2 = 0.1$, $MSE = 321.7665$).

4) Applying the proposed algorithm on the input video sequence is achieved, the result contains high noise components, but these noise components are variant among all frames. Randomly, some of the frames have high values and others have low values depending on the random values of the variance

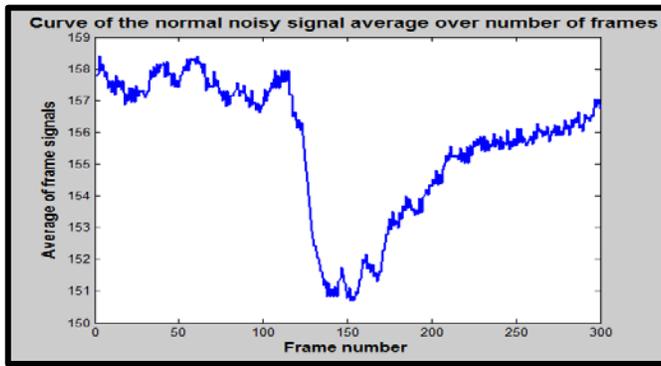
Fig. 3 and Fig. 4 show the curves of comparison between the original video (metro and road) signals and the corresponding synthetic noisy video with different noise levels. The x-axis represents the number of frames and the y-axis represent the average signals of each frame in the sequence.



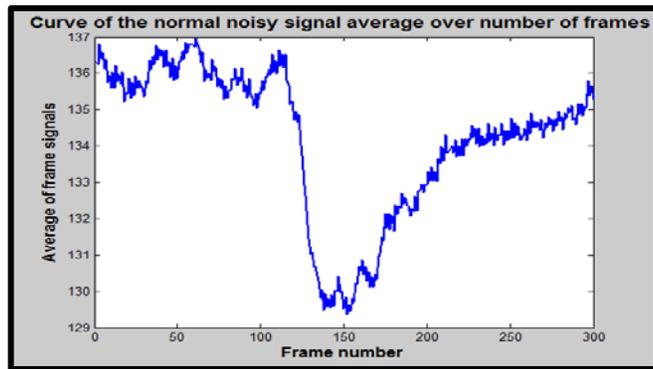
(a)



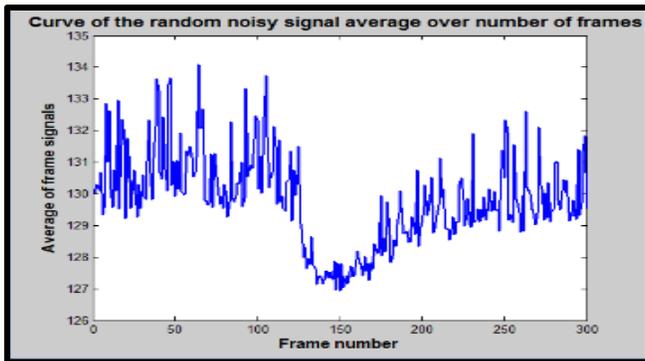
(b)



(c)

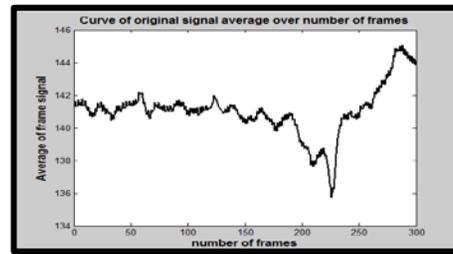


(d)

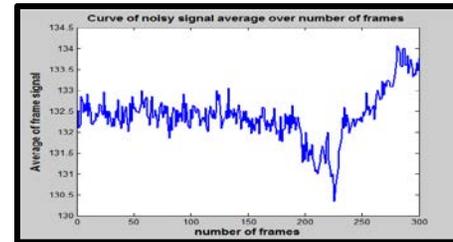


(e)

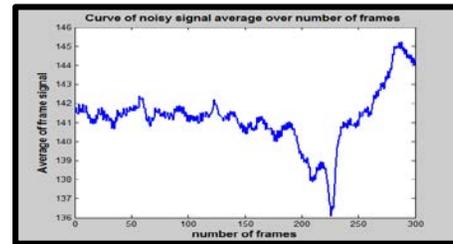
Fig. 3. Sequenced 300 frames from the metro video corrupted by Gaussian noise generated with $\mu = 0$ and different standard deviation values : (a) the original signals. (b) the noisy signals with $\sigma = 1$. (c) the noisy signals with $\sigma = 0.1$. (d) the noisy signals with $\sigma = 0.01$. (e) the noisy signals with $\sigma = randn(1)$



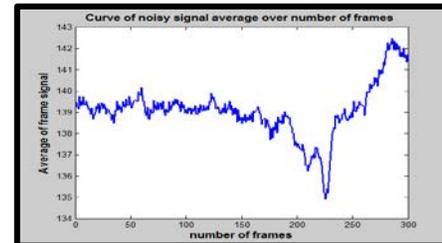
(a)



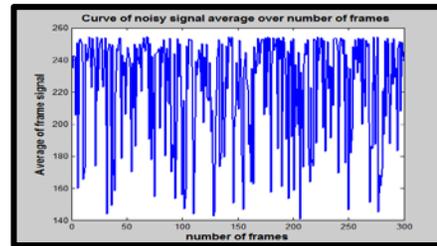
(b)



(c)



(d)



(e)

Fig. 4. Sequenced 300 frames from the road video corrupted by Gaussian noise generated with $\mu = 0$ and different standard deviation values : (a) the original signals. (b) the noisy signals with $\sigma = 1$. (c) the noisy signals with $\sigma = 0.1$. (d) the noisy signals with $\sigma = 0.01$. (e) the noisy signals with $\sigma = randn(1)$

In Fig. 3, and Fig. 4, when the noise values are high, the summarization of the video streams is not occurred because the correlation coefficients are smaller than the threshold ($R < 0.9$). When the noise values are small, all correlation coefficients will be greater than the threshold ($R > 0.9$). In this case, the summarization of the video stream is also not occurred and the output video has redundant frames like the redundant frames in the original video frames.

In order to generate a video sequence containing variant noisy pixels. The proposed algorithm for generating noisy video with dynamic and variant noise values has been applied to solve the above problem. The result of this algorithm is a noisy video which is used with the original (reference) video to form a multi-model dataset. The HOG-correlation based algorithm has been applied on this multi-model dataset. The correlation coefficients (R) have different values. the values that are greater than threshold ($R > 0.9$) are saved, otherwise are removed. The output is a video summary containing important information without redundant and noisy frames.

V. PERFORMANCE ASSESSMENT

The HOG-correlation algorithm for video stream sumrization have been tested on the video sequence (metro scene). The size of short sample of the original metro video is equal to 46.9 MB of 3001 frames. The Table 2 shows the information details before and after applying the HOG-correlation algorithm on the multi-model dataset (the original and the synthetic videos). As shown in Table2, V1 is the video output of HOG-correlation algorithm between the original metro video and the video generated using the traditional noise video generation algorithm, the size of V1 in the memory space is equal 44.9 MB because the original data which is 3D video is converted to 2D video. But the number of frames of V1 is still the same numbers of the original one.

V2 is the output of HOG-correlation algorithm between the original metro video and video generated using the proposed algorithm of dynamic noise generation with using random numbers. The size of V2 is equal to 1.09 MB and the number of frames is only 69 after applying the HOG-correlation based algorithm.

TABLE II. THE INFORMATION DETAILS (VIDEO SIZE AND NUMBER OF FRAMES) BEFORE AND AFTER APPLYING THE PROPOSED CORRELATION BASED METHOD ON THE ORIGINAL METRO VIDEO AND THE SYNTITHICS VIDEOS DATASETS

Metro video sequence	Video size befor	Video size after	Number of frames befor	Number of frames after
V1	101 MB	44.9 MB	3001	3001
V2	200 MB	1.09 MB	3001	69

From the experimental results, the HOG-correlation algorithm gives better results when it is applied on the original video and the video generated by the developed method of dynamic noise generation based on generated random numbers. As a result, it has ability to reduce the size of input video sequence and extract important motion information.

Depending on above information, the performance of HOG-correlation algorithm has been tested. Table3 ilustraites the performance evaluation of the proposed algorithm using

three metrics: Data compression ratio (DCR), Space savings (S_{saving}), and Condensed Ratio (CR).

TABLE III. THE RESULTS OF THE PERFORMANCE EVALUATION OF CORRELATION BASED ALGORITHM FOR METRO VIDEO SEQUENCE

Meto video sequence	DCR	S_{saving}	CR
V1	2.2494	55.54%	0%
V2	183.4862	99.45%	97.70%

As demonstrated in Table3 for the metro video sequence, DCR which is the ratio between the incompact size and compact size gives very good results for V2 against the value of DCR for V1. The S_{saving} , the compact in size relative to the incompact size, for V2 gives better results against the value of S_{saving} for V1. The CR, the ratio between the number of output frames and the number of input frames, also gives very good results (97.70%) for V2 against the value of CR for V1.

From the above performance evaluation results, the HOG-correlation algorithm works better on the noisy video that is ctreated using random numbers.

VI. CONCLUSION AND FUTURE WORK

In real time applications such as surveillance applications, illumination changes or shadowing for motion objects may occur in surveillance video stream. Many video summarization methods that are trying to construct video summary depended on the assumption that noise or illumination values are static along video frames. This leading to the existing video summarization algorithms will stumble in understanding of scene under observation. In this paper, the synthetic noisy video generation algorithm has been developed for testing the proposed video summarization algorithm based on Histogram Oriented Gradient and Correlation Coefficient for multi-model video dataset. The experimental results on the proposed dataset showed good results compared with the classical dataset. For more efficient time and space computation, online video summarization using Histogram of Oriented Gradient and Correlation coefficients techniques will be generated as a future work.

REFERENCES

- [1] Chandola, Varun, and Vipin Kumar. "Summarization-compressing data into an informative representation." Springer, Knowledge and Information Systems 12.3 (2007): 355-378.
- [2] Chandola, Varun, and Vipin Kumar. "Summarization-compressing data into an informative representation." *Knowledge and Information Systems* 12.3 (2007): 355-378.
- [3] Cayllahua-Cahuina, E. J. Y., G. Cámara-Chávez, and D. Menotti. "A static video summarization approach with automatic shot detection using color histograms." Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), 2012.
- [4] Mahmoud, Karim M., Mohamed A. Ismail, and Nagia M. Ghanem. "Vscan: an enhanced video summarization using density-based spatial clustering." *Image Analysis and Processing-ICIAP 2013*. Springer Berlin Heidelberg, 2013. 733-742.
- [5] Li, Y., Meriardo, B., Rouvier, M., & Linares, G."Static and dynamic video summaries." Proceedings of the 19th ACM international conference on Multimedia. ACM, 2011.

- [6] Li, Xiang-Wei, et al Li, X. W., Zhang, M. X., Zhao, S. P., & Zhu, Y. L. I. "A Novel Dynamic Video Summarization Approach Based on Rough Sets in Compressed Domain." (2009).
- [7] Raikwar, Suresh Chandra, Charul Bhatnagar, and Anand Singh Jalal. "A Novel Framework for Efficient Extraction of Meaningful Key Frames from Surveillance Video." *International Journal of System Dynamics Applications (IJSDA)* 4.2 (2015): 56-73.
- [8] Zheng, R., Yao, C., Jin, H., Zhu, L., Zhang, Q., & Deng, W. "Parallel key frame extraction for surveillance video service in a smart city." *PLoS one* 10.8 (2015): e0135694.
- [9] Mei, Shaohui, et al. "Video summarization via minimum sparse reconstruction." *Pattern Recognition* 48.2 (2015): 522-533.
- [10] Damjanovic, U., Fernandez, V., Izquierdo, E., & Martinez, J. M "Event detection and clustering for surveillance video summarization." *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on*. IEEE, 2008.
- [11] Ji, Zhong, Su, Y., Qian, R., & Ma, J. "Surveillance video summarization based on moving object detection and trajectory extraction." *Signal Processing Systems (ICSPS), 2010 2nd International Conference on*. Vol. 2. IEEE, 2010.
- [12] Dammak, Majdi, and Adel M. Alimi. "Viewers Affective Feedback for Video Summarization." *Journal of information processing systems* 11.1 (2015): 76-94.
- [13] Peleg, Shmuel, Pritch, Y., Rav-Acha, A., & Gutman, A. "Method and system for video indexing and video synopsis." U.S. Patent No. 8,818,038. 26 Aug. 2014.
- [14] Potapov, Danila, Douze, M., Harchaoui, Z., & Schmid, C. "Category-specific video summarization." *Computer Vision-ECCV 2014*. Springer International Publishing, 2014. 540-555.
- [15] Pranjal, Pratyush, Lakhdive, R., Vasave, A., Varma, A., & Barve, A. "An Effective Method of Video Segmentation and Summarization for Surveillance." *International Journal of Computer Applications* 95.2 (2014).
- [16] Xu, Xun, Timothy Hospedales, and Shaogang Gong. "Discovery of Shared Semantic Spaces for Multi-Scene Video Query and Summarization." arXiv preprint arXiv:1507.07458 (2015).
- [17] Gygli, Michael, and Helmut Grabner, Luc Van Gool. "Video Summarization by Learning Submodular Mixtures of Objectives." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.
- [18] Kuanar, Sanjay, Kunal Ranga, and Ananda Chowdhury. "Multi-View Video Summarization using Bipartite Matching Constrained Optimum-Path Forest Clustering." (2015).
- [19] Sigari, Mohamad-Hoseyn, Hamid Soltanian-Zadeh, and Hamid-Reza Pourreza. "Fast Highlight Detection and Scoring for Broadcast Soccer Video Summarization using On-Demand Feature Extraction and Fuzzy Inference." *International Journal of Computer Graphics* 6.1 (2015).
- [20] Tang, Feng, Lim, S. H., Chang, N. L., & Tao, H. "A novel feature descriptor invariant to complex brightness changes." *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009.
- [21] Rashwan, Hatem A., Mohamed, M. A., García, M. A., Mertsching, B., & Puig, D. "Illumination robust optical flow model based on histogram of oriented gradients." *Pattern recognition*. Springer Berlin Heidelberg, 2013. 354-363.
- [22] Gabrijel, Ivan, and Andrej Dobnikar. "On-line Inference of Finite Automata in Noisy Environments." Springer Vienna, 2005.
- [23] Pei, Yaling, and Osmar Zaiane. "A synthetic data generator for clustering and outlier analysis." Department of Computing science, University of Alberta, edmonton, AB, Canada (2006).
- [24] Rezatofighi, Seyed Hamid, et al. "A framework for generating realistic synthetic sequences of total internal reflection fluorescence microscopy images." *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*. IEEE, 2013.
- [25] Su, Xin, et al. "Two-step multitemporal nonlocal means for synthetic aperture radar images." *Geoscience and Remote Sensing, IEEE Transactions on* 52.10 (2014): 6181-6196.
- [26] Rozantsev, Artem, Vincent Lepetit, and Pascal Fua. "On rendering synthetic images for training an object detector." *Computer Vision and Image Understanding* (2015).
- [27] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR*, 2005.
- [28] Fleyeh, Hasan, and Janina Roch. "Benchmark Evaluation of HOG Descriptors as Features for Classification of Traffic Signs." (2013).
- [29] Ma, Yingdong, Xiankai Chen, and George Chen. "Pedestrian detection and tracking using HOG and oriented-LBP features." *Network and Parallel Computing*. Springer Berlin Heidelberg, 2011. 176-184.
- [30] Li, Changyan, Lijun Guo, and Yichen Hu. "A new method combining HOG and Kalman filter for video-based human detection and tracking." *Image and Signal Processing (CISP), 2010 3rd International Congress on*. Vol. 1. IEEE, 2010.
- [31] Pedersoli, M., González, J., Chakraborty, B., & Villanueva, J. J. "Enhancing real-time human detection based on histograms of oriented gradients." *Computer Recognition Systems 2*. Springer Berlin Heidelberg, 2007. 739-746.
- [32] Roy, Anuradha. "Estimating correlation coefficient between two variables with repeated observations using mixed effects model." *Biometrical journal* 48.2 (2006): 286-301.
- [33] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE, 2005.
- [34] Davydov, Alexander Y. "Signal and Image Processing with Sinlets." *arXiv preprint arXiv:1206.0692* (2012).
- [35] Dutch, Mike. "Understanding data deduplication ratios." *SNIA Data Management Forum*. 2008.
- [36] Zhu, J., Feng, S., Yi, D., Liao, S., Lei, Z., & Li, S. Z. "High Performance Video Condensation System". *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 25, No. 7, 2015.
- [37] 2007 IEEE International Conference on Advanced Video and Signal based Surveillance, London (United Kingdom), 2007.
- [38] Bian, Jingwen, et al. "Multimedia Summarization for Social Events in Microblog Stream." *Multimedia, IEEE Transactions on* 17.2 (2015): 216-228.
- [39] Fu, Yanwei, et al. "Multi-view video summarization." *Multimedia, IEEE Transactions on* 12.7 (2010): 717-729.
- [40] Zheng, Ran, et al. "Parallel key frame extraction for surveillance video service in a smart city." *PLoS one* 10.8 (2015): e0135694.